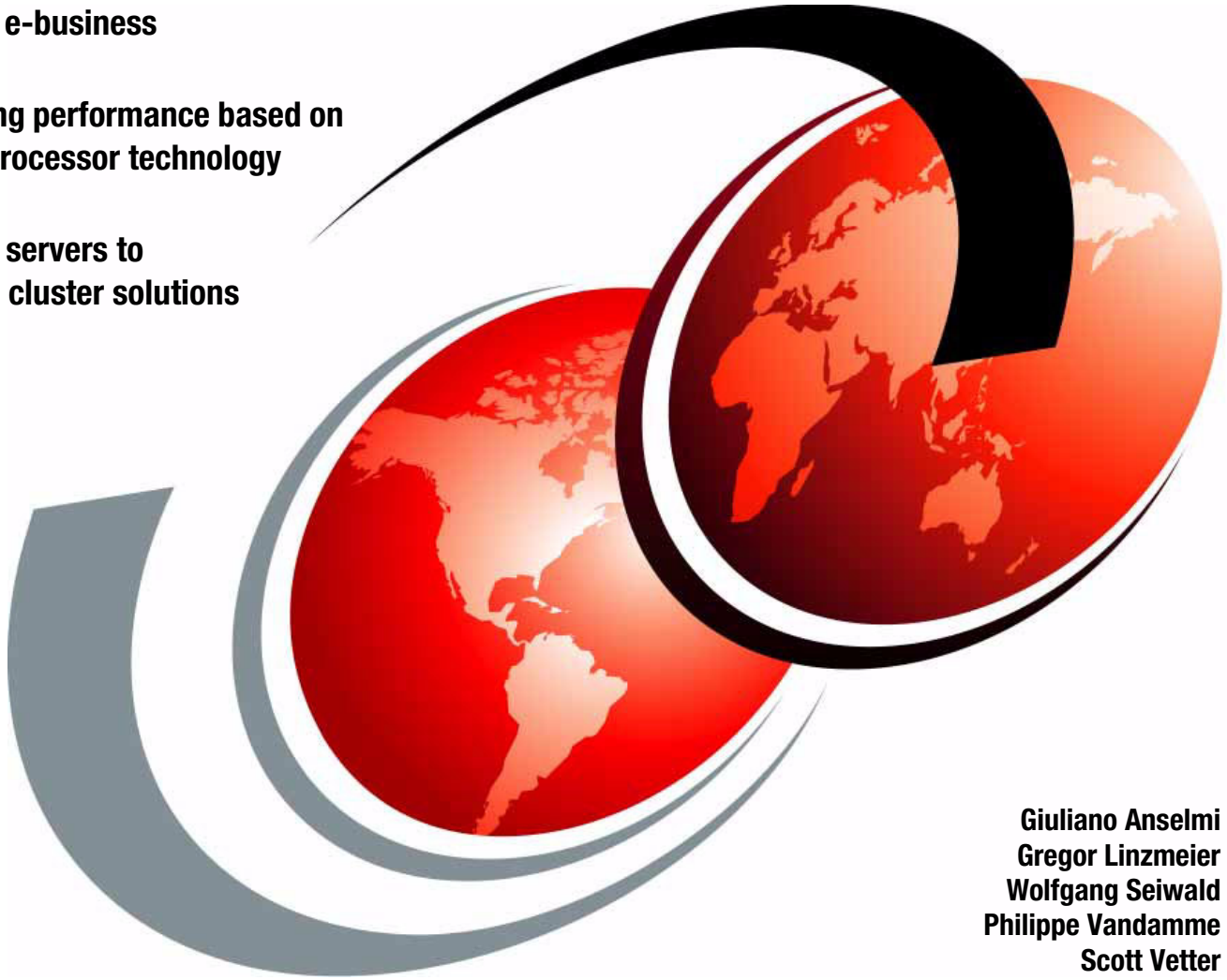
IBM

IBM @server OpenPower 720 Technical Overview and Introduction

- Deskside and rack-mount server for managing e-business
- Outstanding performance based on POWER5 processor technology
- From Web servers to integrated cluster solutions



Giuliano Anselmi
Gregor Linzmeier
Wolfgang Seiwald
Philippe Vandamme
Scott Vetter

Redpaper

ibm.com/redbooks

International Technical Support Organization

IBM @server OpenPower 720 Technical Overview and Introduction

October 2004

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

Second Edition (October 2004)

This edition applies to IBM @server OpenPower 720.

Contents

Notices

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

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

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

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

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

Any references in this information to non-IBM 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.

Trademarks

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

Advanced Micro-Partitioning™
AIX®
AIX 5L™
Chipkill™
Electronic Service Agent™
Enterprise Storage Server®
@server®
@server®
HACMP™
i5/OS™
IBM®
Micro-Partitioning™
OpenPower™
POWER™
POWER4™
POWER4+™
POWER5™
PowerPC®
pSeries®
Redbooks™
Redbooks (logo) ™
RS/6000®
Service Director™
TotalStorage®

The following terms are trademarks of other companies:

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

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

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

Preface

This document is a comprehensive guide covering the IBM @server® OpenPower 720 Linux servers. We introduce major hardware offerings and discuss their prominent functions.

Professionals wishing to acquire a better understanding of IBM @server OpenPower products should 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 documentation by providing a desktop reference that offers a detailed technical description of the OpenPower 720 system.

This publication does not replace the latest IBM @server marketing materials and tools. It is intended as an additional source of information that, together with existing sources, can 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.

Giuliano Anselmi is a certified pSeries Presales Technical Support Specialist working in the Field Technical Sales Support group based in Rome, Italy. For seven years, he was an IBM @server pSeries Systems Product Engineer, supporting Web Server Sales Organization in EMEA, IBM Sales, IBM Business Partners, Technical Support Organizations, and IBM Dublin eServer Manufacturing. Giuliano has worked for IBM for 12 years, devoting himself to RS/6000® and pSeries systems with his in-depth knowledge of the related hardware and solutions.

Gregor Linzmeier is an IBM Advisory IT Specialist for RS/6000 and pSeries workstation and entry servers as part of the Systems and Technology Group in Mainz, Germany 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.

Wolfgang Seiwald is an IBM Presales Technical Support Specialist working for the System Sales Organization in Salzburg, Austria. He holds a Diplomingenieur degree in Telematik from the Technical University of Graz. The main focus of his work for IBM in the past five years has been in the areas of the IBM @server pSeries systems and the IBM AIX operating system.

Philippe Vandamme is an IT Specialist working in pSeries Field Technical Support in Paris, France, EMEA West region. With 15 years of experience in semi-conductor fabrication and manufacturing and associated technologies, he is now in charge of pSeries Pre-Sales Support. In his daily role, he supports and delivers training to the IBM and Business Partner Sales force.

The project that produced this document was managed by:

Scott Vetter
IBM U.S.

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

Ron Arroyo, Barb Hewitt, Thoi Nguyen, Jan Palmer, Charlie Reeves, Craig Shempert, Scott Smylie, Joel Tendler, Ed Toutant, Jane Arbeitman, Tenley Jackson, Andy McLaughlin, Kathy Bennett, Raymond Harney, David Boutcher.
IBM U.S.

Volker Haug
IBM Germany

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 905 Internal Zip 9053D004
  11501 Burnet Road
  Austin, Texas 78758-3493

1

General description

The IBM @server OpenPower 720 deskside and rack-mount servers are designed for greater application flexibility, with innovative technology, designed to help you become an on demand business. With POWER5 microprocessor technology, the OpenPower 720 is the a high-performance entry Linux server that, as an orderable option, includes the next development of the IBM partitioning concept, Micro-Partitioning.

Through the optional POWER Hypervisor feature (FC 1965), LPAR on a 4-way OpenPower 720 (9124-720) allows up to four dedicated partitions. In addition, the POWER Hypervisor feature enables the support up to 40 Micro-Partitions on a 4-way system. Micro-Partitioning technology enables multiple partitions to share a physical processor. The POWER Hypervisor controls dispatching the physical processors to each of the partitions. In addition to Micro-Partitioning and Virtual Ethernet provided by the POWER Hypervisor, the optional Virtual I/O Server feature (5724-K81) allows sharing of physical network adapters and enables the virtualization of SCSI storage. Together, these two features are known as Advanced OpenPower Virtualization.

To dynamically manage the resources on a system, and to configure and use Virtual Ethernet, Shared Ethernet adapters, and Virtual SCSI, an Hardware Management Console is required.

Simultaneous multi-threading (SMT) is a standard feature of POWER5 technology that allows two threads to be executed at the same time on a single processor. SMT is selectable with dedicated or processors from a shared pool using Micro-Partitioning technology.

The symmetric multiprocessor (SMP) OpenPower 720 system features base 1-way or 2-way (1.5 GHz), 2-way or 4-way (1.65 GHz), 64-bit, copper-based POWER5 microprocessors with 36 MB off-chip Level 3 cache configurations standard on 2-way and 4-way models. Main memory starting at 512 MB can be expanded up to 64 GB, based on the available DIMMs, for higher performance and exploitation of 64-bit addressing to meet the demands of enterprise computing, such as large database applications.

Included with the OpenPower 720 are five hot-plug PCI-X slots with Enhanced Error Handling (EEH), one embedded Ultra320 SCSI dual-channel controller, one 10/100/1000 Mbps integrated dual-port Ethernet controller, two serial ports, two USB 2.0 capable ports, two HMC ports, two RIO-2 ports, and two System Power Control Network (SPCN) ports.

The OpenPower 720 includes four front-accessible, hot-swap-capable disk bays in a minimum configuration with an additional four hot-swap-capable disk bays orderable as an optional feature. The eight disk bays can accommodate up to 1.17 TB of disk storage using the 146.8 GB Ultra320 SCSI disk drives. Three non hot-swapable media bays are used to accommodate additional devices. Two media bays only accept slim line media devices, such as DVD-ROM or DVD-RAM, and one half-height bay is used for a tape drive. The OpenPower 720 also has I/O expansion capability using the RIO-2 bus, which allows attachment of the 7311 Model D20 I/O drawers.

Additional reliability and availability features include redundant hot-plug cooling fans and redundant power supplies. Along with these hot-plug components, the OpenPower 720 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.

1.1 System specifications

Table 1-1 lists the general system specifications of the OpenPower 720 system.

Table 1-1 OpenPower 720 specifications

Description	Range
Operating temperature	5 to 35 degrees Celsius (41 to 95 F)
Relative humidity	8% to 80%
Operating voltage	1-2 way: 100 to 127 or 200 to 240 V AC (auto-ranging) 4-way: 200 to 240 V AC
Operating frequency	47/63 Hz
Maximum power consumption	1100 watts maximum
MAximum thermal output	3754 Btu[a]/hour (maximum)

a. British Thermal Unit

1.2 Physical package

The following sections discuss the major physical attributes found on an OpenPower 720 in rack-mounted and deskside versions (Figure 1-1 on page 4), as shown in Table 1-2. The OpenPower 720 is a 4U, 19-inch rack-mounted system or deskside system depending on the feature code selected.

Table 1-2 Physical packaging of the OpenPower 720

Dimension	Rack (FC 7914)	Deskside (FC 7912)
Height	178 mm (7.0 inches)	533 mm (21.1 inches)
Width	437 mm (17.2 inches)	201 mm (7.9 inches)
Depth	731 mm (28.8 inches)	779 mm (30.7 inches)
Minimum configuration	41.4 kg (91 pounds)	
Maximum configuration	57.0 kg (125 pounds)	

Figure 1-1 Rack-mount and deskside versions of the OpenPower 720

1.2.1 IBM eServer OpenPower 720 deskside

The OpenPower 720 is available as a deskside server that is ideal for environments requiring the user to have local access to the hardware. A typical example of this would be applications requiring a native graphics display.

To order an OpenPower 720 system as a deskside version, FC 7912 is required. 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 or Linux 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), the IBM T541H 15-inch TFT Color Monitor (FC 3637), or others.

1.2.2 IBM eServer OpenPower 720 rack-mounted

The OpenPower 720 is available as 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 OpenPower 720, it is possible to place up to 10 systems in an area of 644 mm (25.5 inches) x 1147 mm (45.2 inches).

To order an OpenPower 720 system as a rack-mounted version, FC 7914 is required.The OpenPower 720 can be installed in IBM or OEM racks. Therefore you are required to select one of the following features:

- IBM Rack-mount Drawer Rail Kit (FC 7162)
- OEM Rack-mount Drawer Rail Kit (FC 7163)

Included with the OpenPower 720 rack-mounted server packaging are all of the components and instructions necessary to enable installation in a 19-inch rack using suitable tools.

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 Linux starts the X11-based graphical user interface. Graphical system tools are usable for configuration management if the adapter is connected to a common maintenance console, such as the 7316-TF3 Rack-Mounted Flat-Panel display.

Figure 1-2 shows the basic ports available on the OpenPower 720.

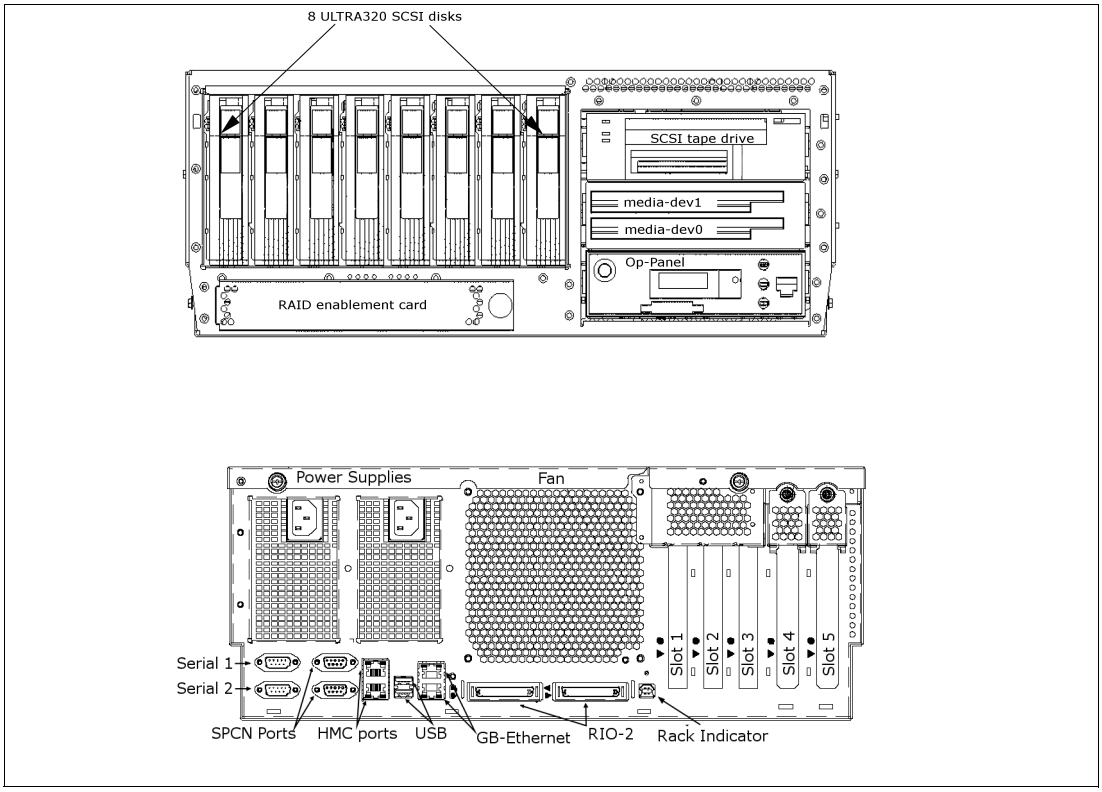


Figure 1-2 Detailed views of the OpenPower 720 rack-mount system

1.3 Minimum and optional features

The OpenPower 720 system is based on a flexible, modular design, and it features:

- Up to two processor books using the POWER5 chip, for a 1-way, 2-way, or 4-way configurations
- From 512 MB to 64 GB of total system memory capacity using DDR1 DIMM technology
- Four SCSI disk drives in a minimum configuration, eight SCSI disk drives with an optional second 4-pack enclosure for an internal storage capacity of 1.17 TB using 146.8 GB drives
- Four PCI-X slots and 1 GX+ slot or five PCI-X slots
- Two slim-line media bays for optional storage devices
- One half-high bay for an optional tape device

The OpenPower 720, including the service processor (SP) described in 2.10.1, “Service processor base” on page 34, supports the following native ports:

- Two 10/100/1000 Ethernet ports

- Two serial ports
- Two USB 2.0 ports
- Two HMC ports
- Two remote I/O (RIO-2) ports
- Two SPCN ports

In addition, the OpenPower 720 features one internal Ultra320 SCSI dual channel controller, redundant hot-swap power supply (optional), cooling fans, and up to two processor power regulators.

