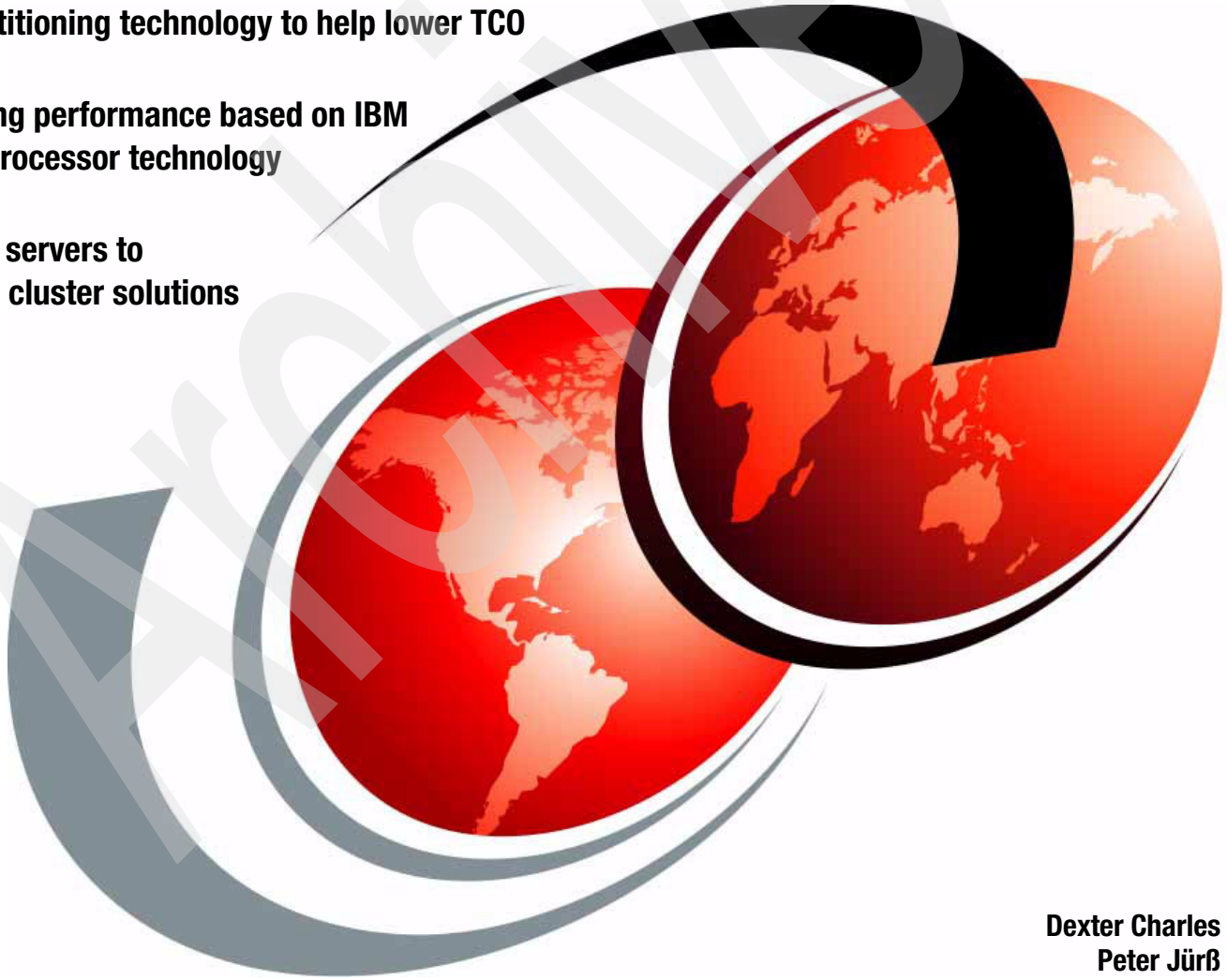
IBM

IBM @server OpenPower 710 Technical Overview and Introduction

- Finer system granulation using Micro-Partitioning technology to help lower TCO
- Outstanding performance based on IBM POWER5 processor technology
- From Web servers to integrated cluster solutions



Dexter Charles
Peter Jürß

Redpaper

ibm.com/redbooks

International Technical Support Organization

IBM @server OpenPower 710 Technical Overview and Introduction

June 2005

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

First Edition (June 2005)

This edition applies to the IBM @server OpenPower 710 server and Linux.

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.

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:

@server®	DB2®	POWER4™
@server®	DB2 Universal Database	POWER4+™
ibm.com®	Enterprise Storage Server®	POWER5™
iSeries™	Hypervisor™	Redbooks™
i5/OS™	HACMP™	RS/6000®
pSeries®	IBM®	Tivoli®
xSeries®	Micro-Partitioning™	Trackpoint®
z/OS®	OpenPower™	TotalStorage®
AIX 5L™	Perform™	WebSphere®
AIX®	PowerPC®	
Chipkill™	POWER™	

The following terms are trademarks of other companies:

Java and all Java-based trademarks and logos are trademarks or registered trademarks of Sun Microsystems, Inc. in the United States, other countries, or both.

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

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™ 710 server. 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:

- Clients
- 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 710 server.

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.

Dexter Charles is an IBM Accredited IT Specialist with seven years in the IT industry. Dexter has worked four years as a pSeries® Specialist with Americas Techline, providing pre-sales technical support for IBM @server® p5, pSeries, Linux® on POWER™, and WebSphere® sizing on pSeries platform. Dexter holds a number of IBM certifications including AIX® System Administration, pSeries Enterprise Technical Support, and AIX 5L™ Installation, Backup and System Recovery.

Peter Jürß is an IBM Accredited Senior IT Specialist for pSeries working for IBM EMEA Techline, Central Region in Germany, supporting IBM Sales and IBM Business Partners with technical pre-sales consulting. He has more than seven years of experience in the IT industry and five years of experience in the area of RS/6000® and pSeries systems and software environment.

The project that created this publication was managed by:
Scott Vetter

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

Arzu Gucer
International Technical Support Organization, Austin Center

John Hilburn, Carolyn K Jones, Thoi Nguyen, Herman Leonard, Tenley Jackson,
Rich LaFrance, Bill Mihaltse, Jim McGaughan, Sharon Dobbs
IBM U.S.

Derrick Daines, Dave Williams, Eric Rowen
IBM U.K.

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 clients.

Your efforts will help increase product acceptance and client 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 IBM Redbook form found at:

  ibm.com/redbooks

- Send your comments in an e-mail to:

  redbook@us.ibm.com

- Mail your comments to:

  IBM Corporation, International Technical Support Organization
  Dept. JN9B Building 905
  11501 Burnet Road
  Austin, Texas 78758-3493

1

General description

The IBM @server® OpenPower™ 710 rack-mount server (9123-710) is a powerful, dense 2-way-capable system that delivers the expandability and availability required for constrained data centers. The OpenPower 710 server's raw computing power positions it well for high-performance engineering and scientific workloads. Its system-level expandability and availability suits edge of network, general infrastructure, Web content serving, and application hosting services applications. The OpenPower 710 server comes in a 2U rack drawer package and is available in a 1-way or 2-way configuration.

The OpenPower 710 server's symmetric multiprocessor (SMP) uses a 1-way or 2-way, state-of-the-art, 64-bit, Silicon on Insulator, copper-based IBM POWER5™ microprocessor running at 1.65 GHz with 36 MB of Level 3 (L3) cache. The base 1 GB of main memory can be expanded to 32 GB for faster performance and exploitation of 64-bit addressing as used in large database applications.

The OpenPower 710 server contains five device bays. The five device bays are front-accessible; four bays are for hot-swap-capable disk drives and can accommodate up to 1.2 TB of disk storage. The fifth bay is used for a slimline DVD-ROM or DVD-RAM. Other integrated features include three 64-bit PCI-X slots, integrated service processor, integrated 10/100/1000 Mbps two-port Ethernet, two service processor communication, two USBs, and two HMC ports, integrated dual-channel Ultra320 SCSI controller, external SCSI port, hot-swappable power and cooling, and optional redundant power.

For partitioning, a Hardware Management Console (HMC) is required. The optional feature code (FC) 1965 POWER Hypervisor™ and Virtual I/O Server are also supported on the OpenPower 710 server, which supports dynamic LPAR[1], up to 20 micro-partitions using Micro-Partitioning™ technology, virtual Ethernet, as well as virtual SCSI.

IBM Cluster Systems Management (CSM) V1.4.1 for Linux on POWER is supported on the OpenPower 710 server.

The OpenPower 710 server is backed by a three-year limited warranty. Check with your IBM representative for particular warranty availbility in your region.

[1] Dynamic LPAR support varies by Linux distribution.

1.1 System specification

Table 1-1 lists the general system specifications of the OpenPower 710 server.

Table 1-1 IBM eServer OpenPower 710 server specifications

Description	Range
Operating temperature	5 to 35 degrees Celsius (41 to 95 degrees Farenheit)
Relative humidity	8 to 80 percent
Operating voltage	100–127 or 200–240 volts AC (auto-ranging)
Operating frequency	50/60 plus or minus 0.5 Hz
Maximum power consumption	475 watts
Maximum thermal output	1622 BTU/hr (British Thermal Unit)

1.2 Physical package

Table 1-2 lists the major physical attributes found on the OpenPower 710 server.