The system supports 32-bit and 64-bit applications.

1.3.1 Processor card features

The OpenPower 720 accommodates 1 and 2-way processor cards with state-of-the-art, 64-bit, copper-based, POWER5 microprocessors running at 1.5 GHz on 1-way, 1.5 GHz 2-way, and 1.65 GHz 2-way cards that share 1.9 MB of L2 on chip cache and eight slots for memory DIMMS using DDR1 technology. 36 MB of L3 cache is available on the 2-way cards. Capacity on Demand (CoD) is not available on the OpenPower 720. For a list of available cards; see Table 1-3.

An initial order must have at least one processor card. Only a single 1-way card may be present on any system, and cards with mixed clock rates cannot be installed on the same system.

Table 1-3 Processor card and feature codes

Processor card FC	Description
1960	1-way 1.5 GHz, no L3 cache, eight DDR1 DIMM sockets
1961	2-way 1.5 GHz, 36 MB L3 cache, eight DDR1 DIMM sockets
5262	2-way 1.65 GHz, 36 MB L3 cache, eight DDR1 DIMM sockets

1.3.2 Memory features

The processor cards used in the OpenPower 720 system have eight sockets for memory DIMMs. Each of the memory features available (at the time of writing) listed in Table 1-4, therefore, provides four DIMMs, except for FC 1936.

Table 1-4 Memory feature codes

Feature code	Description
1936	0.5 GB (2x 256 MB) DIMMs, 250 MHz, DDR1 SDRAM
1937	1 GB (4x 256 MB) DIMMs, 250 MHz, DDR1 SDRAM
1938	2 GB (4x 512 MB), DIMMs, 250 MHz, DDR1 SDRAM
1940	4 GB (4x 1024 MB) DIMMs, 250 MHz, DDR1 SDRAM
1942	8 GB (4x 2048 MB) DIMMs, 250 MHz, DDR1 SDRAM
1945	16 GB (4x 4096 MB) DIMMs, 250 MHz, DDR1 SDRAM

It is recommended that each processor card (if more than a single card is present) have an equal amount of memory installed. Balancing memory across the installed processor cards

allows distributed memory accesses that provide optimal performance. The memory controller will detect a variety of memory configurations of mixed memory sized DIMMs and DIMMs installed in pairs.

1.3.3 Disk and media features

The minimum OpenPower 720 configuration includes a 4-pack disk drive enclosure. A second 4-pack disk drive enclosure can be installed by ordering FC 6592 (integrated controller driven) or FC 6593 (additional adapter driven); therefore, the maximum internal storage capacity is 1.17 TB (using the disk drive features available at the time of writing). The OpenPower 720 also features two slim-line media device bays and one half-height media bay. The minimum configuration requires at least one disk drive. Table 1-5 shows the disk drive feature codes that each bay can contain.

Table 1-5 Disk drive feature code description

Feature code	Description
3273	36.4 GB 10 K RPM Ultra3 SCSI disk drive assembly
3277	36.4 GB 15 K RPM Ultra3 SCSI disk drive assembly
3274	73.4 GB 10 K RPM Ultra3 SCSI disk drive assembly
3278	73.4 GB 15 K RPM Ultra3 SCSI disk drive assembly
3275	146.8 GB 10 K RPM Ultra3 SCSI disk drive assembly

Any combination of DVD-ROM and DVD-RAM drives of the following devices can be installed in the two slim-line bays:

- DVD-RAM drive, FC 5751
- DVD-ROM drive, FC 2640

A logical partition running a supported release of the Linux operating system requires a DVD-ROM drive or DVD-RAM drive to provide a way to boot hardware diagnostics from CD.

Supplementary devices can be installed in the half-height media bay, such as:

- IBM 80/160 GB Internal Tape Drive with VXA Technology, FC 6120
- 60/150 GB 16-bit 8 mm Internal Tape Drive, FC 6134
- 36/72 GB 4 mm Internal Tape Drive, FC 6258

1.3.4 USB diskette drive

For today's administration tasks, an internal diskette drive is not state-of-the-art. In some situations, the external USB 1.44 MB diskette drive for OpenPower 720 systems (FC 2591) is helpful. This super-slim-line and lightweight USB V2 attached diskette drive takes its power requirements from the USB port. A USB cable is provided. The drive can be attached to the integrated USB ports, or to a USB adapter (FC 2738). A maximum of one USB diskette drive is supported per integrated controller/adapter. The same controller can share a USB mouse and keyboard.

1.3.5 I/O drawers

The OpenPower 720 has five internal PCI-X slots, where four of them are long slots and one is a short slot. If more PCI-X slots are needed, especially well-suited to extend the number of

LPARs and Micro-Partitions, up to eight Model 7311 Model D20 I/O drawers can be attached to the rack-mount OpenPower system. Up to four Model D20 drawers can be connected to the two RIO-2 ports on the rear of the system that are provided in a minimum configuration. An additional four Model D20s can be connected by ordering the Remote I/O expansion card (FC 1806). It provides two RIO-2 ports located on a interposer card that occupies the short PCI-X slot.

7311 Model D20 I/O drawer

The 7311 Model D20 I/O drawer is a 4U full-size drawer, which must be mounted in a rack. It features seven hot-pluggable PCI-X slots and optionally up to 12 hot-swappable disks arranged in two 6-packs. Redundant, concurrently maintainable power and cooling is an optional feature (FC 6268). The 7311 Model D20 I/O drawer offers a modular growth path for the OpenPower 720 system with increasing I/O requirements. When a OpenPower 720 is fully configured with eight attached 7311 Model D20 drawers, the combined system supports up to 61 PCI-X adapters (in a full configuration, Remote I/O expansion cards are required) and 104 hot-swappable SCSI disks.

PCI-X and PCI cards are inserted from the top of the I/O drawer down into the slot. The installed adapters are protected by plastic separators, designed to prevent grounding and damage when adding or removing adapters.

The drawer has the following attributes:

- 4U rack-mount enclosure assembly
- Seven PCI-X slots 3.3 volt, keyed, 133 MHz hot-pluggable
- Two 6-pack hot-swappable SCSI devices
- Optional redundant hot-plug power
- Two RIO-2 ports and two SPCN ports

Note: The 7311 Model D20 I/O drawer initial order, or an existing 7311 Model D20 I/O drawer that is migrated from another system, must have the RIO-2 ports available (FC 6417).

7311 Model D20 I/O drawer physical package

The I/O drawer has the following physical characteristics:

- Width: 482 mm (19.0 inches)
- Depth: 610 mm (24.0 inches)
- Height: 178 mm (7.0 inches)
- Weight: 45.9 kg (101 pounds)

Figure 1-3 on page 9 shows the different views of the 7311-D20 I/O drawer.

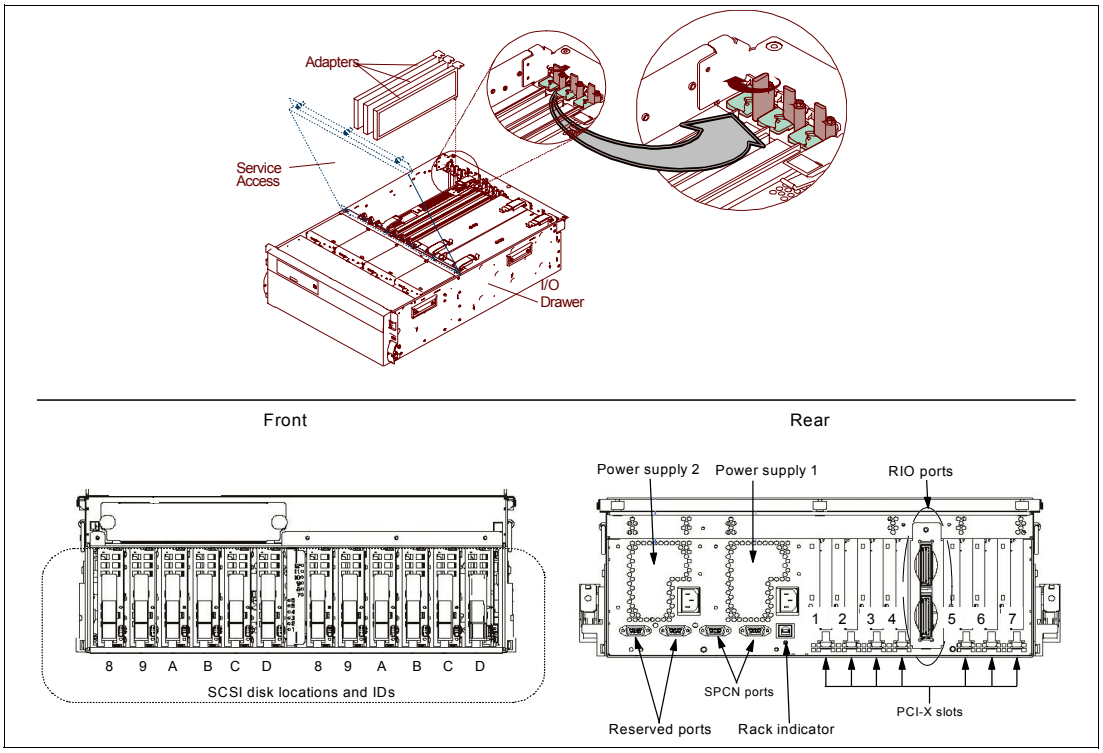


Figure 1-3 7311-D20 I/O drawer

Note: The 7311 Model D20 I/O drawer is designed to be installed by an IBM service representative.

I/O drawers and usable PCI slots

Only the 7311 Model D20 I/O drawer is supported on an OpenPower 720 system. Depending on the system configuration, the maximum number of I/O drawers supported is different. Table 1-6 summarizes the maximum number of I/O drawers supported and the total number of PCI-X slots available.

Table 1-6 Maximum number of I/O drawers supported and total number of PCI slots

OpenPower 720	Max number of I/O drawers	Total number of PCI-X slots
Minimum configuration	4	28 + 5
Additional FC 1806	8	56 + 4 (excluding used slot)

1.3.6 Hardware Management Console models

The Hardware Management Console (HMC) provides a set of functions that is necessary to manage the OpenPower system when dynamic LPAR, Micro-Paritioning, Virtual Ethernet, Virtual SCSI, inventory and remote microcode management, and remote power control functions are needed. These functions include the handling of the partition profiles that define the processor, memory, and I/O resources allocated to an individual partition. The HMC is required when the Virtual I/O Server or the POWER Hypervisor features are ordered.

The 7310 Model CR2 or the 7310 Model C03 HMCs are specifically for POWER5 processor-based systems. However, an existing 7315 Model CR2 and the 7315 Model C03 (POWER4 processor-based systems HMC) can be converted for POWER5 processor-based

systems use when it is loaded with the HMC software required for POWER5 processor-based systems (FC 0961).

POWER5 processor-based system HMCs require Ethernet connectivity. Ensure that sufficient Ethernet adapters are available to enable public and private networks if you need both. The 7310 Model C03 is a desktop model with only one native 10/100/1000 Ethernet port, but three additional PCI slots. The 7310 Model CR2 is a 1U, 19-inch rack mountable drawer that has two native Ethernet ports and two additional PCI slots.

When an HMC is connected to the OpenPower 720, the integrated serial ports are disabled. If you need serial connections, you need to provide an async adapter.

Note: It is not possible to connect POWER4 and POWER5 processor-based systems simultaneously to the same HMC.

1.4 System racks

The Enterprise Rack Models T00 and T42 are 19-inch wide racks for general use with IBM @server OpenPower, p5, pSeries, and RS/6000 rack-based or rack drawer-based systems. The racks provide increased capacity, greater flexibility, and improved floor space utilization.

The OpenPower 720 uses a 4U rack-mounted server drawer.

If an OpenPower 720 system is to be installed in a non-IBM rack or cabinet, you should ensure that the rack conforms to the EIA[1] standard EIA-310-D (see 1.4.6, “OEM rack” on page 13).

Note: 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.

1.4.1 IBM RS/6000 7014 Model T00 Enterprise Rack

The 1.8-meter (71 inch) Model T00 is compatible with past and present OpenPower, p5, 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.
- Standard black or optional white color in OEM format.
- Increased power distribution and weight capacity.
- An optional reinforced (ruggedized) rack feature (FC 6080) provides added earthquake protection with modular rear brace, concrete floor bolt-down hardware, and bolt-in steel front filler panels.

[1] Electronic Industries Alliance (EIA). Accredited by American National Standards Institute (ANSI), EIA provides a forum for industry to develop standards and publications throughout the electronics and high-tech industries.

- Support for both AC and DC configurations.
- DC rack height is increased to 1926 mm (75.8 inches) if a power distribution panel is fixed to the top of the rack.
- Up to four Power Distribution Units (PDUs) can be mounted in the proper bays, but others can fit inside the rack. See 1.4.3, “AC Power Distribution Unit and rack content” on page 11.
- Optional rack status beacon (FC 4690). This beacon is designed to be placed on top of a rack and cabled to servers, such as a OpenPower 720, and other components, such as a 7311 I/O drawer, inside the rack. Servers can be programmed to illuminate the beacon in response to a detected problem or changes in system status.
- A rack status beacon junction box (FC 4693) should be used to connect multiple servers and I/O drawers 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).
- Weight:

  T00 base empty rack: 244 kg (535 pounds)

  T00 full rack: 816 kg (1795 pounds)

1.4.2 IBM RS/6000 7014 Model T42 Enterprise Rack

The 2.0-meter (79.3-inch) 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 pounds)
  - T42 full rack: 930 kg (2045 pounds)

1.4.3 AC Power Distribution Unit and rack content

For rack models T00 and T42 9-outlet PDUs are available.

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.

The OpenPower 720 can be connected to any PDU that is available for the 7014-T00 or 7014-T42 racks.

For detailed power cords requirements and power cord feature codes, see the publication *Site and Hardware Planning Information*, SA38-0508. An online copy can be found at:

`http://publib16.boulder.ibm.com/pseries/en_US/infocenter/base/`

The first four PDUs ordered for a rack will be mounted vertically in the sides of the rack, occupying all the four PDU bays available. Any additional PDUs will be mounted horizontally in the rear of the rack and will occupy 1U of rack space.

Note: Each OpenPower 720, or a system drawer to be mounted in the rack, requires two power cords, which are not included in the base system order.

The Universal PDU (FC 7188) and the optional Universal PDU to be mounted horizontally in the rack (FC 9188) will be available on December 31, 2004, supporting a wide range of country requirements and electrical power specifications. Each Universal PDU provides 12 C13 power outlets for use within a 7014-T00 or 7014-T42 rack, compared to nine C13 power outlets provided by FC 7176 or FC 7177 PDUs. Nine different power cord features can be used to connect the PDU to a wall power outlet. Each power cord provides the unique design characteristics for the different power requirements. To match new power requirements, these power cords can be requested with an initial order of the rack, or with a later upgrade of the rack features.

1.4.4 Rack-mounting rules for OpenPower 720

The primary rules that should be followed when mounting the OpenPower 720 into a rack are:

- The system 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 an OpenPower 720 into the service position, it is essential that the rack manufacturer's safety instructions have been followed regarding rack stability.

Depending on the current implementation and future enhancements of additional 7311 Model D20 drawers connected to the OpenPower 720 system, Table 1-7 shows examples of the minimum and maximum configurations for different combinations of servers and attached 7311 Model D20 I/O drawers.

Table 1-7 Minimum and maximum configurations for 720s and 7311-D20s

	Only 720s	**One 720, one 7311-D20**	**One 720, four 7311-D20s**	**One 720, eight 7311-D20s**
7014-T00 rack	9	4	1	1
7014-T42 rack	10	5	2	1

1.4.5 Additional options for rack

The intention of this section is to highlight some solutions available to provide a single point of management for environments composed of multiple OpenPower 720 servers or other IBM ℮server OpenPower, p5, pSeries, and RS/6000 systems.

IBM 7212 Model 102 IBM TotalStorage Storage device enclosure

The IBM 7212 Model 102 is designed to provide efficient and convenient storage expansion capabilities for select IBM ℮server OpenPower, p5, pSeries, and RS/6000 servers. The IBM 7212 Model 102 is a 1U rack-mountable option to be installed in a standard 19-inch rack using an optional rack-mount hardware feature kit. The 7212 Model 102 has two bays that can accommodate any of the following storage drive features:

- Digital Data Storage (DDS) Gen 5 DAT72 Tape Drive provides physical storage capacity of 36 GB (72 GB with 2:1 compression) per data cartridge.

- VXA-2 Tape Drive provides a media capacity of up to 80 GB (160 GB with 2:1 compression) physical data storage capacity per cartridge.
- DDS-4 tape drive with 20 GB native data capacity per tape cartridge and a native physical data transfer rate of up to 3 MB/sec, using 2:1 compression so that a single tape cartridge can store up to 40 GB of data.
- DVD-ROM drive is a 5 1/4-inch, half-high device. It can read 640 MB CD-ROM and 4.7 GB DVD-RAM media. It can be used for Alternate IPL (IBM-distributed CD-ROM media only) and program distribution.
- DVD-RAM drive with up to 2.7 MB/sec throughput. Using 3:1 compression, a single disk can store up to 28 GB of data. Supported DVD disk native capacities on a single DVD-RAM disk are as follows: up to 2.6 GB, 4.7 GB, 5.2 GB, and 9.4 GB.

Flat panel display options

The IBM 7316-TF3 Flat Panel Console Kit may be installed in the system rack. This 1U 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 picture elements (pels) resolution. The 7316-TF3 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 integrated into the keyboard).

Note: It is recommended that you have the 7316-TF3 installed between EIA 20 to 25 of the rack for ease of use. The 7316-TF3 or any other graphics monitor requires a POWER GXT135P graphics accelerator (FC 2848 or FC 2849) to be installed in the server, or other graphic accelerator, if supported.

Hardware Management Console 7310 Model CR2

The 7310 Model CR2 is a 1U, 19-inch rack mountable drawer supported in the 7014 Model T00 and T42 racks. The 7310 Model CR2 provides one serial port, two integrated Ethernet ports, and two additional PCI slots. The HMC 7310 Model CR2 has USB ports to connect USB keyboard and mouse devices.

Note: The HMC serial port can be used for external modem attachment if the Service Agent call-home function is implemented, and the Ethernet ports are used to communicate to the service processor in OpenPower 720 systems. An Ethernet cable (FC 7801 or 7802) is required to attach the HMC to the OpenPower 720 system it controls.

1.4.6 OEM rack

The OpenPower 720 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.

The key points mentioned in this standard are as follows:

- Any rack used must be capable of supporting 15.9 kg (35 pounds) per EIA unit (44.5 mm [1.75 inches] of rack height).

- To ensure proper rail alignment, the rack must have mounting flanges that are at least 494 mm (19.45 inches) across the width of the rack and 719 mm (28.3 inches) between the front and rear rack flanges.
- It might be necessary to supply additional hardware, such as fasteners, for use in some manufacturer's racks.

Figure 1-4 shows the drawing specifications for OEM racks.

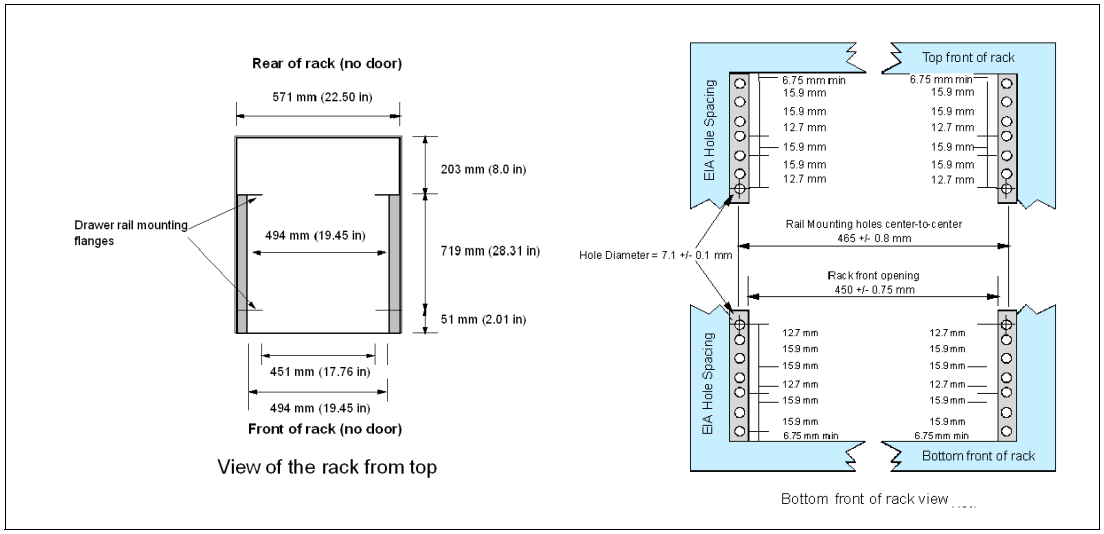


Figure 1-4 Reference drawing for OEM rack specifications

2

Architecture and technical overview

This chapter discusses the overall system architecture represented by Figure 2-1. The major components of this diagram are 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.

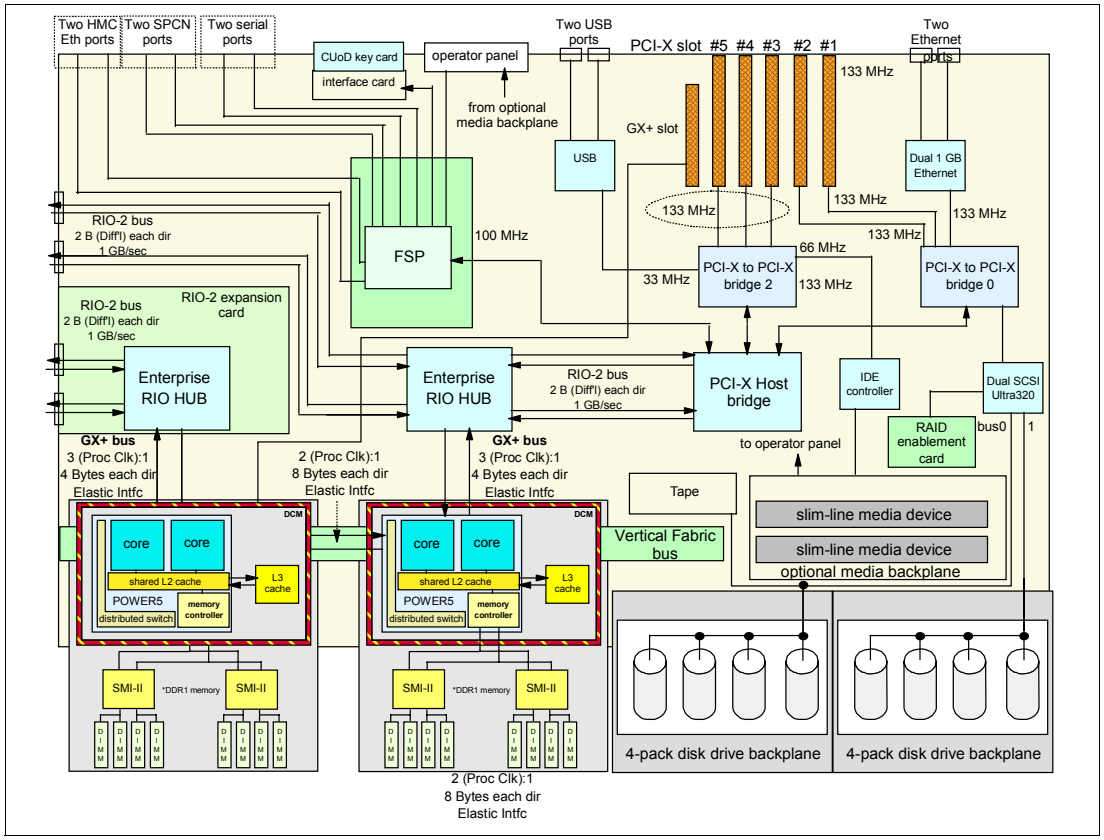


Figure 2-1 OpenPower 720 logic data flow

2.1 The POWER5 chip

The POWER5 chip features single-threaded and multi-threaded execution, providing higher performance in the single-threaded mode than its POWER4 predecessor at equivalent frequencies provides. POWER5 maintains both binary and architectural compatibility with existing POWER4 systems to ensure that binaries continue executing properly and all application optimizations carry forward to newer systems. The POWER5 provides additional enhancements such as virtualization, reliability, availability, and serviceability at both chip and system levels, and it has been designed to support speeds up to 3 GHz.

Figure 2-2 shows the high-level structures of POWER4 and POWER5 processor-based systems. The POWER4 scales up to a 32-way symmetric multiprocessor. Going beyond 32 processors increases interprocessor communication, resulting in high traffic on the interconnection fabric bus. This contention negatively affects system scalability.

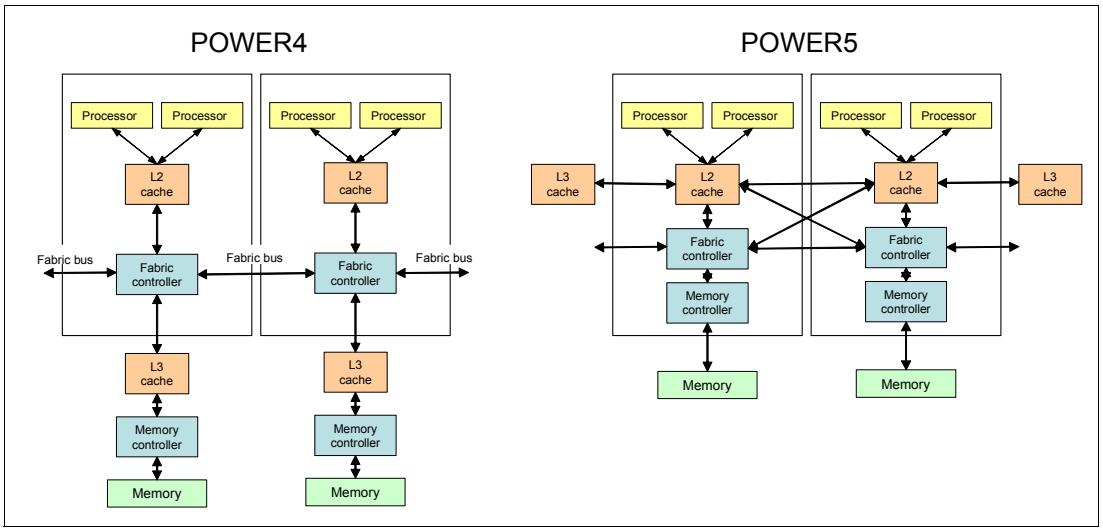


Figure 2-2 POWER4 and POWER5 system structures

Moving the L3 cache (2-way processor cards only) provides significantly more cache on the processor side than previously available, thus reducing traffic on the fabric bus and allowing POWER5 processor-based systems to scale to higher levels of symmetric multiprocessing. The POWER5 supports a 1.9 MB on-chip L2 cache, implemented as three identical slices with separate controllers for each. Either processor core can independently access each L2 controller. The L3 cache, with a capacity of 36 MB, operates as a backdoor with separate buses for reads and writes that operate at half processor speed.

Because of the higher transistor density of the POWER5 130 nm technology, it was possible to move the memory controller on chip and eliminate a chip previously needed for the memory controller function. These changes in the POWER5 also have the significant side benefits of reducing latency to the L3 cache and main memory and reducing the number of chips necessary to build a system.

The POWER5 processor supports the 64-bit PowerPC architecture. A single die contains two identical processor cores, each supporting two logical threads. Only a single core is active on a 1-way processor card. Activation of the second core is not possible.

This architecture makes the chip appear as a 4-way symmetric multiprocessor to the operating system. The POWER5 processor core has been designed to support both enhanced simultaneous multi-threading (SMT) and single threaded (ST) operation modes.

2.1.1 Simultaneous multi-threading

As a permanent requirement for performance improvements at the application level, simultaneous multi-threading (SMT) functionality is embedded in the POWER5 chip technology. Developers are familiar with process-level parallelism (multi-tasking) and thread-level parallelism (multi-threads). SMT is the next stage of processor saturation for throughput-oriented applications to introduce the method of instruction-level parallelism to support multiple pipelines to the processor.

By default, SMT is activated. On a 2-way POWER5 processor-based system, the operating system views the available processors as a 4-way system. To achieve a higher performance level, SMT is also applicable in Micro-Partitioning, capped or uncapped, and dedicated partition environments.

Simultaneous multi-threading is supported on POWER5 systems running the Linux operating system-based at a required 2.6 kernel. For Linux, an additional boot option must be set to activate SMT after a reboot.

The SMT mode maximizes the usage of the execution units. In the POWER5 chip, more rename registers have been introduced (for floating-point operation, rename registers increased to 120) that are essential for out of order execution and then vital for the SMT.

Enhanced SMT features

To improve SMT performance for various workload mixes and provide robust quality of service, POWER5 provides two features:

- Dynamic resource balancing
  - The objective of dynamic resource balancing is to ensure that the two threads executing on the same processor flow smoothly through the system.
  - Depending on the situation, the POWER5 processor resource balancing logic has different thread throttling mechanisms.
- Adjustable thread priority
  - Adjustable thread priority lets software determine when one thread should have a greater (or lesser) share of execution resources.
  - POWER5 supports eight software-controlled priority levels for each thread.

ST operation

Not all applications benefit from SMT. Having threads executing on the same processor will not increase the performance of applications with execution unit limited performance or applications that consume all the chip's memory bandwidth. For this reason, the POWER5 supports the ST execution mode. In this mode, the POWER5 processor gives all the physical resources to the active thread, allowing it to achieve higher performance than a POWER4 processor-based system at equivalent frequencies. Highly optimized scientific codes are one example where ST operation is ideal.

2.1.2 Dynamic power management

In current Complimentary Metal Oxide Semiconductor (CMOS) technologies, chip power is one of the most important design parameters. With the introduction of SMT, more instructions execute per cycle per processor core, thus increasing the core's and the chip's total switching power. To reduce switching power, POWER5 chips use a fine-grained, dynamic clock gating mechanism extensively. This mechanism gates off clocks to a local clock buffer if dynamic power management logic knows the set of latches driven by the buffer will not be used in the

next cycle. This allows substantial power saving with no performance impact. In every cycle, the dynamic power management logic determines whether a local clock buffer that drives a set of latches can be clock gated in the next cycle.

In addition to the switching power, leakage power has become a performance limiter. To reduce leakage power, the POWER5 chip uses transistors with low threshold voltage only in critical paths. The POWER5 chip also has a low-power mode, enabled when the system software instructs the hardware to execute both threads at the lowest available priority. In low power mode, instructions dispatch once every 32 cycles at most, further reducing switching power. The POWER5 chip uses this mode only when there is no ready task to run on either thread.

2.1.3 POWER chip evolution

The OpenPower system complies with the RS/6000 platform architecture, which is an evolution of the PowerPC Common Hardware Reference Platform (CHRP) specifications. Figure 2-3 shows the POWER evolution.