Table 1-2 IBM eServer OpenPower 710 server physical packaging

Dimension	
Height	89 mm (3.5 inches)
Width	483 mm (19 inches)
Depth	686 mm (27.0 inches)
Weight	
Minimum configuration	16.8 kg (37 lb)
Maximum configuration	23.2 kg (51 lb)

1.3 OpenPower 710 rack-mount server

Figure 1-1 shows the OpenPower 710 server that has been removed from a rack.

Figure 1-1 The OpenPower 710 rack-mount server

The OpenPower 710 server is a 2U-high, rack-mounted server, designed to be installed in a 19-inch rack. There is no deskside model available.

One of the following feature codes (FC) must be ordered along with the system:

- FC 7998 IBM Rack-mount Drawer Bezel
- FC 7999 OEM Rack-mount Drawer Bezel and Hardware

The OpenPower 710 server can be installed in either IBM or OEM racks. There is only one adjustable rack-mount drawer rail kit available for both IBM and OEM racks:

- FC 7166 IBM/OEM Rack-mount Drawer Rail Kit

It is possible to place up to 21 systems in an area of 644 mm (25.5 inches) x 1147 mm (45.2 inches) if one 42U-high rack is used. The 14 ft. length of the jumper cords (between the drawer and the PDU) will require additional cable management. The extra length of a cord not used in its routing between its two ends will have to be accounted for inside rack space.

Included with the OpenPower 710 rack-mounted server packaging are all of the components and instructions necessary to enable installation in a 19-inch rack.

Figure 1-2 shows a more detailed view of the OpenPower 710 rack-mount server, including connectors, location codes, SCSI IDs, and components.

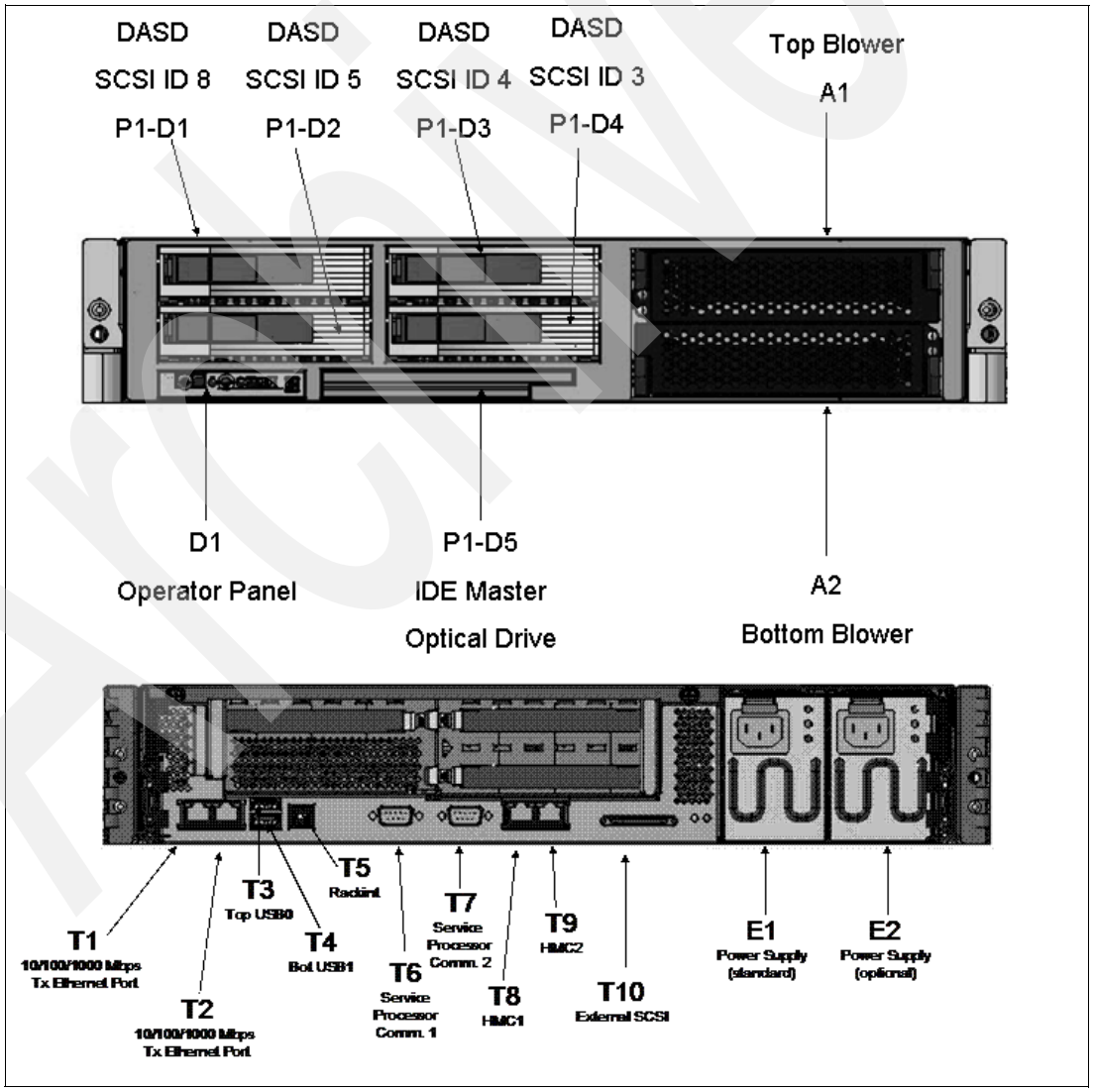


Figure 1-2 Front and rear view of the OpenPower 710 server including location codes

1.4 Minimum and optional features

The IBM @server OpenPower 710 server is based on a flexible, modular design featuring:

- One-way and 2-way symmetric multiprocessor (SMP) design using one POWER5 chip packaged in a processor module soldered directly to the system planar
- 1 GB of 266 MHz DDR1 ECC memory, expandable to 32 GB
- Four hot-swappable disk drive bays
- Three 64-bit, 3.3 volt, long, 133 MHz PCI-X slots
- One slimline media bay

The OpenPower 710 server supports the following integrated ports:

- Dual ported 10/100/1000 Ethernet
- Dual channel Ultra320 SCSI controller (one internal and one external VHDCI LVD connector)
- Service processor
- Hot-plug and redundant fans
- Two USB ports
- Two service processor communications ports
- Two HMC ports

The OpenPower 710 server supports the Linux® operating system (OS) and requires the following specific levels:

- SUSE LINUX Enterprise Server 9 (SLES 9) for POWER, or later
- Red Hat Enterprise Linux AS for POWER Version 3 (RHEL AS 3), or later

1.4.1 Processor features

The OpenPower 710 server features one POWER5 chip with one or two active processor cores running at 1.65 GHz. The OpenPower 710 server is available as a one-way or two-way system. Table 1-3 lists the available processor features.

Table 1-3 Available processor options

Feature code	Description
1963	1-way 1.65 GHz POWER5 Processor Card, 36 MB L3 Cache
1966	2-way 1.65 GHz POWER5 Processor Card, 36 MB L3 Cache

The OpenPower 710 POWER5 chip is mounted on the system planar and directly interfaced to the memory buffer SMI-II chips.

1.4.2 POWER Hypervisor technology

The POWER Hypervisor technology for the IBM @server OpenPower 710 server includes static and dynamic logical partitioning, as well as Micro-Partitioning and virtualization technology. Table 1-4 shows the required feature code.

Table 1-4 Advanced POWER Virtualization option

Feature code	Description
1965	POWER Hypervisor and Virtual I/O Server

For a detailed discussion on logical partitioning and virtualization, please refer to “The POWER Hypervisor and Virtual I/O Server technologies” on page 26.

1.4.3 Memory features

The minimum memory requirement for the OpenPower 710 server is 1 GB, and the maximum capacity is 32 GB. Memory DIMMs are installed into eight DIMM sockets located on the system planar. The supported memory is 266 MHz DDR1 DIMMs. Table 1-5 lists the available memory features.

Table 1-5 Memory feature codes

Feature code	Description
1949	1024 MB (2 x 512 MB) DIMMS, 208-pin, 266 MHz DDR SDRAM
1950	2048 MB (2 x 1024 MB) DIMMS, 208-pin, 266 MHz DDR SDRAM
1951	4096 MB (2 x 2048 MB) DIMMS, 208-pin, 266 MHz DDR SDRAM
1952	8192 MB (2 x 4096 MB) DIMMS, 208-pin, 266 MHz DDR SDRAM

1.4.4 Disk and media features

The OpenPower 710 server features four disk bays and one slimline media bay. The minimum configuration requires at least one disk drive. Table 1-6 shows the disk drive feature codes that each bay can contain.

Table 1-6 Hot-swappable disk drive options

Feature code	Description
1970	36.4 GB 15,000 rpm Ultra320 SCSI Disk Drive Assembly
1968	73.4 GB 10,000 rpm Ultra320 SCSI Disk Drive Assembly
1971	73.4 GB 15,000 rpm Ultra320 SCSI Disk Drive Assembly
1969	146.8 GB 10,000 rpm Ultra320 SCSI Disk Drive Assembly
1972	146.8 GB 15,000 rpm Ultra320 SCSI Disk Drive Assembly
1973	300 GB 10,000 rpm Ultra320 SCSI Disk Drive Assembly

A DVD-ROM or DVD-RAM drive can be installed in the slimline bay:

- DVD-ROM drive, FC 1994 (default)
- DVD-RAM drive, FC 1993

A logical partition running a supported release of the Linux operating system requires a DVD drive to provide a method to run the hardware diagnostics from the CD.

1.4.5 USB diskette drive

In some situations, an external USB 1.44 MB diskette drive for OpenPower 710 servers (FC 2591) is helpful. This 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 controller. The same controller can share a USB mouse and keyboard.

1.4.6 Hardware Management Console models

The Hardware Management Console is a dedicated workstation that allows you to configure and manage partitions. The hardware management application helps you configure and partition the server through a graphical user interface (GUI).

Functions performed by the HMC includes:

- Creating and maintaining a multiple-partition environment
- Displaying a virtual operating system session terminal for each partition
- Displaying a virtual operator panel of contents for each partition
- Detecting, reporting, and storing changes in hardware conditions
- Powering managed systems on and off
- Acting as a service local point for service representatives to determine an appropriate service strategy

See “Hardware Management Console” on page 33 for detailed information on the HMC.

Table 1-7 lists the HMC options for POWER5 processor-based systems available at the time this paper was written.

Table 1-7 Supported HMC

Type-model	Description
7310-C04	IBM 7310 Model C04 Desktop Hardware Management Console
7310-CR3	IBM 7310 Model CR3 Rack-Mount Hardware Management Console

1.5 Express Product Offerings

New specially priced Express Product Offerings are now available for the OpenPower 710 servers. These Express Product Offerings feature popular, easy-to-order preconfigured servers with attractive financial incentives. Express Product Offerings are available only as an initial order.

OpenPower Express servers are complemented by pre-tested solutions that provide recommended system configurations with installation and sizing aids, for a range of business requirements. Built on the solid base of OpenPower servers, the Linux operating system, and popular application software packages, these offerings are designed to help smaller and mid-sized companies solve a variety of business problems—application consolidation, e-mail security, and infrastructure for Enterprise Resource Planning (ERP).

Available solutions include:

- IBM @server OpenPower Network E-Mail Security Express Solution
- IBM @server OpenPower and IBM DB2® Universal Database for SAP Solution

- IBM @server OpenPower Consolidation Express Solution

Express Product Offerings consist of the following processor requirements: Either a one-way POWER5 1.65 GHz processor (FC 1939) or two-way POWER5 1.65 GHz processor (FC 1941), and a defined minimum configuration. If any of the features in an Express Product Offering are changed, the Express Product Offering identification feature (FC 936X) will be removed from the order.

Table 1-8 lists the available Express Product Offerings configurations and minimum requirements.

Table 1-8 Express Product Offerings configurations

Express Offering Identifier	Processors	Memory (MB)	Disk	Config Number (SKU)[a]
Entry Offering 9361	1-way, FC 1939, 1.65 GHz	1 x 2048 (FC 1950)	2 x 73.4 GB disk drive, FC 1968	91231D1
Value Offering 9262	2-way, FC 1941, 1.65 GHz	2 x 2048 (FC 1950)	2 x 73.4 GB disk drive, FC 1968	91231D2
Consolidation Offering[b] 9363	2-way, FC 1941, 1.65 GHz	2 x 4096 (FC 1951)	2 x 73.4 GB disk drive, FC 1971	91231D3

a. SKU Identifier = Stock Keeping Unit Identifier.
b. The Consolidation Offering includes 2 x FC 1965 (POWER Hypervisor and Virtual I/O Server) as a minimum requirement.

The Express Product Offerings also includes:

- DVD-ROM (FC 1994)
- IBM Rack-mount drawer bezel and hardware (FC 7998)
- Rack-mount drawer rail kit (FC 7166)
- Power supply, 700 watt (FC 7989)
- Language group specify (FC 9300 or 97xx)
- Power cord

Note: Keyboard, mouse, operating system (OS) license, and OS media are not included in these configurations.

Note: If a build-to-order (BTO) configuration meets all the requirements of an OpenPower Express configuration, Express configuration pricing will be applied.

When an OpenPower Express configuration is ordered, the nine-digit reference number called a SKU Identifier will be printed on the packing list and on a label (readable and barcode) on the outside of the box. Also, it will appear on invoices and billing statements. The SKU Identifier helps improve IBM distributors' inventory management. The SKU Identifier numbers for OpenPower Express configurations are listed in Table 1-8 above.

Note: Only Express Product Offerings configurations will have SKU Identifiers. No BTO configuration, even if it meets the definition of an Express Product Offering configuration, will have an SKU Identifier. Any modifications to an Express hardware configuration will suppress the SKU Identifier.

1.6 System racks

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

If a IBM ℯserver OpenPower system is to be installed in a non-IBM rack or cabinet, you must ensure that the rack conforms to the EIA[2] standard EIA-310-D (see “OEM rack” on page 12).

Note: It is the client’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.

1.6.1 IBM RS/6000 7014 Model T00 Enterprise Rack

The 1.8-meter (71-inch) Model T00 is compatible for past and present IBM ℯserver p5, OpenPower, and pSeries systems that fit into 19-inch 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 business black or optional white color in OEM format.
- Increased power distribution and weight capacity.
- 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.
- 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 PDU bays (see Figure 1-3 on page 10), but others can fit inside the rack. See “AC Power Distribution Unit and rack content” on page 9.
- An optional rack status beacon (FC 4690). This beacon is designed to be placed on top of a rack and cabled to servers, such as an OpenPower 710 server, and other components inside the rack. Servers can be programmed to illuminate the beacon in response to a detected problem or changes in system status.

[2] 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.

- A rack status beacon junction box (FC 4693) should be used to connect multiple servers to the beacon. This feature provides six input connectors and one output connector for the rack. To connect the servers or other components to the junction box or the junction box to the rack, status beacon cables (FC 4691) are necessary. Multiple junction boxes can be linked together in a series using daisy chain cables (FC 4692).
- Weights:
  - T00 base empty rack: 244 kg (535 pounds)
  - T00 full rack: 816 kg (1795 pounds)

1.6.2 IBM RS/6000 7014 Model T42 Enterprise Rack

The 2.0-meter (79.3-inch) Model T42 is the rack that will address a client requirement for 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 (42 U) of usable space (6 U of additional space)
- Model T42 supports AC only
- Weights:
  - T42 base empty rack: 261 kg (575 pounds)
  - T42 full rack: 930 kg (2045 pounds)

1.6.3 AC Power Distribution Unit and rack content

For rack models T00 and T42, 12-outlet PDUs (FC 9188 and FC 7188) are available.

Four PDUs can be mounted vertically in the 7014 racks. See Figure 1-3 for placement of the four vertically mounted PDUs. In the rear of the rack, two additional PDUs can be installed horizontally in the T00 and three in the T42 rack. The four vertical mounting locations will be filled first. Mounting PDUs horizontally consumes 1U per PDU and reduces the space available for other racked components. When mounting PDUs horizontally, it is recommended that fillers be used in the EIA units occupied by these PDUs to facilitate proper air-flow and ventilation in the rack.

For detailed power cord requirements and power cord feature codes, see the publication *IBM @server Planning Information,* SA38-0508. An online copy can be found at **Maps of pSeries books to the information center → Planning → Printable PDFs → Planning**:

http://publib16.boulder.ibm.com/infocenter/eserver/v1r2s/en_US/index.htm

Note: Ensure the appropriate power cord feature is configured to support the power being supplied.

The Base/Side Mount Universal PDU (FC 9188) and the optional, additional, Universal PDU (FC 7188), support a wide range of country requirements and electrical power specifications. It receives power through a UTG0247 connector. Each PDU requires one PDU to wall power cord. Nine power cord features are available for different countries and applications by varying the PDU to Wall Power Cord, which must be ordered separately. Each power cord provides the unique design characteristics for the different power requirements. To match new power requirements and save previous investments, these power cords can be requested with an initial order of the rack, or with a later upgrade of the rack features.

The PDU has twelve IEC 320-C13 outlets. There are six groups of two outlets fed by six circuit breakers. Each outlet is rated up to 10 amps, but each group of two outlets is fed from one 15 amps circuit breaker.

Note: The maximum number of OpenPower 710 servers supported on a 30 amp service (PDU derated to 24 amps) is nine per PDU.

The maximum number of OpenPower 710 servers supported on a 60 amp service (PDU derated to 48 amps) is 12 per PDU.

The Universal PDUs are compatible with previous IBM @server p5, OpenPower, and pSeries models.