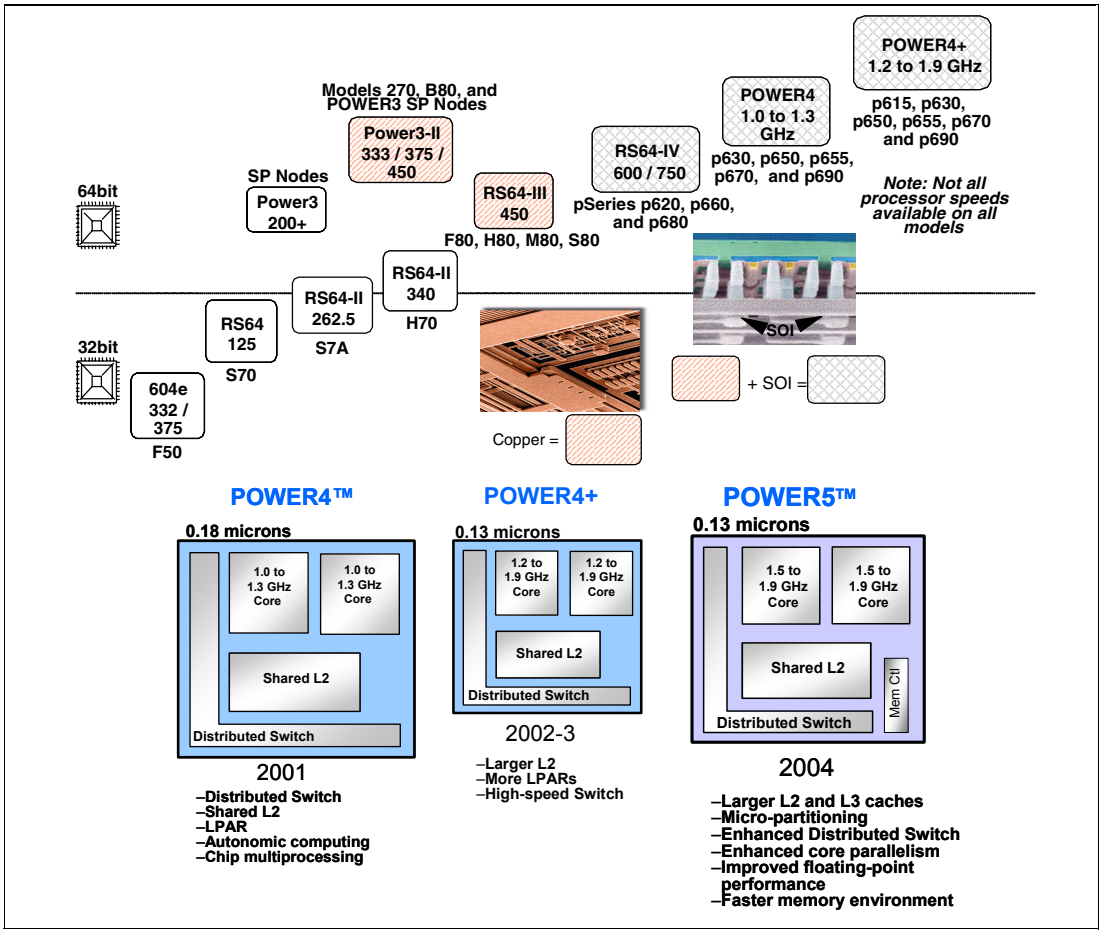


Figure 2-3 POWER chip evolution

2.1.4 CMOS, copper, and SOI technology

The POWER5 processor design is a result of a close collaboration between *IBM Systems and Technology Group* and *IBM Microelectronics Technologies* that enables IBM ℯserver OpenPower systems to give customers improved performance, reduced power

consumption, and decreased IT footprint size through logical partitioning. The POWER5 processor chip takes advantage of IBM leadership technology. It is made using IBM 0.13-µm-lithography CMOS technology. The POWER5 processor also uses copper and 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.

2.2 Processor cards

In the OpenPower system, the POWER5 chip has been packaged with the L3 cache chip (on 2-way cards) into a cost-effective Dual Chip Module (DCM) package. The storage structure for the POWER5 chip is a distributed memory architecture, which provides high memory bandwidth. Each processor can address all memory and sees a single shared memory resource. As such, a single DCM and its associated L3 cache and memory are packaged on a single processor card. Access to memory behind another processor is accomplished through the fabric buses. The OpenPower 720 supports up to two processor cards (each card is a 2-way) or a single 1-way card. Each 2-way processor card has a single DCM containing a POWER5 processor chip and a 36 MB L3 module. On all cards I/O connects to the Central Electronic Complex (CEC) subsystem through the GX+ bus. Each DCM provides a single GX+ bus for a total system capability of two GX+ buses. The GX+ bus provides an interface to a single device such as the RIO-2 buses, as shown in Figure 2-4.

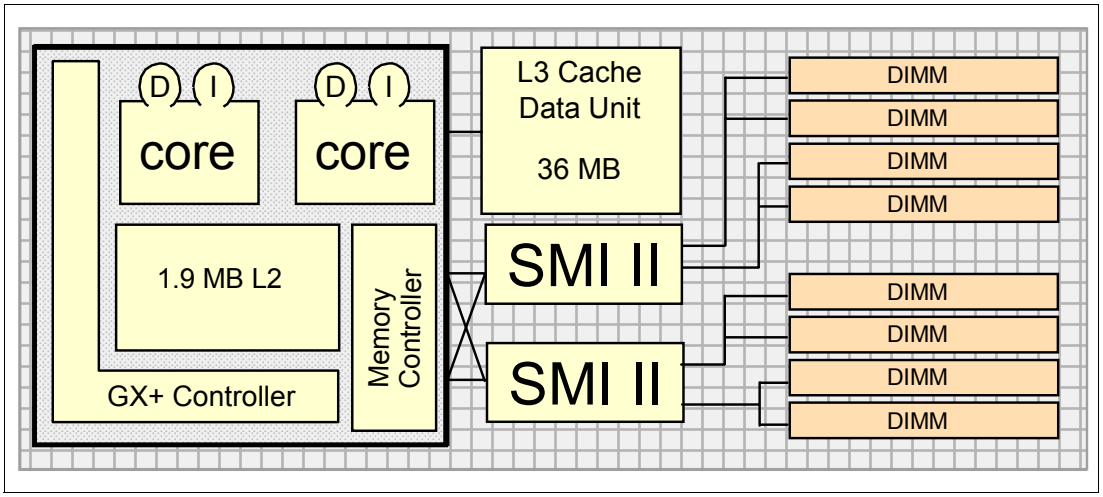


Figure 2-4 OpenPower 720 DCM diagram

Each processor card contains a single DCM and the local memory storage subsystem for that DCM. The processor card also contains LEDs for each FRU[1] on the CPU card including the CPU card itself. See Figure 2-5 on page 20 for a processor card layout view.

[1] Field Replacement Unit

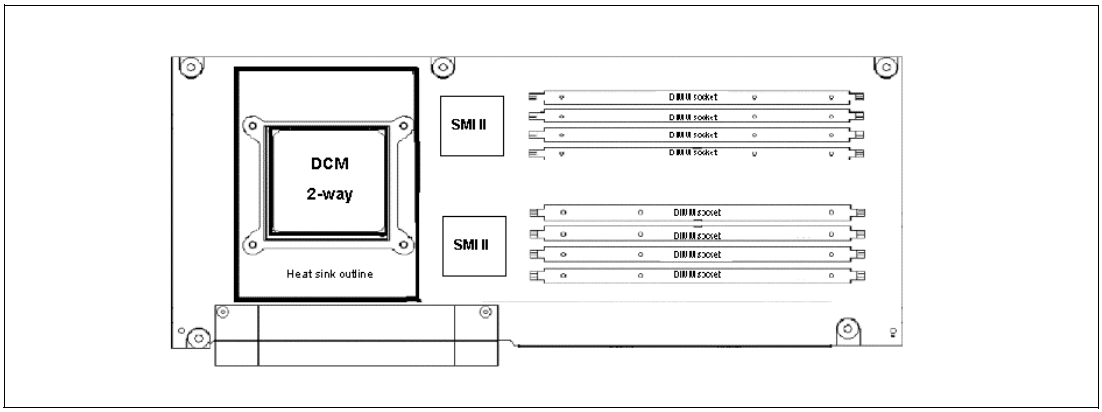


Figure 2-5 Processor card with DDR1 memory socket layout view

2.2.1 Available processor speeds

The OpenPower 720 operates with a processor clock rate of 1.65 GHz for 2 and 4-way POWER5 processor-based systems and 1.5 GHz for 1 and 2-way systems.

All processors installed on a single system must operate at the same clock speed.

2.3 Memory subsystem

The OpenPower 720 system offers pluggable DIMMs for memory. Each processor card contains eight slots for up to eight pluggable DIMMs. The minimum memory for an OpenPower 720 processor-based system is 0.5 GB and 64 GB as a maximum installable memory option. The memory capacity is dependant on the number of processors ordered for the system and on the type of memory and functionality required. Figure 2-6 on page 21 shows the offerings and memory slot availability.

2.3.1 Memory placement rules

The memory features available at the time of writing for the OpenPower 720 are listed in 1.3.2, “Memory features” on page 6.

Memory features that contain four DIMMs, or quad, and must be installed according to Figure 2-6 on page 21. The first quad slots are J0A, J1A, J0C, and J1C, and for the second quad, the slots are J0B, J1B, J0D, and J1D. The 0.5 GB memory feature must be installed in J0A and J1A.

Note: A quad must consist of a single feature (that is made of identical DIMMs). Mixed DIMM capacities in a quad will result in reduced RAS.

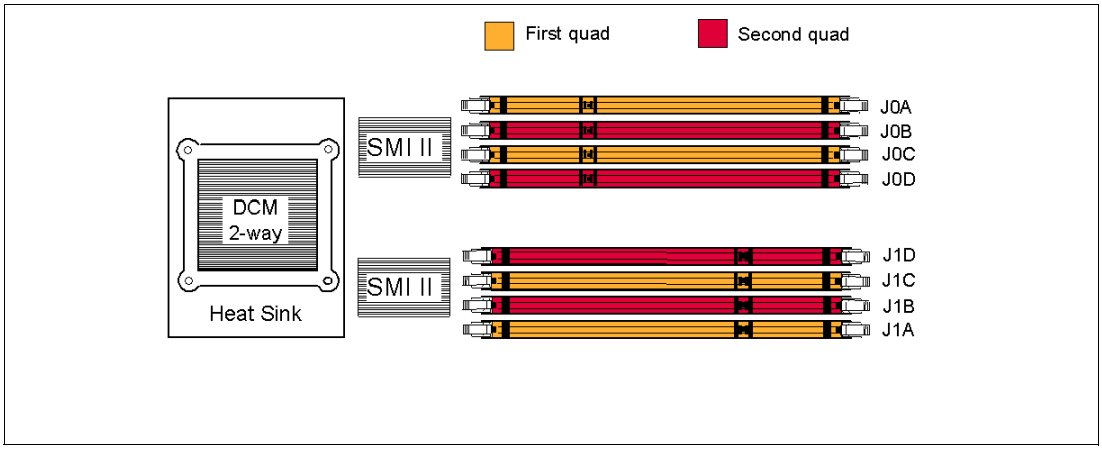


Figure 2-6 Memory placement for the OpenPower 720

2.3.2 Memory throughput

The memory subsystem throughput is based on the speed of the memory, not the speed of the processor. An elastic interface, contained in the POWER5 chip, buffers reads and writes to and from memory and the processor. There are two SMIs per processor card, each with a single 8 byte write and 2 byte read DDR bus to the processor. A DDR bus enables double read or writes per clock cycle. If 250 MHz memory is installed, the throughput is (16 x 2 x 250) + (4 x 2 x 250) or 9.76 GB/second per processor card. Therefore, a 4-way system would have a maximum theoretical throughput of 19.53 GB/second.

The POWER5 processor's integrated memory controller further reduces latency over the previous outboard controller on POWER4 systems to the SMI chips by requiring fewer cycles in order to set up memory addressing in the hardware.

2.3.3 Memory restrictions

The OpenPower does not officially support OEM memory, and there is no exception to this rule. OEM memory is never certified for the use in OpenPower servers. If the OpenPower 720 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.

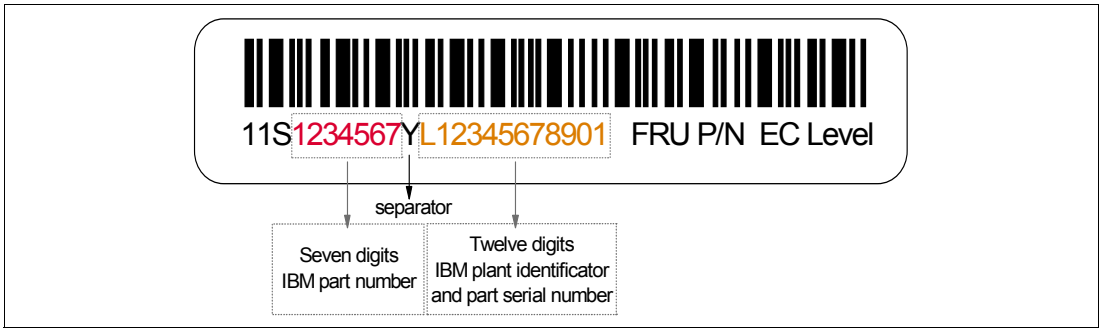


Figure 2-7 IBM memory certification label

Sometimes OEM vendors put a label reporting the IBM memory part number but not the barcode or the alphanumeric string, or both 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.

2.4 System buses

The following sections provide additional information related to the internal buses.

2.4.1 RIO buses and GX card

Each DCM provides a GX+ bus that is used to connect to an I/O subsystem. In a fully populated OpenPower 720, there are two GX+ buses, one from each processor card. Each system has one GX+ slot with a single GX+ bus. The GX+ slot is not active unless the second processor card is installed. The GX+ cards are hot-pluggable. The OpenPower 720 provides two external RIO-2 ports that can operate up to 1 GHz. An add-in GX adapter card (Remote I/O expansion card, FC 1806) when plugged into the GX+ slot adds two more RIO-2 ports. The RIO-2 ports are used for I/O expansion to external I/O drawers. The supported I/O drawer is 7311 Model D20.

2.5 Internal I/O subsystem

The internal I/O subsystem resides on the system planar, and the service processor (SP) is packaged on a separate service processor card. Each card is a separate FRU. There is an internal RIO-2 bus imbedded in the system planar. The system planar contains both the Enterprise RIO-2 hub and the PCI-X Host bridge chip to connect to the integrated I/O packaged on the system planar. Two RIO-2 ports of the Enterprise hub chip are used for the integrated I/O, and the remaining two ports are routed to external connectors.

The system planar provides five PCI-X slots and several integrated PCI devices that interface to two EADS-X chips that function as PCI-X to PCI-X bridges to the primary PCI-X buses on the PCI-X Host bridge chip.

PCI-X slot 5 can accept a short PCI-X or PCI card and its space is shared with the Remote I/O expansion card. The remaining PCI-X slots are full length cards. The dual 10/100/1000 Mbps Ethernet adapter is integrated on the system planar.

2.5.1 PCI-X slots and adapters

PCI-X, where the X stands for extended, is an enhanced PCI bus, delivering a bandwidth of up to 1 GB/sec, running a 64-bit bus at 133 MHz. PCI-X is backward compatible, so OpenPower 720 systems can support existing 3.3 volt PCI adapters.

The PCI-X slots in the system support hot-plug and Extended Error Handling (EEH). In the unlikely event of a problem, EEH-enabled adapters respond to a special data packet generated from the affected PCI-X 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.

64-bit and 32-bit adapters

IBM offers 64-bit adapter options for the OpenPower 720, as well as 32-bit adapters. 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, some 64-bit adapters cannot be used in 32-bit slots. For a full list of the adapters that are supported on the OpenPower 720 systems, and for important information regarding adapter placement, see the *IBM @server Hardware Information Center*. You can find it at:

http://www.ibm.com/openpower/en_US/infocenter/base/

2.5.2 LAN adapters

The dual port internal 10/100/1000 Mbps RJ-45 Ethernet controller integrated on the system planar can be used to connect an OpenPower 720 to a local area network (LAN).

See the Table 2-1 for the list of additional LAN adapters available at the time of writing. IBM supports an installation with NIM using an Ethernet adapter (CHRP[2] is the platform type).

Table 2-1 Available LAN adapters

Feature code	Adapter description	Slot	Size	Max
4962	10/100 Ethernet	32 or 64	short	5
5700	Gigabit Ethernet	64	short	5
5701	10/100/1000 Ethernet	64	short	5
5706	2-port 10/100/1000 Ethernet	64	short	5
5707	2-port Gigabit Ethernet - SX	64	short	5

2.5.3 Graphic accelerators

The OpenPower 720 supports up to two enhanced POWER GXT135P (FC 2849) 2D graphic accelerators. The POWER GXT135P is a low-priced 2D graphics accelerator for OpenPower, pSeries, and p5 servers. It can be configured to operate in either 8-bit or 24-bit color modes running at 60 Hz to 85 Hz. This adapter supports both analog and digital monitors. The adapter requires one short 32-bit or 64-bit PCI-X slot.

2.5.4 SCSI adapters

To connect to external SCSI devices, the following adapters provided in Table 2-2 are available, at the time of writing, to be used in OpenPower 720 system.

Table 2-2 Available SCSI adapters

Feature code	Adapter description	Slot	Size	Max
5703	Ultra320 SCSI RAID, bootable	64	long	4
5712	Ultra320 SCSI	64	short	5

There is also the option to make the internal Ultra320 SCSI channel externally accessible on the rear side of the system by installing FC 4273. No additional SCSI adapter is required in this case. If FC 4273 is installed, a second 4-pack disk enclosure (FC 6592 or FC 6593)

[2] CHRP stands for Common Hardware Reference Platform, a specification for PowerPC-based systems that can run multiple operating systems.

cannot be installed, which limits the maximum number of internal disks to four. FC 4273 also requires one PCI-X slot.