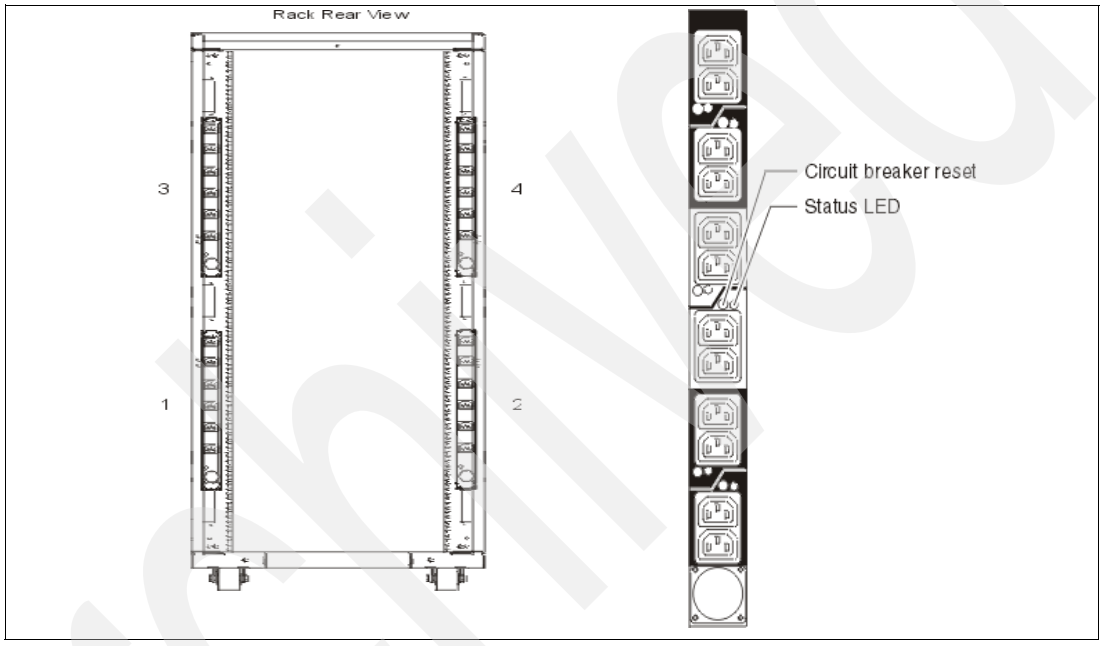


Figure 1-3 PDU placement and PDU view

1.6.4 Rack-mounting rules for OpenPower 710 server

The OpenPower 710 server is a 2U rack-mounted server drawer. The primary rules that should be followed when mounting the OpenPower 710 server into a rack are:

- The OpenPower 710 server is designed to be placed at any location in the rack. For rack stability, it is advisable to start filling a rack from the bottom.
- Any remaining space in the rack can be used to install other systems or peripherals, provided that the maximum permissible weight of the rack is not exceeded and the installation rules for these devices are followed.
- Before placing a OpenPower 710 server into the service position, it is essential that the rack manufacturer's safety instructions are followed regarding rack stability.
- A maximum of 18 model OpenPower 710 servers fit in the T00 rack, and 21 model OpenPower 710 servers in T42 rack.

  The 14 ft. length of the jumper cords (between the drawer and the PDU) will require additional cable management. The extra length of a cord not used in its routing between its two ends will have to be accounted for inside rack space.

1.6.5 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 710 servers or other IBM @server p5 and pSeries 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 p5, OpenPower, and pSeries 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) of physical data storage capacity per cartridge.
- Digital Data Storage (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 that uses a 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[3] (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 can 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 pels[4] 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 Travel Keyboard mounts in the rack keyboard tray (Integrated Trackpoint® and UltraNav).

IBM PS/2 Travel Keyboards are supported on the 7316-TF3 for use in configurations where only PS/2 keyboard ports are available.

The IBM 7316-TF3 Flat Panel Console Kit provides an option for the IBM USB Travel Keyboards with UltraNav. The USB keyboard allows the 7316-TF3 to be connected to systems that do not have PS/2 keyboard ports. The IBM USB Travel Keyboard may be direct attached to an available integrated USB port or a supported USB adapter (2738) on IBM @server OpenPower servers or IBM 7310-CR3 and 7315-CR3 Hardware Management Consoles.

The Netbay LCM (Keyboard/Video/Mouse) Switch (FC 4202) allows users single-point access and control of up to 64 servers from a single console. The Netbay LCM Switch has a maximum video resolution of 1600 x 1280 and mounts in a 1U drawer behind the 7316-TF3

[3] Initial Program Load
[4] Picture elements

monitor. A minimum of one LCM feature (FC 4268) or USB feature (FC 4269) is required with a Netbay LCM Switch (FC 4202). Each feature can support up to four systems. When connecting to an OpenPower 710 server, FC 4269 provides connection to the POWER5 USB ports. Only the PS/2 keyboard is supported when attaching the 7316-TF3 to the LCM Switch.

The following should be considered when selecting the LCM Switch:

- The KCO cable (FC 4268) is used with systems with PS/2 style keyboard, display, and mouse ports.
- The USB cable (FC 4269) is used with systems with USB keyboard or mouse ports.
- The switch offers four ports for server connections. Each port in the switch can connect a maximum of 16 systems.
  - One KCO cable (FC 4268) or USB cable (FC 4269) is required for every four systems supported on the switch.
  - A maximum of 16 KCO cables or USB cables per port may be used with the Netbay LCM Switch (FC 4202) to connect up to 64 servers.

Note: A server microcode update may be required on installed systems for boot-time SMS menu support of the USB keyboards. The update may also be required for the LCM switch on the 7316-TF3 console (FC 4202). Microcode updates are located at the URL below.

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

We recommend 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 2849) to be installed in the server, or other graphic accelerator, if supported.

Hardware Management Console 7310 Model CR3

The 7310 Model CR3 Hardware Management Console (HMC) is a 1U, 19-inch rack-mountable drawer supported in the 7014 Model T00 and T42 racks. For additional HMC specifications see “Hardware Management Console” on page 33.

1.6.6 OEM rack

The OpenPower 710 server can be installed in a suitable OEM rack, provided that the rack conforms to the EIA-310-D standard for 19-inch racks. This standard is published by the Electrical Industries Alliance, and a summary of this standard is available in the publication *IBM @server Planning,* SA38-0508.

The key points mentioned in this documentation are as follows:

- The front rack opening must be 451 mm wide + 0.75 mm (17.75 inches + 0.03 inches), and the rail-mounting holes must be 465 mm + 0.8 mm (18.3 inches + 0.03 inches) apart on center (horizontal width between vertical columns of holes on the two front-mounting flanges and on the two rear-mounting flanges). See Figure 1-4 for a top view showing specifications dimensions.
- The vertical distance between mounting holes must consist of sets of three holes spaced (from bottom to top) 15.9 mm (0.625 inches), 15.9 mm (0.625 inches), and 12.67 mm (0.5 inches) on center (making each three-hole set of vertical hole spacing 44.45 mm (1.75 inches) apart on center). Rail-mounting holes must be 7.1 mm + 0.1 mm (0.28 inches + 0.004 inches) in diameter. See Figure 1-5 and Figure 1-6 for top and bottom front specification dimensions.

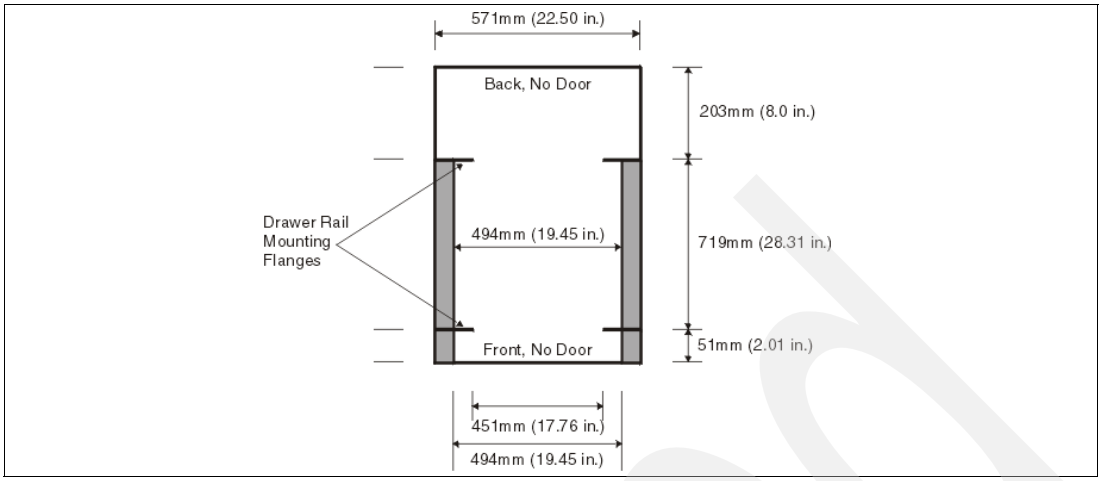


Figure 1-4 Top view of non-IBM rack specifications dimensions

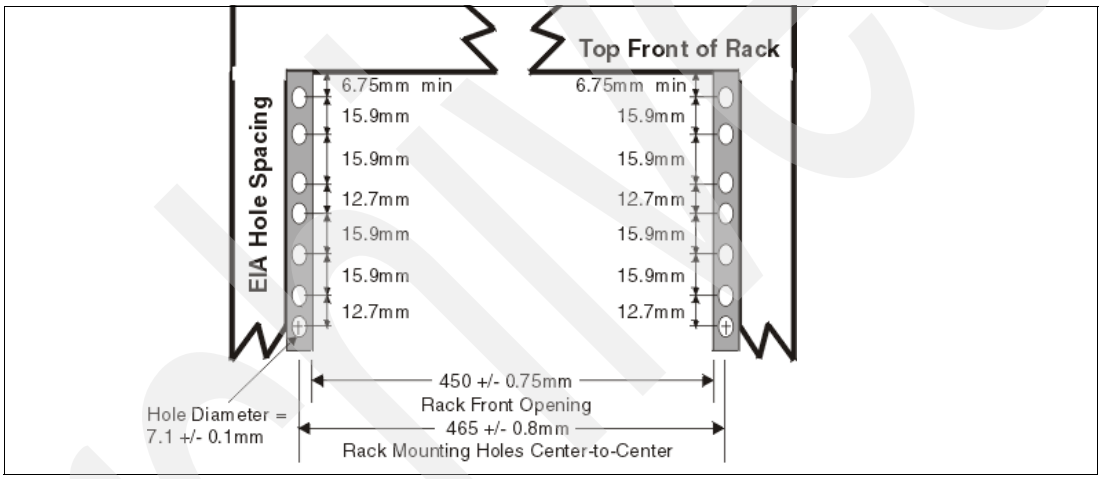


Figure 1-5 Rack specifications dimensions, top front view

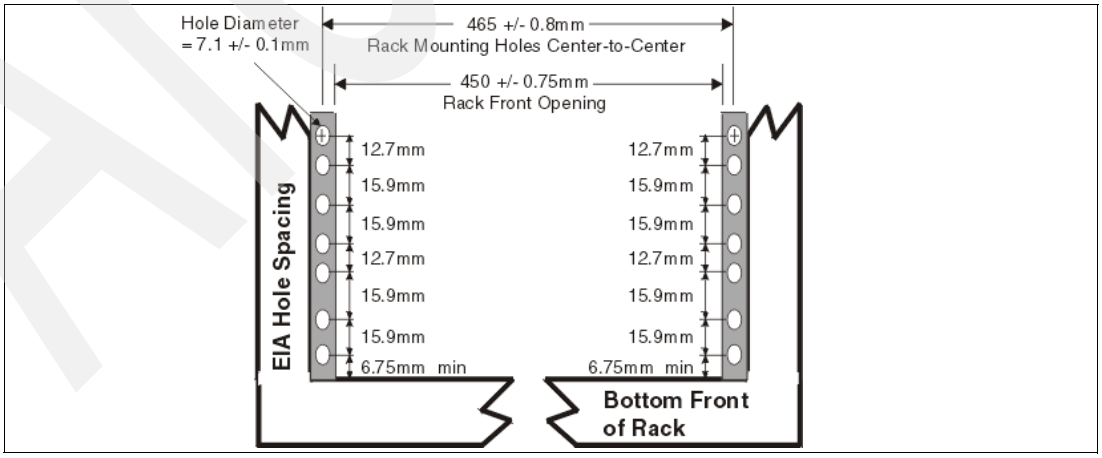


Figure 1-6 Rack specifications dimensions, bottom front view

- It might be necessary to supply additional hardware, such as fasteners, for use in some manufacturer's racks.

- The rack or cabinet must be capable of supporting an average load of 15.9 kg (35 lb.) of product weight per EIA unit.
- The rack or cabinet must be compatible with drawer mounting rails, including a secure and snug fit of the rail-mounting pins and screws into the rack or cabinet rail support hole.

Note: The OEM rack must only support AC powered drawers. It is strongly recommended to use a power distribution unit (PDU) that meets the same specifications as IBM PDUs to supply rack power. Rack or cabinet power distribution devices(s) must meet the drawer power requirements, as well as that of any additional products that will be connected to the same power distribution device.

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. We always recommend that you obtain real-world performance measurements using production workloads.

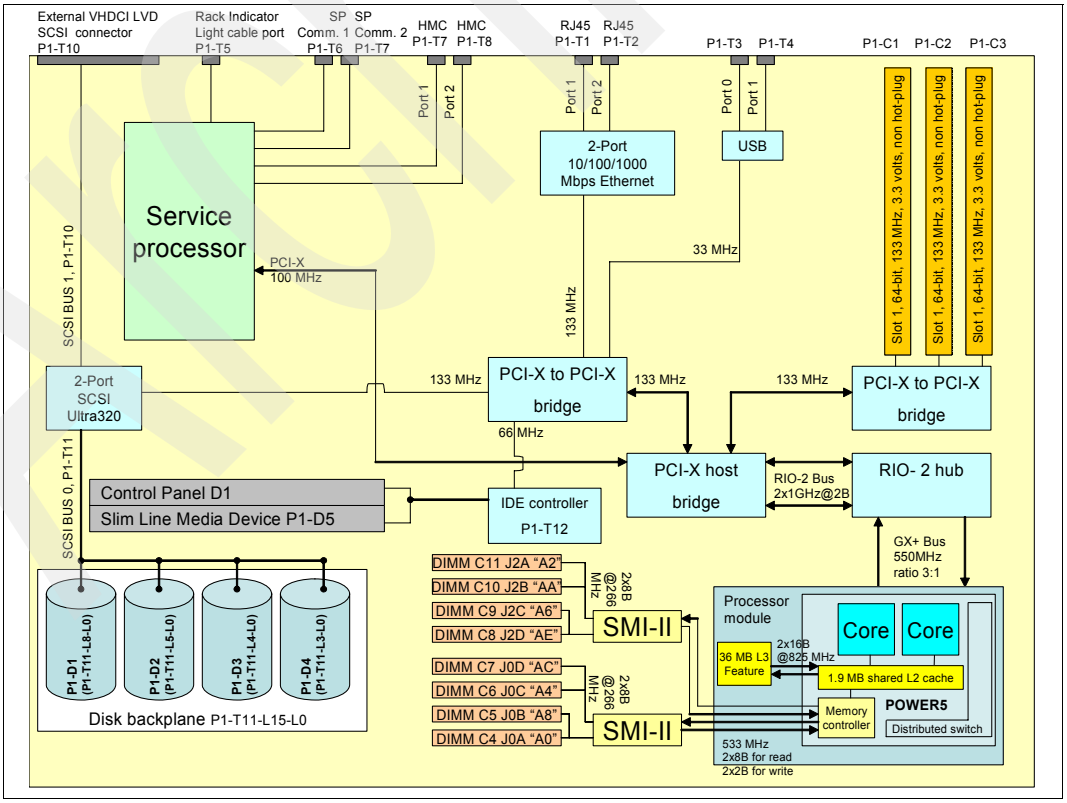


Figure 2-1 OpenPower 710 server logic data flow

2.1 The POWER5 chip

The POWER5 chip features single and simultaneous multi-threaded execution, providing higher performance in the single-threaded mode than its POWER4™ predecessor at equivalent frequencies. The POWER5 processor maintains both binary and architectural compatibility with existing POWER4 processor-based systems and is designed to ensure that binaries continue executing properly and application optimizations carry forward to newer systems. Table 2-1 shows highlights and changes between the POWER4 and the POWER5 processor.

Table 2-1 POWER4 to POWER5 comparison

	POWER4 design	POWER5 design
L1 data cache	2-way set associative FIFO[a]	4-way set associative LRU[b]
L2 cache	8-way set associative 1.44 MB	10-way set associative 1.9 MB
L3 cache	32 MB 118 clock cycles	36 MB ~80 clock cycles
Memory bandwidth	4 GB/second /chip	~16 GB/second /chip
Simultaneous multi-threading	No	Yes
Processor addressing	1 processor	1/10th of processor
Dynamic power management	No	Yes
Size	412 mm	389 mm

a. FIFO stands for First In First Out
b. LRU stands for Least Recently Used

POWER5 design provides additional enhancements such as virtualization, reliability, availability, and serviceability (RAS) features at both chip and system levels.

Key enhancements introduced into the POWER5 processor and system design include:

- Simultaneous multi-threading
- Dynamic resource balancing to efficiently allocate system resources to each thread
- Software-controlled thread prioritization
- Dynamic power management to reduce power consumption without affecting performance
- Micro-Partitioning technology
- Virtual storage, virtual Ethernet
- Enhanced scalability, parallelism
- Enhanced memory subsystem

Figure 2-2 on page 17 shows the high-level structures of POWER5 processor-based systems. POWER5 processor 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.

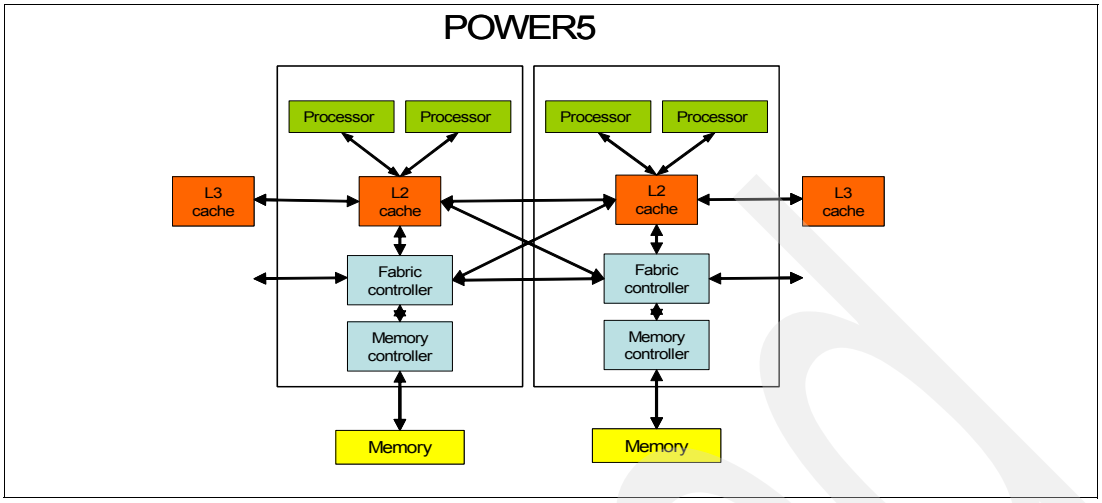


Figure 2-2 POWER5 system structure

Because of the higher transistor density of the POWER5 0.13-µm technology, it was possible to move the memory controller on the processor chip and eliminate a chip previously needed for the memory controller function. These changes in the POWER5 processor also have the significant side benefits of reducing latency to the L3 cache and main memory, as well as reducing the number of chips necessary to build a system.