For more information about the internal SCSI system, see 2.7, “Internal storage” on page 24.

2.6 Internal serial ports

The serial ports S1 and S2, at the rear of the system, are only available if the system is not managed using a Hardware Management Console (HMC). In this case, the S1 and S2 ports support the attachment of serial console and modem.

If an HMC is connected, a *virtual serial console* is provided by the HMC (logical device vsa0 under AIX) and also a modem can be connected to the HMC. The S1 and S2 ports are not usable in this case.

2.7 Internal storage

There is one dual channel Ultra320 SCSI controller managed by the EADS-X chips, integrated into the system planar, that is used to drive the internal disk drives. The eight internal drives plug into the disk drive backplane, which has two separate SCSI buses with four disk drives per bus.

The internal disk drive can be used in two different modes based on whether the SCSI RAID Enablement Card (FC 5709) is installed (see 2.7.3, “RAID options” on page 25).

The OpenPower 720 supports a split 8-pack disk drive backplane that is designed for hot-pluggable disk drives. The disk drive backplane docks directly to the system planar. The virtual SCSI Enclosure Services (VSES) hot-plug control functions are provided by the Ultra320 SCSI controllers.

2.7.1 Internal media devices

The OpenPower 720 provides two slim-line media bays for an optional DVD-ROM (FC 2640) and optional DVD-RAM (FC 5751), and one media bay for a tape drive. These features are summarized in Table 2-3.

Table 2-3 Available tape drives

Feature code	Description
6258	4-mm 36/72 GB tape (LVD)
6134	8-mm 60/150 GB tape (LVD)
6120	VXA 80/160 GB tape (LVD)

2.7.2 Internal SCSI disks

The OpenPower 720 can have up to eight disk drives plugged in the two 4-pack disk drives backplanes. The two 4-pack disk drives backplanes can accommodate the devices listed in Table 2-4 on page 25.

Table 2-4 Disk drive options

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

At the time of writing, if a new order is placed with two 4-pack DASD backplanes (FC 6592) 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.

Hot-swap disks and Linux

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

2.7.3 RAID options

Internal hardware RAID is available on the OpenPower 720. Three options are available:

- Install the Dual Channel SCSI RAID Enablement Card (FC 5709). Install four disk drives in the first 4-pack DASD backplane (FC 6592). This will allow RAID 0, 5, or 10 capabilities within a single 4-pack of DASD with one RAID controller.
- Install FC 5709. Install a second FC 6592. Install four additional disk drives in the second 4-pack DASD backplane. This will allow RAID 0, 5, or 10 capabilities across two 4-packs of DASD with one RAID controller.
- Install feature number 5709. Install the Ultra320 SCSI 4-Pack Enclosure for Disk Mirroring (FC 6593). Install the PCI-X Dual Channel Ultra320 SCSI RAID Adapter (FC 5703). Install the SCSI Cable which connects the PCI Adapter to the second 4-pack DASD backplane (FC 4267). This will allow RAID 0, 5, or 10 capabilities within each 4-pack of DASD with two RAID controllers.

Note: Because the OpenPower 720 has eight disk drive slots, customers performing upgrades must perform appropriate planning to ensure the correct handling of their RAID arrays.

2.8 External I/O subsystem

This section describes the external I/O subsystem, the 7311 Model D20 I/O drawer that is supported on the OpenPower 720 system.

2.8.1 I/O drawers

As described in Chapter 1, “General description” on page 1, the OpenPower 720 system has five internal PCI-X slots, which is enough for average customer configurations. If more PCI-X

slots are needed to dedicate more adapters to a partition or to increase the bandwidth of network adapters, up to eight 7311 model D20 I/O drawers can be added to the OpenPower 720 system.

The OpenPower 720 system contains a RIO-2 bus to connect the internal PCI-X slots through the PCI-X to PCI-X bridges and support up to four external I/O drawers. To support more I/O drawers in one OpenPower 720, a RIO-2 expansion card FC 1806 is needed occupying PCI-X slot 5. The RIO-2 expansion card supports up to four additional I/O drawers.

2.8.2 7311 Model D20 I/O drawer

The 7311 Model D20 I/O drawer must have the RIO-2 loop adapter (FC 6417) to be connected to the system. The PCI-X host bridge inside the I/O drawer provides two primary 64-bit PCI-X buses running at 133 MHz. Therefore, a maximum bandwidth of 1 GB/sec is provided by each of the buses.

Figure 2-8 shows a conceptual diagram of the 7311 Model D20 I/O drawer subsystem.

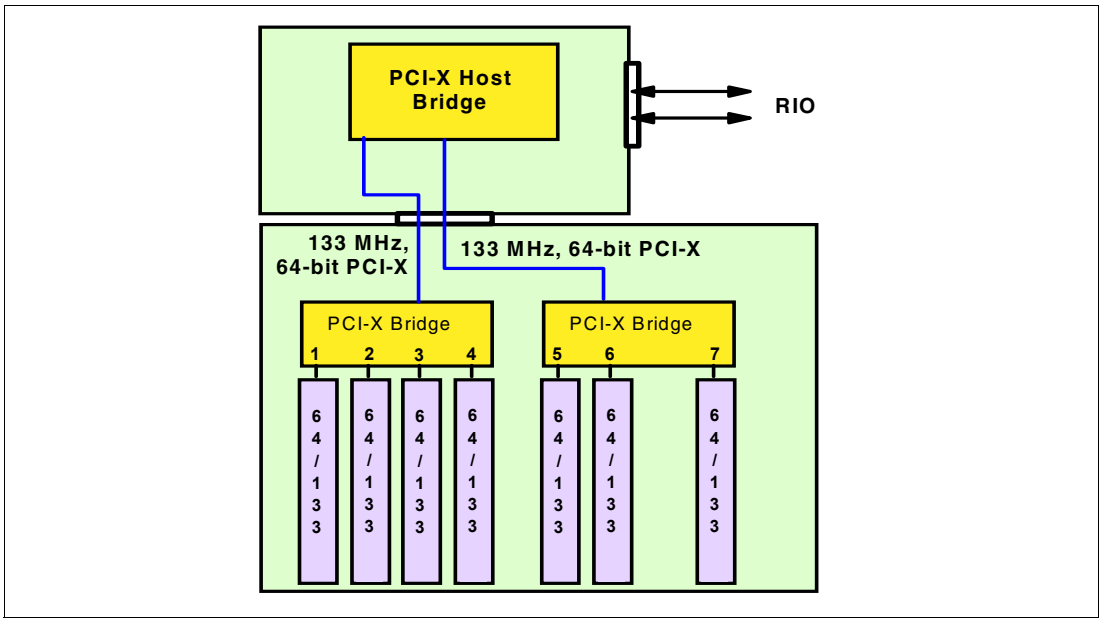


Figure 2-8 Conceptual diagram of the 7311-D20 I/O drawer

7311 Model D20 internal SCSI cabling

A 7311 Model D20 supports hot-swappable disks using two 6-pack disk bays for a total of 12 disks. Additionally, the SCSI cables (FC 4257) are used to connect a SCSI adapter (that can have various features) in slot 7 to each of the 6-packs, or two SCSI adapters, one in slot 4 and one in slot 7 (see Figure 2-9 on page 27).

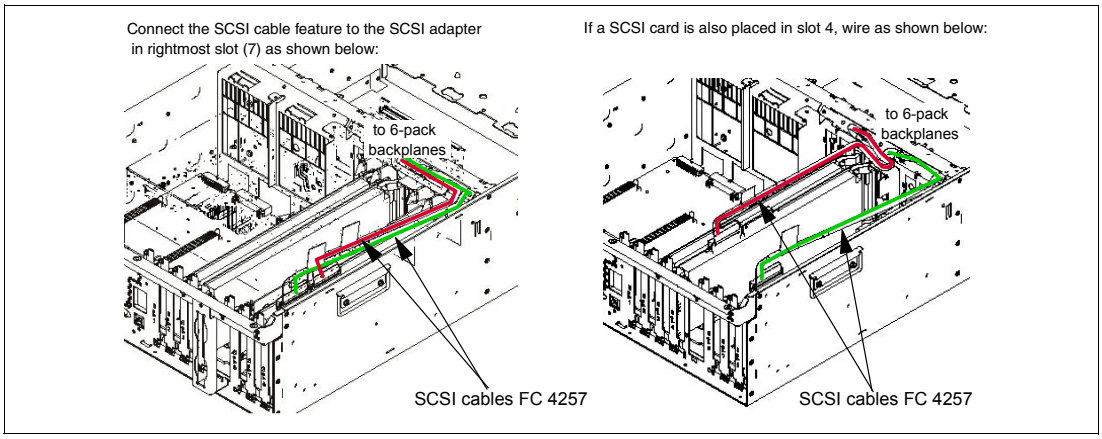


Figure 2-9 7311 Model D20 internal SCSI cabling

Note: Any 6-pack and the related SCSI adapter can be assigned to a single partition. If one SCSI adapter is connected to both 6-packs, both 6-packs can be assigned only to the same partition.

2.8.3 7311 I/O drawer's RIO-2 cabling

As described in the 2.8, "External I/O subsystem" on page 25, you can connect up to four I/O drawers in the same loop and up to eight I/O drawers to the OpenPower 720 system.

Each RIO-2 port can operate at 1 GHz in bidirectional mode and is capable of passing data in each direction on each cycle of the port. Therefore, the maximum data rate is 4 GB/s per I/O drawer in double barrel mode.

There is one default primary RIO-2 loop in any OpenPower 720 system. This feature provides two remote I/O ports for attaching up to four 7311 Model D20 I/O drawers to the system in a single loop. The optional RIO-2 expansion card can be used to increase the number of I/O drawers connected to one OpenPower 720 system, and the same rules of the default RIO-2 loop must be considered. Figure 2-10 on page 28 shows how you could logically connect four I/O drawers to one OpenPower 720 system. Actual cables should be routed according to installation instructions.

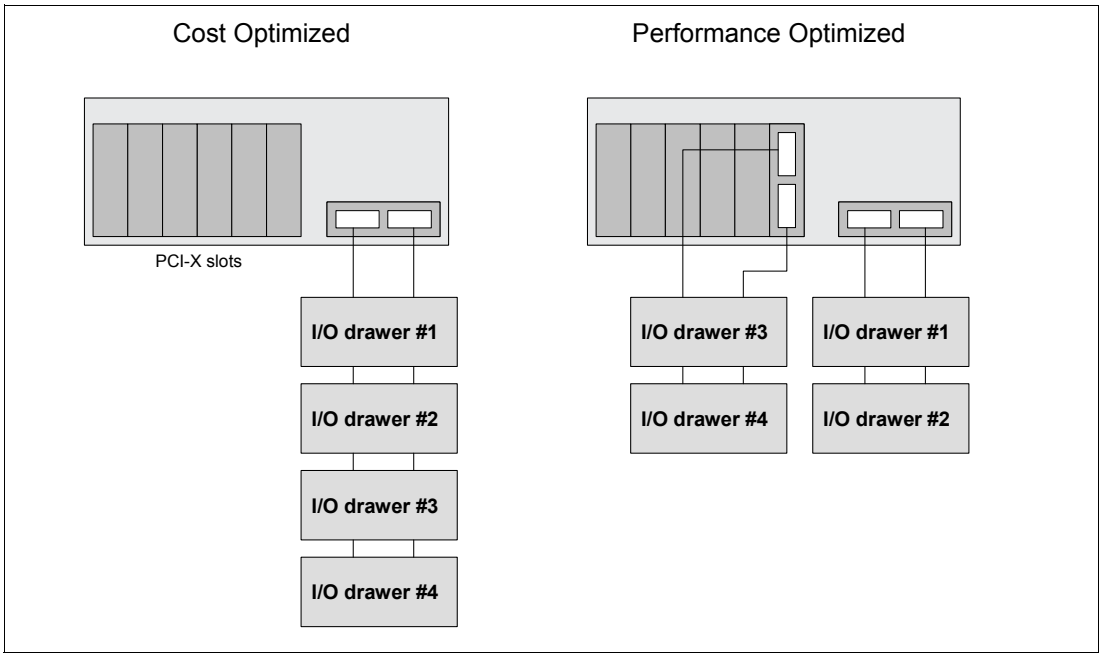


Figure 2-10 RIO-2 connections

The RIO-2 cables used have different lengths to satisfy the different connection requirements:

- Remote I/O cable, 3.5 m (FC 3147)
- Remote I/O cable, 10 m (FC 3148)

2.8.4 7311 Model D20 I/O drawer SPCN cabling

SPCN is used to control and monitor the status of power and cooling within the I/O drawer. The SPCN is a loop (Figure 2-11 on page 29), the cabling starts from SPCN port 0 on the OpenPower system to SPCN port 0 on the first I/O drawer. The loop is closed connecting the SPCN port 1 of the I/O drawer back to the port 1 of the OpenPower 720 system. If you have more than one I/O drawer, you continue the loop connecting the following drawer (or drawers) with the same rule.

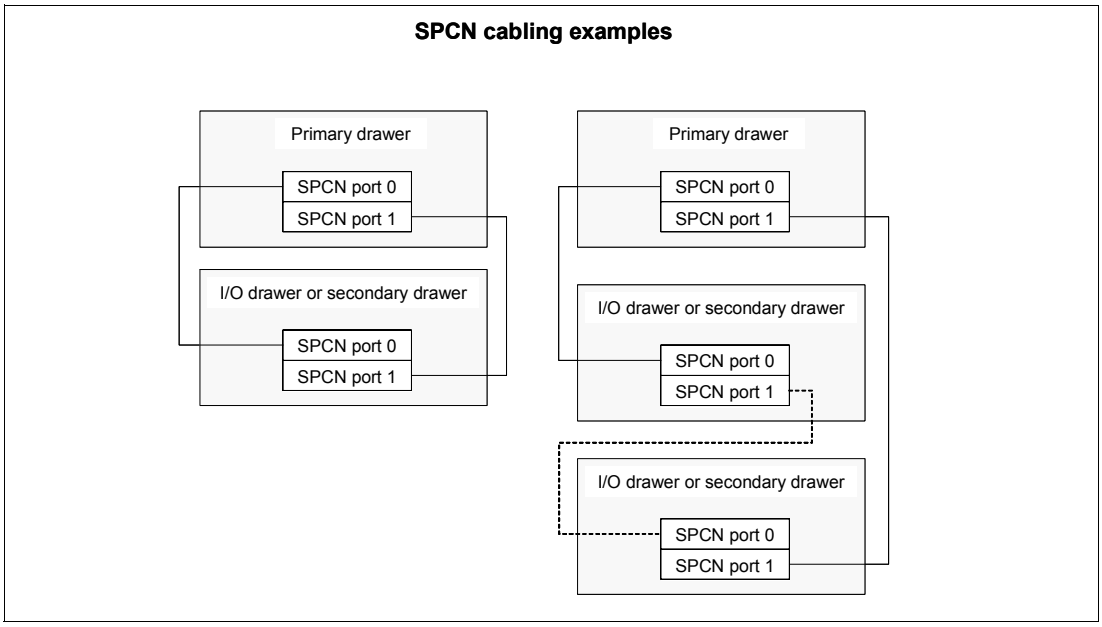


Figure 2-11 SPCN cabling examples

There are different SPCN cables to satisfy different length requirements:

- SPCN cable drawer to drawer, 2 m (FC 6001)
- SPCN cable drawer to drawer, 3 m (FC 6006)
- SPCN cable rack to rack, 6 m (FC 6008)
- SPCN cable rack to rack, 15 m (FC 6007)

2.8.5 External disk subsystems

The OpenPower 720 system has internal hot-swappable drives. Internal disks are usually used for the Linux partition and paging space. The two 4-pack backplanes can be split when the SCSI RAID Enablement Card is not part of the OpenPower configuration, but specific customer requirements can be satisfied with the several external disks possibilities that the OpenPower 720 supports.

IBM 2104 Expandable Storage Plus

The IBM 2104 Expandable Storage Plus Model DS4 is a low-cost 3U disk subsystem that supports up to 14 Ultra320 SCSI disks from 18.2 GB up to 146.8 GB, at the time this publication was written. This subsystem can be used in splitbus mode, meaning the bus with 14 disks could be split into two buses with seven disks each. In this configuration, two additional LPARs (using dedicated devices when configured with the POWER Hypervisor feature) could be provided with up to seven disks for rootvg by using one Ultra3 SCSI adapter (FC 5712) for each LPAR.

For further information about the IBM 2104 Expandable Storage Plus subsystem, visit the following Web site:

http://www.storage.ibm.com/hardsoft/products/expplus/expplus.htm

IBM TotalStorage FAStT Storage servers

The IBM TotalStorage FAStT Storage server family consists of five models: Model 100, 600, 700, and 900. The Model 100 is the smallest model, which scales up to 14 TB, and Model 900

is the largest, which scales up to 32 TB of disk storage, at the time this publication was written. Model 600 provides up to 16 bootable partitions that are attached with the Gigabit Fibre Channel adapter (FC 5716). Model 700 provides up to 64 bootable partitions. In most cases, both the FAStT Storage server and the OpenPower 720 or the 7311 Model D20 I/O drawers are connected to a storage area network (SAN). If only space for the rootvg is needed, the FAStT Model 100 is a good solution.

For support of additional features and for further information about the FAStT family, refer to the following Web site:

http://www.storage.ibm.com/hardsoft/disk/fastt/index.html

IBM TotalStorage Enterprise Storage Server

The IBM TotalStorage Enterprise Storage Server (ESS) is the high-end premier storage solution for use in storage area networks. The 2105 Model 800 provides from 582 GB up to 55.9 TB of usable disk capacity. An ESS system can also be used to provide disk space for booting LPARs or partitions using Micro-Partitioning technology. An ESS is usually connected to a SAN to which the OpenPower 720 is also connected by using Gigabit Fibre Channel adapters (FC 5716).

For further information about ESS, refer to the following Web site:

http://www.storage.ibm.com/hardsoft/products/ess/index.html

2.9 Advanced OpenPower Virtualization

Advanced OpenPower Virtualization is a collection of optional features. The POWER Hypervisor (FC 1965) and the Virtual I/O Server (5724-K81).

To interoperate with the Advanced OpenPower Virtualization features requires either:

- SUSE LINUX Enterprise Server 9 for POWER
- Red Hat Enterprise Linux AS for POWER Version 3

The POWER Hypervisor feature (FC 1965) provides:

- Firmware enablement for LPAR
- Firmware enablement for dynamic LPAR
- Firmware enablement for partitions using Micro-Partitioning technology
- Support of virtual Ethernet

Installation of the Virtual I/O Server software (5724-K81) provides:

- Ethernet adapter sharing
- Support of virtual SCSI

2.9.1 Logical partitioning and dynamic logical partitioning

Introduced with the POWER4 processor product line the logical partition (LPAR) became available. The technology offered the capability to divide a pSeries system into separate systems, where each LPAR runs an operating environment on dedicated attached devices, such as processors, memory, and I/O components. When the customer requested system flexibility to change the system topology on demand, it was achieved by modifying the system layout on the required HMC.

Note: On the OpenPower 720, LPAR, dynamic LPAR, and other advanced virtualization features such as Micro-Partitioning, virtual Ethernet, are available only when a system is configured with the POWER Hypervisor feature, FC 1965.

Later, dynamic LPAR increased the flexibility, allowing selected system resources, such as processors, memory, and I/O components to be added and deleted from dedicated partitions while they are executing. This requires an attached HMC, with the proper level of software, to control the system resources and an updated system firmware level to electronically isolate systems resources. The ability to reconfigure dynamic LPARs encourages system administrators to dynamically redefine all available system resources to reach the optimum capacity for each defined dynamic LPAR.

Dynamic logical partitioning is supported by SUSE LINUX Enterprise Server 9, or later, but with reduced functionality (changing memory attributes dynamically is not supported at the time of writing). Dynamic logical partitioning is not supported by current version of Red Hat Enterprise Linux AS for POWER Version 3.

USB resources are considered a single group, as are slimline devices. Groups can only be allocated to a single partition at a time.

2.9.2 Virtualization

As an optional feature of the OpenPower 720, logical partitions requiring dedicated resources may now be able to take advantage of a new technology that allows resources to be virtualized, allowing for a better overall balance of global system resources and their effective utilization.

Virtual Ethernet

To enhance intercommunication between partitions, either dedicated partitions or partitions using Micro-Partitioning technology, Virtual Ethernet allows in-memory connections at a high bandwidth from partition to partition. Virtual Ethernet working on LAN technology allows a transmission speed in the range of 1 to 3 GB/sec depending on the MTU[3] size. Virtual Ethernet connections (up to 256) are supported in a partition, where a single Virtual Ethernet resource can be connected to another Virtual Ethernet, a real network adapter, or both in a partition.

Micro-Partitioning technology

Based on the partitioning concepts of a stable and well-known mainframe technology and existing LPAR/dynamic LPAR implementation on POWER4 and POWER4+ servers, The POWER5 systems introduce an enhanced partitioning model available as a hardware feature.

The Micro-Partitioning model offers a virtualization method of system resources. In POWER5 processor-based systems, physical resources are abstracted into virtual resources that are available to partitions. This sharing method is the primary feature of this new partitioning concept and it happens automatically after a initial system setup.

POWER5 Micro-Partitioning specifies processor capacity in processing units. One processing unit represents 1% of one physical processor. A partition defined with 220 processing units is equivalent to the power of 2.2 physical processors. Creating a partition using Micro-Partitioning technology, the minimum capacity is 10 processing units, or 1/10 of a

[3] Maximum Transmission Unit

physical processor. A maximum of 10 micro-partitions for each physical processor can be defined, but on a loaded system the practical limit is less.

Partitions using Micro-Partitioning technology can also be defined with the capped and uncapped attributes. A capped micro-partition is not allowed to exceed the defined share, while an uncapped partition is allowed to consume additional capacity with fewer restrictions. Uncapped partitions can be configured to the total idle capacity of the server or a percentage of it. Configuration through the HMC menus sets the allowed share and the capped or uncapped attribute.

The POWER5 processor-based systems use the POWER Hypervisor, which is the new active Hypervisor to execute Micro-Partitioning technology. The Hypervisor of existing POWER4 processor-based systems is working on a demand basis, as the result of machine interrupts and callbacks to the operating system.

The Virtual I/O Server described in 2.9, “Advanced OpenPower Virtualization” on page 30, builds on all the POWER5 and POWER Hypervisor enhancements to reach the highest level of granularity of installed system resources.

Figure 2-12 shows the Micro-Partitioning technology and LPAR concepts.

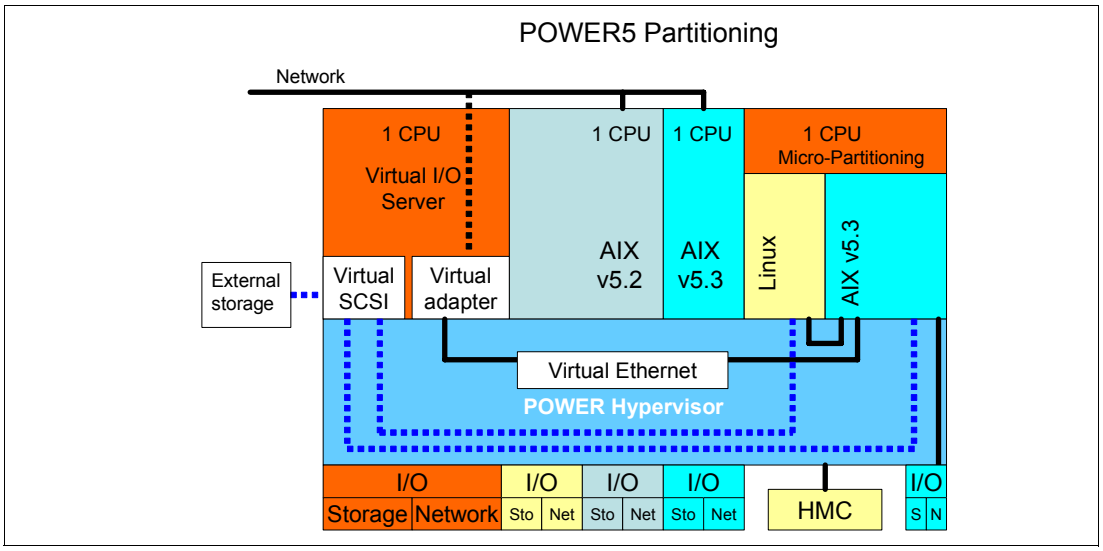


Figure 2-12 LPARs using Micro-Partitioning technology

Virtual I/O Server

The Virtual I/O Server is a special purpose partition to provide virtual I/O resources to client partitions. It is available as software (5724-K81) and requires an HMC for installation and configuration. The Virtual I/O server owns the real resources that will be shared with the other clients. The Virtual I/O technology allows a physical adapter assigned to a partition to be shared by one or more partitions, enabling clients to minimize their number of physical adapters. Virtual I/O will be used to reduce costs by eliminating the requirement that each partition has a dedicated network adapter, disk adapter, and disk drive.

It is preferred that you use the Virtual I/O server in a partition with dedicated resources to help ensure stable performance.

The two major functions that are provided from the Virtual I/O Server are described in the following sections.

Shared Ethernet adapter

The Shared Ethernet adapter is a new service that acts as a layer 2 network switch to route network traffic from a Virtual Ethernet to a real network adapter. The Shared Ethernet Adapter must run in a Virtual I/O Server partition.

The advantages of the virtual Ethernet services is that partitions to communicate outside the system without having a physical network adapter attached to the partition. At the time of writing, up to 16 Virtual Ethernet x 18 VLANs can be shared on a single network interface. The amount of network traffic will limit the number of client partitions served through a single network interface.

Virtual SCSI

Access to real storage devices is implemented through the Virtual SCSI services, a part of the Virtual I/O Server partition. Logical volumes created and exported on the Virtual I/O Server partition will be shown at the Virtual Storage Client partition as a SCSI disk. All current storage device types, such as SAN, SCSI, and RAID, are supported. iSCSI and SSA are not supported.

Note: The Shared Ethernet adapter and Virtual SCSI server function is provided in the Virtual I/O Server that is an optional orderable software feature. An HMC is required.

2.10 Service processor

The service processor (SP) is an embedded controller based on a PowerPC 405GP processor (PPC405) implementation running the SP internal operating system. The SP operating system contains specific programs and device drivers for the SP hardware (Figure 2-13 on page 33).

The OpenPower 720 uses the SP implementation. The key components include a FSP-Base (FSP-B) and an Extender chipset (FSP-E). FSP-B and FSP-E are implemented on a dedicated card.

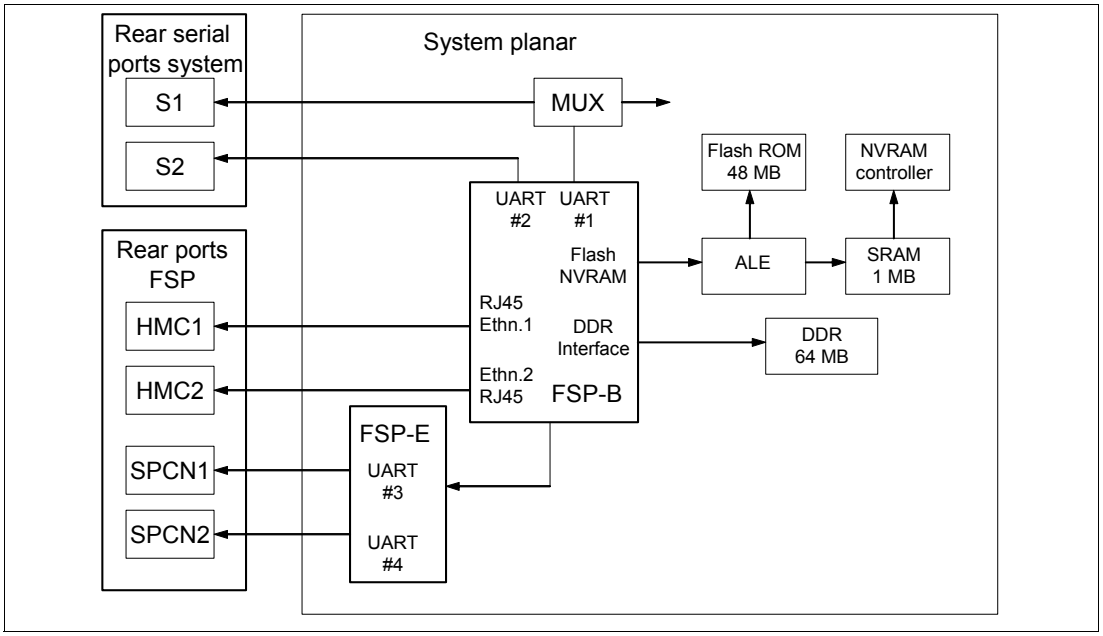


Figure 2-13 Service processor logical diagram

2.10.1 Service processor base

The PPC405 core is 5-stage pipeline instruction processor and contains 32-bit general purpose registers. The Flash ROM contains a compressed image of a software load.

The SP base unit offers the following connections:

- Two Ethernet Media Access Controller3 (MAC3) cores, which is a generic implementation of the Ethernet Media Access (MAC) protocol compliant with ANSI/IEEE 802.3, IEEE 802.3u, ISO/IEC 8802.3 CSMA/CD standard. The Ethernet MAC3 supports both half duplex (CSMA/CD) and full duplex operation at 10/100 Mbps. Both Ethernet port are visible only to the service processor.

- Two serial interfaces, accessible only though the serial ports of OpenPower 720 on the rear side. At the time of writing, the System Management Interface (SMI) is usable if a connection is established to serial port 1. Terminals connected to serial port 2 receive only boot sequence information without manual interaction. When the HMC is connected to the SP, the serial ports are disabled and do not provide any external connection.

2.10.2 Service processor extender

The SP extender unit offers two system power control network (SPCN) ports that are used to control the power of the attached I/O subsystems. The SPCN control software and the service processor software are run on the same PPC405 processor.

2.11 Boot process

With the implementation of the POWER5 chip technology in the pSeries platform, the boot process is enhanced for the flexibility that the POWER5 processor-based hardware features. Depending on the customer configuration, a system may or may not require the use of an HMC to manage the system. The boot process, based on the Initial Program Load (IPL) setup, will depend on the hardware setup and on the way we will use the features that POWER5 processor-based systems provide.

The IPL process starts when power is connected to the system. Immediately after, the SP starts an internal self test based on integrated diagnostic programs (Built-In-Self-Test, BIST). Only if all the test units have been successfully passed, the system status changes to standby.

2.11.1 IPL flow without an HMC attached to the system

When the system status is standby, the SP provides a System Management Interface (SMI) that can be accessed by pressing any key on an attached serial console keyboard, or the Advanced System Management Interface (ASMI) using a Web browser[4] on a client system that is connected to the SP on an Ethernet network.

The SP and the ASMI are standard on all POWER5 processor-based hardware. Both system management interfaces require you to enter the general or admin ID password and allow you to set flags that affect the operation of the system, according to the provided password, such

[4] Supported browsers are Netscape (Version 7.1), Microsoft Internet Explorer (version 6.0), and Opera (Version 7.23). At the time of writing, previous versions of these browsers are not supported. JavaScript™ and cookies must be enabled.

as auto power restart, to view information about the system (such as the error log and VPD), network environment access setup, and to control the system power.

You can start and shut down the system in addition to setting IPL options. This server has a permanent firmware boot side, or A side, and a temporary firmware boot side, or B side. New levels of firmware should be installed on the temporary side first in order to test the update's compatibility with your applications. When the new level of firmware has been approved, it can be copied to the permanent side.

In the SMI and ASMI, you can view and change the following IPL settings:

- System boot speed.

  Fast or Slow. Fast boot results in skipped diagnostic tests and shorter memory tests during the boot.

- Firmware boot side for next boot.

  Permanent or Temporary. Firmware updates should be tested by booting from the temporary side before being copied into the permanent side.

- System operating mode.

  Manual or Normal. Manual mode overrides various automatic power-on functions, such as auto-power restart, and enables the power switch button.

- AIX/Linux partition mode boot, available only if the system is not managed by the HMC.
  - Service mode boot from saved list. This is the preferred way to run concurrent AIX diagnostics.
  - Service mode boot from default list. This is the preferred way to run stand-alone AIX diagnostics.
  - Boot to open firmware prompt.
  - Boot to System Management Service (SMS) to further select the boot devices or network boot options.

- Boot to server firmware:

  Select the state for the server firmware: Standby or Running.