2.1.1 Simultaneous multi-threading

As a requirement for performance improvements at the application level, simultaneous multi-threading functionality is embedded in the POWER5 chip technology. Applications developed to use process-level parallelism (multi-tasking) and thread-level parallelism (multi-threads) can shorten their overall execution time. Simultaneous multi-threading 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.

The simultaneous multi-threading 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), which are essential for out-of-order execution, and then vital for simultaneous multi-threading.

If simultaneous multi-threading is activated:

- More instructions can be executed at the same time.
- The operating system views twice the number of physical processors installed in the system.
- Provides support in mixed environments:
  - Capped and uncapped partitions
  - Virtual partitions
  - Dedicated partitions
  - Single partition systems

Note: Simultaneous multi-threading is supported on POWER5 processor-based systems running Linux operating system-based systems at an appropriate level.

The simultaneous multi-threading policy is controlled by the operating system and is thus partition specific.

For Linux, an additional boot option must be set to activate simultaneous multi-threading after a reboot.

Simultaneous multi-threading features

To improve simultaneous multi-threading performance for various workloads and provide robust quality of service, the POWER5 processor provides two features:

- Dynamic resource balancing

  Dynamic resource balancing is designed 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 (a thread reached threshold of L2 cache misses will be throttled to allow other threads to pass the stalled thread).

- Adjustable thread priority

  Adjustable thread priority that allows software to determine when one thread should have a greater (or lesser) share of execution resources. The POWER5 processor supports eight software-controlled priority levels for each thread.

Single threading operation

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 processor supports the single threading 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 a single threading operation may provide more throughput.

2.1.2 Dynamic power management

In current Complementary Metal Oxide Semiconductor (CMOS) technologies, chip power is one of the most important design parameters. With the introduction of simultaneous multi-threading, 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 are dispatched 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.

2.2 Processor and cache

The OpenPower 710 server supports one processor module (either a 1-way or 2-way) and integrated 36 MB L3 cache module.

Note: Since the POWER5 processor modules are directly soldered to the system planar, special care must be taken for sizing and selecting the ideal CPU configuration.

The storage structure for the POWER5 chip is a distributed memory architecture that provides high memory bandwidth, although each processor can address all memory and sees a single shared memory resource. They are interfaced to eight memory slots, controlled by two SMI-2 controllers, which are located in close physical proximity to the processor modules. I/O connects to the OpenPower 710 processor module using the GX+ bus. The processor module provides a single GX+ bus. The GX+ bus provides an interface to I/O devices through the RIO-2 connections.

The theoretical maximum troughput of the L3 cache is 16-byte read, 16-byte write at a bus frequency of 825 MHz, which equates to 26400 MB/second or 25.78 GB/s.

The processor core contains a single processor module and the local memory storage subsystem for that processor module. Figure 2-3 shows a POWER5 processor core layout view.

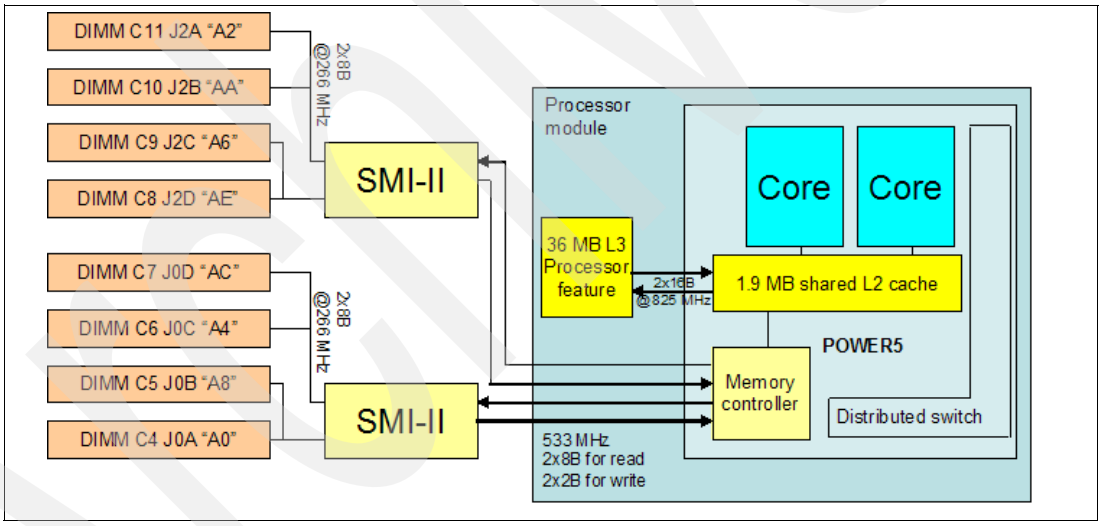


Figure 2-3 POWER5 processor core with DDR1 memory socket layout view

2.2.1 Available processor speeds

At the time of writing, the IBM @server OpenPower 710 server operates at a processor clock rate of 1.65 GHz.

2.3 Memory subsystem

The OpenPower 710 server offers pluggable DIMMs for memory. The system planar provides eight slots for up to eight pluggable DIMMs. The minimum memory for a OpenPower 710 server is 1 GB and 32 GB as maximum installable memory option. Figure 2-4 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 710 server are listed in “Memory features” on page 5.

Memory DIMMs must be ordered and installed in pairs. Memory feature numbers may be mixed within a system. The DIMMs slots are accessed by first removing the PCI riser book.

Memory is installed in the following order: J2A and J0A, J2C and J0C, J2B and J0B, and J2D and J0D. Memory must be balanced across the DIMM slots. The Service Information label, located on the top cover of the system, provides memory DIMMs slot location information.

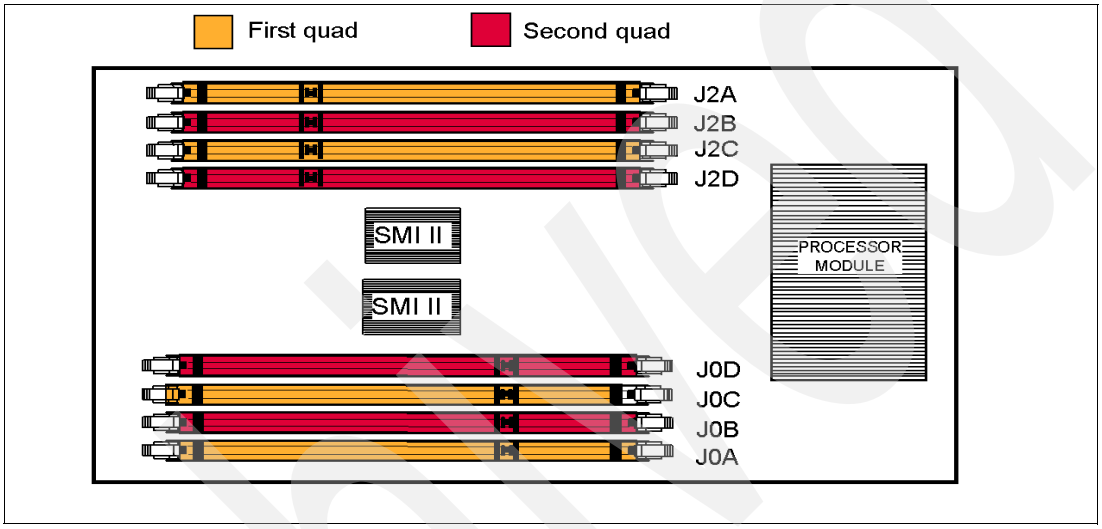


Figure 2-4 Memory placement for the OpenPower 710 server

2.3.2 Memory restriction

OEM memory is not supported by IBM on the OpenPower 710 server. OEM memory is not certified by IBM for the use in OpenPower servers. If the OpenPower 710 server is populated with OEM memory, you could experience unexpected and unpredictable behavior, especially when the system is using Micro-Partitioning technology.

All IBM memory is identified by an IBM logo and a white label printed with a barcode and an alphanumeric string, illustrated in Figure 2-5.