  When the server is in the server firmware standby state, partitions can be set up and activated

2.11.2 Hardware Management Console

Depending on the model, the HMC provides a number of native serial ports and Ethernet port. One serial port can be used to attach a modem for the Service Agent. The Service Agent Connection Manager can be used instead if the HMC has a TCP/IP port 80 connection to the Internet. The HMC provides an Ethernet port (or ports) to connect to partitions on its POWER5 processor-based managed systems. The network connection is mandatory for the HMC to OpenPower systems, and highly recommended between the HMC and partitions. It supports the following functions:

- Logical partition configuration and management
- Dynamic logical partitioning
- Capacity and resource management
- System status
- HMC management
- Service functions (for example, Microcode Updates and Service Focal Point)
- Remote HMC interface

Note: The same HMC cannot be attached to POWER4 and POWER5 processor-based systems simultaneously, but for redundancy purposes, one POWER5 processor-based server can be attached to two HMCs.

All the managed servers must be authenticated from the HMC. If a new attached system is discovered, the HMC will prompt you to set two passwords using the HMC interface:

- Advanced System Management general user ID password
- Advanced System Management admin ID password
- HMC access password

2.11.3 IPL flow with an HMC attached to the system

When the system status is in standby, you can use the HMC to open a virtual terminal and access the SMI, or launch a Web browser to access the ASMI.

Using the SMI or the ASMI, you can view or modify the proper IPL settings in order to set the boot mode to partition standby and then turn the system on. However, the HMC can be also used to power on the managed system (and is highly recommended). Using the HMC to turn the system on requires you to select one of the following choices:

- Partition Standby
  - The Partition Standby power-on mode allows you to create and activate logical partitions.
  - When the Partition Standby power-on is completed, the operator panel on the managed system displays *LPAR...*, indicating the managed system is ready for you to use the HMC to partition its resources and, possibly, activate them.

- When a partition is activated, the HMC requires you to select the boot mode of the single partition.

- System Profile

  The System Profile option powers on the system according to a predefined set of profiles. The profiles are activated in the order in which they are shown in the system profile.

- Partition autostart

  This option powers on the managed system to partition standby mode and then activates all partitions that have been designated autostart.

After the system boots with any of the above choices, the HMC can be used to manage the system, such as continuing to boot from the operating system or manage the logical partitions. See 2.11.2, “Hardware Management Console” on page 36.

2.11.4 Definitions of partitions

Describing the detailed process to work with the HMC and the management tasks to create and manage a logical partition (LPAR or dynamic LPAR) is not the intention of this documentation. The following section describes the additional functionality used to create partitions that are using fractional elements of available system resources, namely Micro-Partitioning.

For a better understanding of the partitioning concept, this section contains an overview of common terminology. There are two important components:

- Managed systems
- Profiles

Managed systems

Managed systems are physical systems that are managed by the HMC, whereby one HMC can manage more managed systems at a time.

Profiles

A profile defines the configuration of a logical partition or managed system. There are three types of profiles that can be used to create multiple profiles for each logical partition or managed system:

- Partition profile
  - A partition profile includes the collection of resource specifications, such as processing units, memory, and I/O resources, because a logical partition is not aware of a resource until it is activated.
  - A logical partition can have more than one partition profile, but at least one is a minimum requirement.
- All resources dedicated partition profile

  A partition profile that contains the entire resource list of the machine, using all physical resources working as one system.
- System profile
  - A system profile is an ordered list of partition profiles. When you activate a system profile, the managed system will attempt to activate the partition profiles in the defined order.

– To enhance the flexibility to use the system within several different logical configuration, a system profile can be defined to collect more than one partition profile to provide the requested system behavior.

2.11.5 Hardware requirements for partitioning

To implement Micro-Partition technology on a POWER5 processor-based system, resource planning is important to have a base configuration and enough flexibility to make desirable changes to the running logical partitions. To configure a partition, the minimum requirements needed are processors, memory, and possibly an expansion unit to define more partitions than possible in a single system.

Processors

Within POWER5 technology, and depending on performance requirements, a logical partition can be created by using a shared processor pool or a dedicated processor.

Shared processors can be defined by a fractional number of 1/10 as minimum requirement of a real processor. To calculate the required processor power, a real processor is divided in 100 processing units, and 1/10 of a processor is equal to 10 processing units.

Dedicated processors are entire processors that can be assigned to a single logical partition without the capability to share free capacity with other logical partitions.

Memory

Depending on given application and performance requirements, a logical partition requires memory to execute the installed operating system and application. To create partitions, the minimum memory requirement is 128 MB per logical partition.

Expansion unit

Expansion units extend the flexibility of the server system to enlarge the number of possible logical partitions by adding additional hardware, such as storage or network devices.

2.11.6 Specific partition definitions used for Micro-Partitioning technology

In addition to the base definitions for a partitions, new parameters must be defined to receive more flexibility and capacity usage of logical partitions included in POWER5 technology.

Capped and uncapped partitions

A capped partition indicates that the local partition will never exceed its assigned capacity. A partition using the uncapped mode indicates that if the maximum capacity is reached, the shared pool of processing power is used for more resources if available.

To manipulate the behavior of uncapped partitions, the parameter uncapped weight, in the range from 0 through 255, must be defined. To prevent an uncapped partition from receiving extra capacity, the uncapped weight parameter should be 0.

The default uncapped weight is 128.

2.11.7 System Management Service

Either booting up a full partition system or a logical partition to System Management Services (SMS), the ASCII[5] interface or the GUI are identical in contents and functionality.

The OpenPower 720 (or the logical partition) must be 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 native serial ports, or the attached HMC to use the SMS menus. It is possible to view information about the system (or the single logical partition) and perform tasks such as setting a password, changing the boot list, and setting the network parameters.

If the system or the partition has been activated without flagging the option to stop to the SMS, there is the option to press the 1 key on the terminal, or in the graphic window, after the word `keyboard` appears and before the word `speaker` appears. In the terminal, or in the GUI, the system or the partitions will require you to enter the password defined for admin or general access. After the text-based SMS starts (either for terminal or graphic window), a screen similar to the one shown in Figure 2-14 opens.

```
Version SF220_004
SMS 1.5 (c) Copyright IBM Corp. 2000,2003 All right reserved
------------------------------------------------------------------------------
Main Menu
 1.  Select Language
 2.  Setup Remote IPL (Initial Program Load)
 3.  Change SCSI Settings
 4.  Select Console
 5.  Select Boot Options









------------------------------------------------------------------------------
Navigation Keys:

                                          X = eXit System Management Services
------------------------------------------------------------------------------
Type the number of the menu item and press Enter or select Navigation Key:
```

Figure 2-14 System Management Services main menu

Note: The version of system firmware currently installed in your system is displayed at the top of each screen. Processor and other device upgrades might 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 review or set the boot list information, or to set up the network environment parameters if you want the system to boot from a NIM server.

2.11.8 Boot options

The OpenPower 720 handles the boot process in a way that is similar to other IBM UNIX 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. After the machine or the logical partition

[5] American Standard Code for Information Interchange: This is the world-wide standard for the code numbers used by computers to represent all the uppercase and lowercase Latin letters, numbers, punctuation and so forth.

reaches the SMS menus, all of the necessary tests have been performed and the machine is scanning the bus for a boot source.

Most system backplanes 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. Disk drive bays 1 and 5 are hardwired to spin-up immediately. The left-most and accordingly bottom-most slot of the 4-pack disk backplanes (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 the preferred boot device. See Figure 2-15 to locate all of the disk bays.

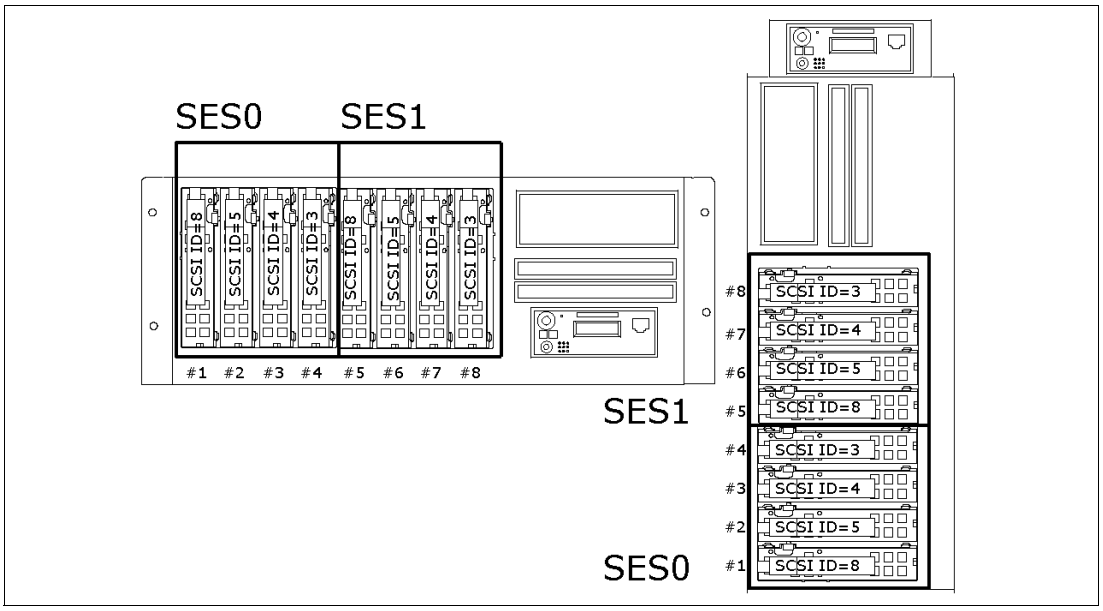


Figure 2-15 Disk bays and SCSI addresses within a OpenPower 720

When SMS menus are available, the Select Boot Options menu can be used to view and set various options regarding the installation devices and boot devices:

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.

2. Select Boot Devices

   Enables you to set the boot list.

3. Multiboot Startup

   Toggles the multiboot startup flag, which controls whether the multiboot menu is invoked automatically on startup.

2.11.9 Additional boot options

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, or a logical partition (if the resource is available to the specific partition), so that a system can be loaded, system

	maintenance performed, or stand-alone diagnostics performed.
SCSI disk, and Virtual 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 SCSI-attached disk could be used if required. As described in previous sections, Virtual SCSI devices are also available to a logical partition.
SAN boot	It is possible to boot the OpenPower 720 system from a SAN using a 2 GB Fibre Channel Adapter (FC 6239), or it is possible to boot one partition using the dedicated 2 GB Fibre Channel Adapter or the Virtual SCSI device related to this adapter. 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. Logical partitions can use both a dedicated Ethernet adapter or Virtual Ethernet to accomplish that.

2.11.10 Security

The OpenPower 720 system allows you to set two different types of passwords to limit the access to these systems. These are defined in the ASMI menus. This password is usually used by the system administrator. The *general ID password* provides limited access to the system functions and is usually available to all users who are allowed to power on the server, especially remotely.

2.12 Operating system requirements

OpenPower servers support appropriate versions of Linux.

2.12.1 Linux

For the OpenPower 720, Linux distributions are available through SUSE and Red Hat at the time this publication was written. A software pre-install feature (FC 5005) is available. The OpenPower 720 requires the following version of Linux distributions:

- SUSE LINUX Enterprise Server 9 for POWER systems, or later
- Red Hat Enterprise Linux AS for POWER Version 3

The Advanced OpenPower Virtualization features and DLPAR require SUSE SLES 9. Red Hat distributions support all the Advanced OpenPower Virtualization functions.

In Japan, Turbolinux is also available. In the Latin America sales region, Conectiva is also available. For related information and an overview, see:

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

To find full information about Red Hat Enterprise Linux AS for POWER Version 3 at:

http://www.redhat.com/software/rhel/as/

To find full information about SUSE LINUX Enterprise Server 9 for POWER, see:

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, 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

RAS and manageability

The following sections provide more detailed information about IBM @server OpenPower design features that will help lower the total cost of ownership (TCO).

3.1 Reliability, availability, and serviceability

Excellent quality and reliability are inherent in all aspects of the IBM @server OpenPower design and manufacturing. The fundamental objective 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 might occur. This is achieved using capabilities provided by both the hardware and the operating system.

The OpenPower, as a POWER5 server, enhances the RAS capabilities implemented in POWER4-based systems. RAS enhancements available on POWER5 servers are:

- Most firmware updates allow the system to remain operational.
- The ECC has been extended to inter-chip connections for the fabric and processor bus.
- Partial L2 cache deallocation is possible.
- The number of L3 cache line deletes improved from 2 to 10 for better self-healing capability.

The following sections describe the concepts that form the basis of leadership RAS features of IBM @server OpenPower systems in more detail.

3.1.1 Fault avoidance

The OpenPower systems are built on a quality-based design to keep errors from ever happening. This design includes the following features:

- Reduced power consumption, cooler operating temperatures for increased reliability, enabled by copper chip circuitry, silicon-on-insulator, and dynamic-clock-gating
- Mainframe-inspired components and technologies

3.1.2 First Failure Data Capture

If a problem should occur, the ability to correctly diagnose it is a fundamental requirement upon which improved availability is based. The OpenPower 720 incorporates advanced capability in start-up diagnostics and in run-time First Failure Data Capture (FDDC) 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 (SP). The SP in the OpenPower 720 has the capability to access system components using special purpose service processor ports or by access to the error registers (Figure 3-1 on page 45).

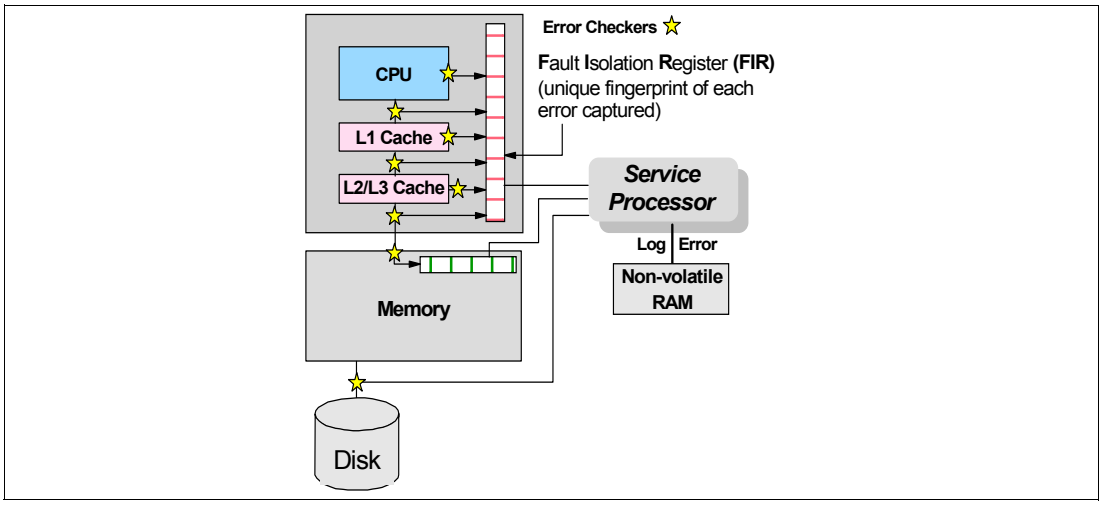


Figure 3-1 Schematic of Fault Isolation Register implementation

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 can include dynamic deallocation of potentially failing components.

Errors are logged into the system non-volatile random access memory (NVRAM) and the SP event history 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. For more information, see “Dynamic or persistent deallocation” on page 47.

3.1.3 Permanent monitoring

The SP included in the OpenPower 720 provides a means to monitor the system even when the main processor is inoperable. See the following subsections for a more detailed description of monitoring functions in an OpenPower 720.

Mutual surveillance

The SP can monitor the operation of the firmware during the boot process, and it can monitor the operating system for loss of control. This allows the service processor to take appropriate action, including calling for service, when it detects that the firmware or the operating system has lost control. Mutual surveillance also allows the operating system to monitor for service processor activity and can request a service processor repair action if necessary.

Environmental monitoring

Environmental monitoring related to power, fans, and temperature is done by the System Power Control Network (SPCN). Environmental critical and non-critical conditions generate Early Power-Off Warning (EPOW) events. Critical events (for example, Class 5 AC power loss) trigger appropriate signals from hardware to impacted components so as to prevent any data loss without the operating system or firmware involvement. Non-critical environmental events are logged and reported through Event Scan.

The operating system cannot program or access the temperature threshold using the SP.

EPOW events can, for example, trigger the following actions:

- Temperature monitoring, which increases the fans speed rotation when ambient temperature is above a preset operating range.
- Temperature monitoring warns the system administrator of potential environmental-related problems. It also performs an orderly system shutdown when the operating temperature exceeds a critical level.
- Voltage monitoring provides warning and an orderly system shutdown when the voltage is out of the operational specification.