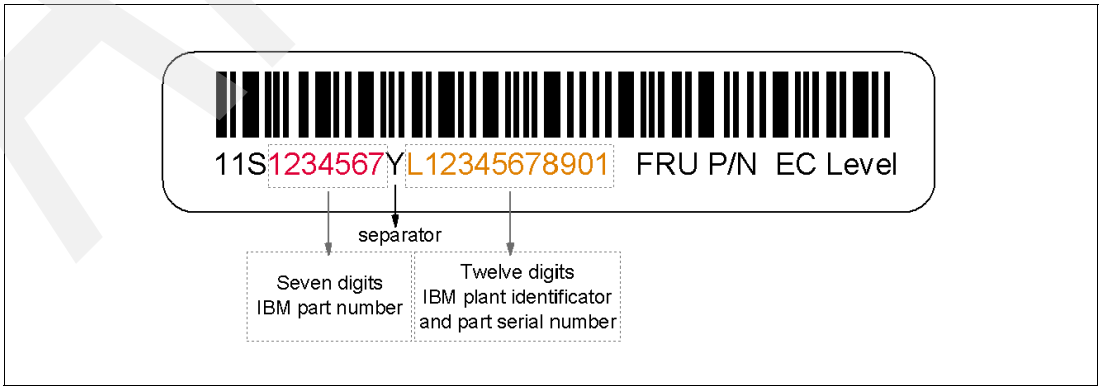


Figure 2-5 IBM memory certification label

2.3.3 Memory throughput

The memory subsystem throughput is based on the speed of the memory. An elastic interface, contained in the POWER5 chip, buffers reads and writes to and from memory and the processor. There are two SMIs, each with a single 8-byte read and 2-byte write DDR bus to the processor. A DDR bus allows double reads or writes per clock cycle. Since 266 MHz memory is installed (operating at 266.5 MHz), the throughput is (16 x 2 x 266.5) + (4 x 2 x 266.5) or 10660 MB/second or 10.41 GB/second between the processor and memory controller. These values are maximum theoretical throughputs for comparison purposes only.

There are four 8-byte paths to the memory DIMMs from the SMIs; therefore, the throughput is 8.32 GB/second.

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

2.4 System buses

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

2.4.1 GX+ and RIO-2 buses

The processor module provides a GX+ bus that is used to connect to the I/O subsystem. GX+ bus clock frequency is 550 MHz with a CPU to GX+ ratio of 3:1.

The GX+ bus is connected to an internal Remote I/O-2 bus on the system board through a RIO-2 hub. Two RIO-2 buses are available inside the system. Each RIO-2 bus provides 1 byte at 1 GHz in each direction, which leads to a theoretical maximum bandwidth of 4 GB/s.

Note: The OpenPower 710 server has no external RIO-2 ports and therefore additional external storage must be attached using other connections, such as network or SCSI.

2.5 Internal I/O subsystem

The internal I/O subsystem and the service processor reside directly on the system planar. There is an internal RIO-2 bus imbedded in the system planar. The system planar contains both the 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 RIO-2 hub chip are used for the integrated I/O, and the remaining two ports are unused.

The PCI -X riser card provides three PCI-X slots and is plugged directly into the system planar. The PCI-X riser card is connected exclusively to one of the two PCI-X to PCI-X bridges. The remaining integrated PCI-X devices interface to the second PCI-X to PCI-X bridge. Both PCI-X to PCI-X bridges are connected to the primary PCI-X buses on the PCI-X host bridge chip.

All PCI-X slots (1 trough 3) can accept long PCI-X or PCI cards. They are all 64-bit, 3.3 volts, 133 MHz.

2.5.1 PCI-X slots and adapter

PCI-X, where the X stands for extended, is an enhanced PCI bus, delivering a theoretical peak bandwidth of up to 1 GB/sec, running a 64-bit bus at 133 MHz. PCI-X is backward compatible, so the OpenPower 710 servers can support existing 3.3 volt PCI adapters.

Restriction: The integrated PCI-X card slots do not support hot-plug.

2.5.2 LAN adapters

When an OpenPower 710 server is connected to a local area network (LAN), the internal dual port 10/100/1000 Mbps RJ-45 Ethernet controller integrated on the system planar can be used. Except for the VPD module, the integrated dual 10/100/1000 Mbps Ethernet subsystem is identical to the FC 5706 IBM 2-Port 10/100/1000 Base-TX Ethernet PCI-X Adapter.

Table 2-2 provides a list of additional LAN adapters available at the time of writing. IBM supports an installation with Network Installation Manager (NIM) using Ethernet adapters (CHRP[1] is the platform type).

Table 2-2 Available LAN adapter

Feature code	Adapter description	Slot priority	Size	Max
1985	IBM 10/100 Mbps Ethernet PCI Adapter II	1, 2, 3	Short	3
1978	IBM Gigabit Ethernet-SX PCI-X Adapter	1, 2, 3	Short	3
1979	IBM 10/100/1000 Base-TX Ethernet PCI-X Adapter	1, 2, 3	Short	3
1983	IBM 2-Port 10/100/1000 Base-TX Ethernet PCI-X Adapter	1, 2, 3	Short	3
1984	IBM 2-Port Gigabit Ethernet-SX PCI-X Adapter	1, 2, 3	Short	3
1981	IBM 10 Gigabit Ethernet-SR PCI-X Adapter[a]	1, 2, 3	Short	1
1982	IBM 10 Gigabit Ethernet-LR PCI-X Adapter[a]	1, 2, 3	Short	1

a. Check with your IBM representative for the latest availability and support using Red Hat Enterprise Linux AS for POWER Version 4 or later.

2.5.3 SCSI adapters

The OpenPower 710 server has one internal dual port Ultra320 capable SCSI adapter integrated on the system planar. Unlike the other IBM eserver p5 or OpenPower models, the OpenPower 710 server uses only one of the available SCSI busses for internal connection to the four disk drives. The second SCSI bus has an external VHDCI LVD connector on the rear of the system.

Note: It is not possible to add hardware RAID capabilities to the onboard adapter.

Table 2-3 shows a list of additional SCSI adapters available at the time of writing. All listed adapters can be used as boot adapters.

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

Table 2-3 Available SCSI adapters

Feature code	Adapter description	Slot priority	Size	Max
1975	PCI-X Dual Channel Ultra320 SCSI RAID Adapter	1, 2, 3	Long	3
1974	PCI-X Dual Channel Ultra320 SCSI Adapter	1, 2, 3	Short	3

2.5.4 Graphic accelerator

The OpenPower 710 server supports one enhanced POWER GXT135P 2D graphic accelerator (FC 1980). It can be configured to operate in either 8-bit or 24-bit color modes. This adapter supports both analog and digital monitors. The adapter requires one short 32-bit or 64-bit PCI-X slot. The GXT135P is supported for SMS, firmware, and other functions, as well as when a Linux operating system starts the X11-based graphical user interface (GUI).

2.5.5 Service processor communication ports

The service processor communications ports S1 and S2, at the rear of the system, provide a virtual serial connection.

This virtual connection provides support on a non HMC-managed server. When no HMC is attached, these ports can be used for a local console, remote console, and call home modem attachment. The virtual ports do not support serial devices like printers or uninterruptible power supplies (UPS).

When a HMC is connected, the S1 and S2 ports are disabled.

If serial ports are needed, optional PCI adapters are available. Table 2-4 shows a list of available serial adapters available at the time of writing.

Table 2-4 Available serial adapter

Feature code	Adapter description	Slot priority	Size	Max
5723	2-Port Asynchronous EIA-232 PCI Adapter	1, 2, 3	Short	2

2.6 Internal storage

There is one integrated dual channel Ultra320 SCSI controller managed by an EADS-X chip, that is used to drive the internal disk drives. The OpenPower 710 server supports four bays that are designed for hot-swappable disk drives. The disk drive backplane docks directly to the system planar. The virtual SCSI Enclosure Services (VSES) hot-swappable control functions are provided by the integrated Ultra320 SCSI controller.

The four internal drives are on SCSI bus 0, which is connected to the internal port on the integrated Ultra320 SCSI controller. See “Dynamic LPAR minimum requirements” on page 30 for dynamic LPAR considerations.

2.6.1 Internal media devices

The OpenPower 710 server provides one slimline media bay for mandatory DVD drives. Table 2-5 lists available and supported internal media devices. The DVD device or an Internet connection can be used to maintain or update system microcode to the latest required level.

The control panel and DVD drive share the same bay. The front bezel cover must be removed to remove the DVD device and the control panel/media bay. This control panel/media bay is controlled by the integrated IDE controller.

Table 2-5 Available and supported internal media devices

Feature code	Description
1994	IDE Slimline DVD-ROM Drive
1993	IBM 4.7 GB IDE Slimline DVD-RAM Drive

2.6.2 Internal hot-swappable SCSI disks

The OpenPower 710 server can have up to four hot-swappable disk drives. The hot-swap process is controlled by the SCSI enclosure service (SES), which is provided by the integrated SCSI Ultra320 controller. Table 2-6 lists available and supported hot-swappable drives.

Note: The hot-swappable disk drive is only available with Linux 2.6 kernel. An OpenPower 710 server not running the Linux 2.6 kernel must be shut down and powered off before you replace any disk drives.

Table 2-6 Hot-swappable disk drive options

Feature code	Description
1970	36.4 GB 15,000 rpm Ultra320 SCSI Disk Drive Assembly
1968	73.4 GB 10,000 rpm Ultra320 SCSI Disk Drive Assembly
1971	73.4 GB 15,000 rpm Ultra320 SCSI Disk Drive Assembly
1969	146.8 GB 10,000 rpm Ultra320 SCSI Disk Drive Assembly
1972	146.8 GB 15,000 rpm Ultra320 SCSI Disk Drive Assembly
1973	300 GB 10,000 rpm Ultra320 SCSI Disk Drive Assembly

The system configuration shipped will have the first two SCSI disks installed in disk drive SCSI IDs 8 and 5. The drives at ID 8 and ID 5 are hardwired to spin up immediately during startup sequencing. The remaining drives will spin up under software control (typically 5 seconds intervals). The disk drive placement priority is SCSI ID 8, 5, 4, and then 3. See Figure 1-2 on page 3 for each SCSI ID location.