3.1.4 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:

- *Bit steering* to redundant memory in the event of a failed memory module to keep the server operational.
- *Bit-scattering*, thus allowing for error correction and continued operation in the presence of a complete chip failure (*Chipkill recovery*).
- Single bit error correction using ECC without reaching error thresholds for main, L2, and L3 cache memory.
- L3 cache line deletes extended from 2 to 10 for additional self-healing.
- ECC extended to inter-chip connections on fabric and processor bus.
- *Memory scrubbing* to help prevent soft-error memory faults.
- *Dynamic processor deallocation*, a deallocated processor can be replaced by an unused Capacity on Demand processor to keep the system operational.

Memory reliability, fault tolerance, and integrity

The OpenPower 720 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 use *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.

3.1.5 N+1 redundancy

The use of redundant parts allows the OpenPower 720 to remain operational with full resources:

- Redundant spare memory bits in L1, L2, L3, and main memory
- Redundant fans

- Redundant power supplies (optional)

3.1.6 Fault masking

If corrections and retries succeed and do not exceed threshold limits, the system remains operational with full resources, and no client or IBM customer engineer intervention is required. This technology is used in the following faults:

- CEC bus retry and recovery
- PCI-X bus recovery
- ECC Chipkill soft error

3.1.7 Resource deallocation

If recoverable errors exceed threshold limits, resources can be deallocated with system remaining operational, allowing deferred maintenance at a convenient time.

Dynamic or persistent deallocation

Dynamic deallocation (available on SUSE LINUX Enterprise Server 9 for POWER) of potentially failing components is non-disruptive, allowing the system to continue to run. Persistent deallocation occurs when a failed component is detected, which is then deactivated at a subsequent reboot.

Dynamic deallocation functions include:

- Processor
- L3 cache line delete
- Partial L2 cache deallocation
- PCI-X 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 two).

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. 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.

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 OpenPower 720 supports PCI Extended Error Handling (EEH) if it is supported by the PCI-X 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 OpenPower system,

hardware, system firmware, and system interaction have 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.

Note: 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)

3.1.8 Serviceability

Increasing service productivity means the system is up and running for a longer time. OpenPower improves service productivity by providing the functions described in the following subsections:

Error indication and LED indicators

The OpenPower 720 is designed for customer setup of the machine and for the subsequent addition of most hardware features. The OpenPower 720 also allows customers to replace service parts (Customer Replaceable Unit). To accomplish this, the system provides internal LED diagnostics that will identify parts that require service. 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, and ending with an LED near the failing Field Replaceable Unit.

For more information about Customer Replaceable Units, including videos, see:

http://publib.boulder.ibm.com/eserver

System Attention LED

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.

Additional LEDs on I/O components such as PCI-X slots and disk drives provide status information such as power, hot-swap, and need for service.

Concurrent Maintenance

Concurrent Maintenance provides replacement of the following parts while the system remains running:

- Disk drives
- Cooling fans
- Power subsystems
- PCI-X adapter cards

3.2 Manageability

The functions and tools provided for IBM @server OpenPower systems are described in the next sections.

3.2.1 Service processor

With system in power standby mode, or with an operating system in control of the machine, or controlling the related partition, the SP is working and checking the system for errors, ensuring the connection to the HMC for manageability purposes. With system up and running, the SP provides the possibility to view and change the Power-On settings, using the Advanced System Management Interface (ASMI). Also, the surveillance function of the SP is monitoring the operating system to check that it is still running and has not stalled.

See Figure 3-2 on page 50 for an example of the Advanced System Management Interface accessed from a Web browser.

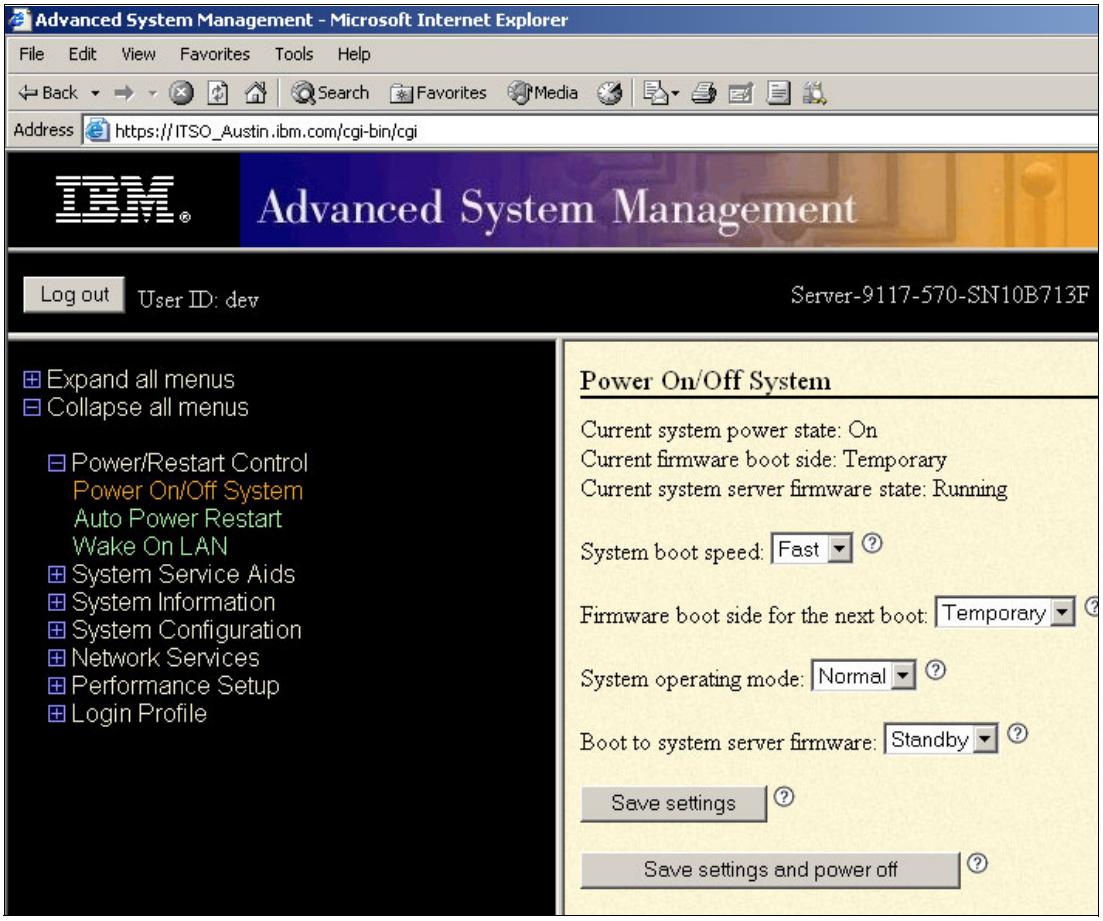


Figure 3-2 Advanced System Management main menu

3.2.2 Service Agent

Service Agent is an application program that operates on an OpenPower, p5, pSeries or IBM RS/6000 computer and monitors them for hardware errors. It reports detected errors, assuming they meet certain criteria for criticality, to IBM for service with no customer intervention. It is an enhanced version of Service Director with a graphical user interface.

Key things you can accomplish using Service Agent for an OpenPower, p5, pSeries, and RS/6000 include:

- Automatic problem analysis
- Problem-definable threshold levels for error reporting
- Automatic problem reporting; service calls placed to IBM without intervention
- Automatic customer notification

In addition:

- Commonly viewed hardware errors. You can view hardware event logs for any monitored machine on the network from any Service Agent host user interface.
- Network environment support with minimum telephone lines for modems.
- VPD data can be sent to IBM using Performance Management.

Machines are defined by using the Service Agent user interface. After the machines are defined, they are registered with the IBM Service Agent Server (SAS). During the registration process, an electronic key is created that becomes part of your resident Service Agent program. This key is used each time the Service Agent places a call for service. The IBM Service Agent Server checks the current customer service status from the IBM entitlement database; if this reveals that you are not on Warranty or MA, the service call is refused and posted back through an e-mail notification.

Service Focal Point

Service Focal Point is used by service technicians to start and end their service calls. It provides service representatives with event, Vital Product Data (VPD), and diagnostic information. The HMC can also notify service representatives of hardware failures automatically by using the Service Agent features. You can configure the HMC to use the Service Agent call-home feature to send IBM event information. This information is stored, analyzed, and then acted upon by the service representative. Some parts of Service Focal Point need to be configured so that the proper information is sent to IBM.

You can download the latest version of Service Agent at:

`ftp://ftp.software.ibm.com/aix/service_agent_code`

3.2.3 OpenPower Customer-Managed Microcode

The Customer-Managed Microcode is a methodology that enables you to manage and install microcode updates on an OpenPower, p5, pSeries and RS/6000 systems and associated I/O adapters. The IBM pSeries Microcode Update Web site can be found at:

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

IBM provides service tools that can assist you in determining microcode levels and updating systems with the latest available microcode. To determine which tool to use in a specific environment, visit:

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

3.3 Cluster 1600

Today's IT infrastructure requires that systems meet increasing demands, while offering the flexibility and manageability to rapidly develop and deploy new services. IBM clustering hardware and software provide the server, with availability, scalability, security, and single-point-of-management control, to satisfy these needs.

IBM @server Cluster 1600 is a POWER-based AIX 5L and Linux Cluster targeting scientific and technical computing, large-scale databases, and workload consolidation.

IBM Cluster Systems Management (CSM) is designed to provide a robust, powerful, and centralized way to manage a large number of POWER5-based systems all from one single point of control. CSM can help lower the overall cost of IT ownership by helping to simplify the tasks of installing, operating, and maintaining clusters of servers. CSM can provide one consistent interface for managing both AIX and Linux nodes (physical systems or logical partitions), with capabilities for remote parallel network install, remote hardware control, and distributed command execution.

The OpenPower 720 is supported with the Cluster 1600 running CSM for AIX, V1.3.1. To attach a OpenPower 720 to a Cluster 1600, an HMC is required. One HMC can also control

several OpenPower systems that are part of the cluster. If a OpenPower 720 is configured in partition mode (with physical or virtual resources) is part of the cluster, all partitions must be part of the cluster. (Check with your IBM support representative for the availability of this feature).

It is not possible to use selected partitions as part of the cluster and use others for non-cluster use. The HMC uses a dedicated connection to the OpenPower 720 to provide the functions needed to control the server, such as powering the system on and off. The HMC must have an Ethernet connection to the Control Work Station (CWS). Each partition in an OpenPower 720 must have an Ethernet adapter to connect to the CWS *trusted* LAN.

Information regarding HMC control, cluster building block servers, and cluster software available at:

http://www.ibm.com/servers/eserver/clusters/

The benefits of a clustered environment based on logical partitions

The evolution of processor and storage technologies has a great impact on the architecture of IT infrastructures. This is the most significant challenge for the infrastructure. During the first half of the 1990s, one single central instance of an application per node was suitable, moreover, most productive systems needed additionally associated nodes, so-called application servers.

Increasing performance and reliability by simply replicating application server nodes led to complex environments that often resulted in poor system management. The reason for these complex constructions was the limited computing power of a single node. This limitation was softened during the second half of the 1990s.

Large symmetric multiprocessor (SMP) nodes with higher clock rates and increased memory provided the possibility to install more than one system on a node. More systems made release planning more difficult as operations had to pay attention to different databases or application versions to avoid unresolved conflicts.

Some customers expanded the usage of their dedicated systems and consequently model more business processes. This often caused an increased number of dedicated systems used and a stronger demand on flexibility. In addition, the life cycle of these systems differed extremely. Renaming, removal, or deletion became more and more common system administration tasks.

In 2001, the pSeries hardware technology with logical partitioning was generally available. Logical partitioning provides the ability to define the logical partitions (LPARs) that are adapted to customer needs regarding the number of processors, assigned memory, and I/O adapters. The OpenPower 720 offers the flexibility to increase the usage of the resources even more using Micro-Partitioning technology, and reduce the total cost of ownership (TCO).

Partitions with associated physical resources or virtual resources are not different from a collection of stand-alone nodes.

Today, server consolidation is a must for many IT sites. Minimized TCO and complexity, with the maximum amount of flexibility, is a crucial goal of nearly all customers. LPARs allow a flexible distribution of resources with LPAR boundaries. Each logical partition can be configured according to the specific needs of the occupant application. LPARs provide a protection boundary between the systems. More test and development systems can exist on the same server in separate partitions.

The CSM value points

The CSM allows the management of different hardware platforms from one single point of control and it has consistent interfaces to manage systems and logical partitions running both AIX and Linux. The management is achieved across multiple switch and interconnect topologies. PSSP (a solution preceding CSM) had tools that only allowed system administrators to perform tasks in specific way (such as NIM, and SP user management). The CSM provides assistance on these specific tasks, allows the system administrator to tailor their system to their own needs, and it has the ability to manage systems across different geographical sites.

Monitoring is much easier to use and the system administrator can monitor all the network interfaces, not just the switch and administrative interfaces. The management server pushes information out to the nodes, which allows the management server to not have to trust the node. In addition, the nodes do not have to be network connected to each other either. This means that giving root access on one node does not mean giving root access on all the nodes. The base security setup is all done automatically at install time.

CSM Version 1.4 on AIX and Linux (planned 4Q04)

The CSM V1.4 on Linux introduces an optional IBM CSM High Availability Management Server feature, designed to allow automated failover of the CSM management server to a backup management server. In addition, sample scripts for setting up NTP[1], and the capability to copy files across nodes or node groups in the cluster can improve cluster ease of use and site customization.

[1] Network Time Protocol

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 56. Note that some of the documents referenced here may be available in softcopy only.

- *Linux Handbook: A Guide to IBM Linux Solutions and Resource*, SG24-7000
- *The Complete Partitioning Guide for IBM pSeries Servers*, SG24-7039
- *Practical Guide for SAN with pSeries*, SG24-6050
- *AIX and Linux Interoperabilty*, SG24-6622
- *Linux with xSeries and FAStT: Essentials*, SG24-7026
- *Linux Clustering with CSM and GPFS*, SG24-6601
- *Deploying Linux on IBM eServer pSeries Clusters*, SG24-7014

Other publications

These publications are also relevant as further information sources:

- *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 can be installed.
- *7316-TF3 17-Inch Flat Panel Rack-Mounted Monitor and Keyboard Installation and Maintenance Guide*, SA38-0643, contains information regarding the 7316-TF3 Flat Panel Display, which can be installed in your rack to manage your system units.
- *IBM @server Hardware Management Console for pSeries Installation and Operations Guide*, SA38-0590, provides information to operators and system administrators on how to use a IBM Hardware Management Console for pSeries (HMC) to manage a system. It also discusses the issues associated with logical partitioning planning and implementation.
- *Planning for Partitioned-System Operations*, SA38-0626, provides information to planners, system administrators, and operators about how to plan for installing and using a partitioned server. It also discusses some issues associated with the planning and implementing of partitioning.

Online resources

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

- Ceramic Column Grid Array (CCGA), see IBM Chip Packaging

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

  http://www.ibm.com/chips/technology/technologies/copper/
- Hardware documentation

  http://publib16.boulder.ibm.com/openpower/en_US/infocenter/base/
- IBM @server Information Center

  http://publib.boulder.ibm.com/eserver/
- IBM @server pSeries and RS/6000 microcode update

  http://techsupport.services.ibm.com/server/mdownload2/download.html
- 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
- 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
- SCSI T10 Technical Committee

  http://www.t10.org
- Silicon-on-insulator (SOI) technology

  http://www.ibm.com/chips/technology/technologies/soi/
- SUSE LINUX Enterprise Server 8 for pSeries information

  http://www.suse.de/us/business/products/server/sles/i_pseries.html
- The LVT is a PC based tool intended assist you in logical partitioning

  http://www-1.ibm.com/servers/eserver/iseries/lpar/systemdesign.htm

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

Help from IBM

IBM Support and downloads

ibm.com/support

IBM Global Services

ibm.com/services

IBM @server OpenPower 720 Technical Overview and Introduction

Redpaper

Deskside and rack-mount server for managing e-business

Outstanding performance based on POWER5 processor technology

From Web servers to integrated cluster solutions

This document is a comprehensive guide covering the IBM @server OpenPower 720 UNIX servers. We introduce major hardware offerings and discuss their prominent functions. Professionals wishing to acquire a better understanding of IBM @server OpenPower products should 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 documentation by providing a desktop reference that offers a detailed technical description of the OpenPower 720 system. This publication does not replace the latest @server marketing materials and tools. It is intended as an additional source of information that, together with existing sources, can be used to enhance your knowledge of IBM 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