Prior to the hot-swap of a disk in the hot-swappable capable bay, all necessary operating system actions must be undertaken to ensure that the disk is capable of being deconfigured. After the disk drive has been deconfigured, the SCSI enclosure device will power off the bay, enabling safe removal of the disk. You should ensure that the appropriate planning has been given to any operating-system-related disk layout.

Note: We recommend that you use the following procedure after the disk has been deconfigred, when removing a hot-swappable disk:

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

2.7 External disk subsystem

The OpenPower 710 server has internal hot-swappable drives. Specific client requirements can be satisfied with the several external disks possibilities that the OpenPower 710 server supports.

Note: External I/O drawers 7311-D11 and 7311-D20 are not supported on the OpenPower 710 server.

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 36.4 GB up to 300 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, one additional LPAR (using dedicated devices) could be provided with up to seven disks for rootvg by using one Ultra3 SCSI adapter (FC 1974) or one Ultra3 SCSI RAID adapter (FC 1975).

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

http://www.ibm.com/servers/storage/disk/expplus/index.html

IBM TotalStorage Storage servers

The IBM TotalStorage DS4000 Storage server family consists of five models: Model DS4100, DS4300, DS4400, and DS4500. The Model DS4100 is the smallest model, which scales up to 28 TB; and Model DS4500 is the largest, which scales up to 32 TB of disk storage at the time this publication was written. Model DS4300 provides up to 16 bootable partitions that are attached with the Gigabit Fibre Channel adapter (FC 1977). Model DS4400 provides up to 64 bootable partitions. In most cases, both the TotalStorage DS family and the OpenPower 710 server are connected to a storage area network (SAN). If only space for the rootvg is needed, the Model DS4100 is a good solution.

For support of additional features and for further information about the IBM TotalStorage DS4000 Storage Server family, refer to the following Web site:

http://www.ibm.com/servers/storage/disk/ds4000/index.html

IBM TotalStorage Enterprise Storage Server®

The IBM TotalStorage Enterprise Storage Server (ESS) Models DS6000 and DS8000 are the high-end premier storage solution for use in storage area networks. The TotalStorage DS6000 provides Enterprise class capabilities in a space-efficient modular package. It scales to 67.2 TB of physical storage capacity by adding storage expansion enclosures. The Model DS8000 series is the flagship of the TotalStorage DS family. The DS8000 scales to 192 TB; however, the system architecture is designed to scale to over one petabyte. The Model DS6000 and DS8000 systems can also be used to provide disk space for booting LPARs or partitions using Micro-Partitioning technology. ESS is usually connected to a SAN, to which the OpenPower 710 server is also connected by using Gigabit Fibre Channel adapters (FC 1977).

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

http://www.ibm.com/servers/storage/disk/enterprise/ds_family.html

2.8 The POWER Hypervisor and Virtual I/O Server technologies

The optional POWER Hypervisor and Virtual I/O Server feature (FC 1965) includes the POWER Hypervisor enablement and Virtual I/O Server (VIOS) software for OpenPower systems.

The POWER Hypervisor and Virtual I/O Server feature (FC 1965) provides:

- The POWER Hypervisor:
  - Firmware enablement for LPAR
  - Firmware enablement for dynamic LPAR
  - Firmware enablement for partitions using Micro-Partitioning technology
  - Support for virtual SCSI (VIOS software)
  - Virtual Ethernet
  - Virtual TTY
- The Virtual I/O Server software:
  - Ethernet adapter sharing
  - Virtual SCSI

2.8.1 POWER Hypervisor

When FC 1965 is ordered for the OpenPower 710 server on an initial order, the POWER Hypervisor is shipped enabled. On MES orders the firmware enablement code must be downloaded from the following Web page:

`http://www.ibm.com/servers/eserver/openpower/cod`

The firmware enablement code is entered into the HMC to enable the POWER Hypervisor technology. The POWER Hypervisor technology includes static LPAR, dynamic LPAR[2], Micro-Partitioning, and virtual adapters.

Note: On the OpenPower 710 server, LPAR, dynamic LPAR, and other advanced virtualization features, such as Micro-Partitioning and virtual Ethernet, are available only when a system is configured with the POWER Hypervisor and Virtual I/O Server feature, FC 1965.

Logical partitioning and dynamic logical partitioning

The POWER Hypervisor technology offers the capability to divide the OpenPower 710 server into logical partitions, where each LPAR runs an operating environment on dedicated attached devices, such as processors, memory, and I/O components. The HMC is required when partitioning the OpenPower 710 server. The HMC is required to create and modify the partitions.

Static LPAR is supported by Red Hat Enterprise Linux AS 3 for POWER. Static and dynamic LPAR are supported by SUSE LINUX Enterprise Server 9 for POWER, or later, and Red Hat Enterprise Linux AS 4 for POWER, or later. Static LPARs require a partition reboot for modifications to be applied. Dynamic LPAR increases flexibility, allowing selected system resources, such as processors and I/O components, to be added and deleted from dedicated partitions without reboot. 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.

[2] Dynamic LPAR support varies by Linux distribution.

Note: Dynamic logical partitioning is supported by SUSE 9 and RHEL AS 4 or later. Changing memory attributes dynamically is not supported at the time of writing.

Micro-Partitioning technology

The OpenPower 710 POWER Hypervisor, when enabled, provides Micro-Partitioning virtualization technology. With Micro-Partitioning, a partition may be defined with a processor capacity as small as 10 processor units. This represents 1/10 of a physical processor. Each processor can be shared by up to 10 shared processor partitions and can increment fractionally as little as 1/100th of the processor. The shared processor partitions are created and managed by the HMC. Dedicated and micro-partitioned processors can co-exist on the same POWER5 server as long as resources allow. Table 2-7 lists processor partitioning information related to the OpenPower 710 server.

Table 2-7 Processor partitioning overview of the OpenPower 710 server

Partitioning implementation (based on 2-way configuration)	OpenPower 710
Processors	2
Dedicated processor partitions	2
Shared processor partitions	20

It is important to point out that the maximums stated are supported by the hardware, but the practical limits based on production workload demands may be lower. Table 2-8 on page 30 lists operating systems supported with Micro-Partitioning.

Virtual I/O adapters

The OpenPower 710 POWER Hypervisor, when enabled, also supports virtual I/O adapter virtualization technology. Three types of virtual I/O adapters are supported by the POWER Hypervisor.

Virtual SCSI

The POWER5 server uses SCSI as the mechanism for virtual storage devices. This is accomplished using two paired adapters: A virtual SCSI server adapter and a virtual SCSI client adapter. Virtual SCSI requires FC 1965, the appropriate level of Linux, and an HMC to define the virtual SCSI devices, as follows.

Virtual Ethernet

The POWER Hypervisor provides a virtual Ethernet switch function that allows partitions on the *same server* a means for fast and secure communication. 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 requires FC 1965, the appropriate level of Linux, and an HMC to define the virtual Ethernet devices. Virtual Ethernet does not require the Virtual I/O Server.

Virtual Ethernet features include:

- A partition supports 256 virtual Ethernet connections, where a single virtual Ethernet resource can be connected to another Virtual Ethernet, a real network adapter, or both in a partition. Each Virtual Ethernet adapter can also be configured as a trunk adapter.
- Each partition operating system sees the virtual local area network (VLAN) switch as an Ethernet adapter, without the physical link properties and asynchronous data transmit operations. Layer-2 bridging to a physical Ethernet adapter is also included in the virtual

[3] Maximum transmission unit

Ethernet features. The virtual Ethernet network is extendable outside the server to a physical Ethernet network.

Note: Virtual Ethernet is based on the IEEE 802.1Q VLAN standard. No physical I/O adapter is required when creating a VLAN connection between partitions, and no access to an outside network is required.

Virtual (TTY) console

Each partition needs to have access to a system console. Tasks such as operating system install, network setup, and some problem analysis activities require a dedicated system console. The POWER Hypervisor, when enabled, provides virtual console using a virtual TTY or serial adapter and a set of Hypervisor calls to operate on them. Virtual TTY does not require the Virtual I/O Server to be running on the system.

Depending on the system configuration, the operating system console can be provided by the Hardware Management Console, virtual TTY, or from a terminal emulator connected to physical serial ports on the system's service processor.

The virtual TTY console support should not be confused with the virtual serial available with the service processor, as described in "Service processor communication ports" on page 23. The virtual TTY support discussed here is configured using the HMC.

Note: When enabled, the POWER Hypervisor is active when the server is running in partition and non-partition mode, and also when not connected to the HMC. The Hypervisor memory requirements should be considered when planning the system memory specifications. Use the LPAR Validation Tool for calculating the Hypervisor memory requirements.

Virtual I/O Server

The OpenPower 710 POWER Hypervisor and Virtual I/O Server technologies also include the VIOS software. The VIOS software is not required for static or dynamic LPAR, Micro-Partitioning, virtual Ethernet, or virtual TTY. The Virtual I/O Server is a special purpose partition that provides virtual I/O resources to client partitions. The Virtual I/O Server will own the real resources that will be shared with the other LPARs. 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 the number of physical adapters. The Virtual I/O Server eliminates the requirement that each partition own a dedicated network adapter, disk adapter, and disk drive. The Virtual I/O Server software is only supported in Virtual I/O Server partitions.

Figure 2-6 shows a view of Linux virtual partitions including the Virtual I/O Server. The figure also includes virtual SCSI and Ethernet connections and mixed operating system partitions.

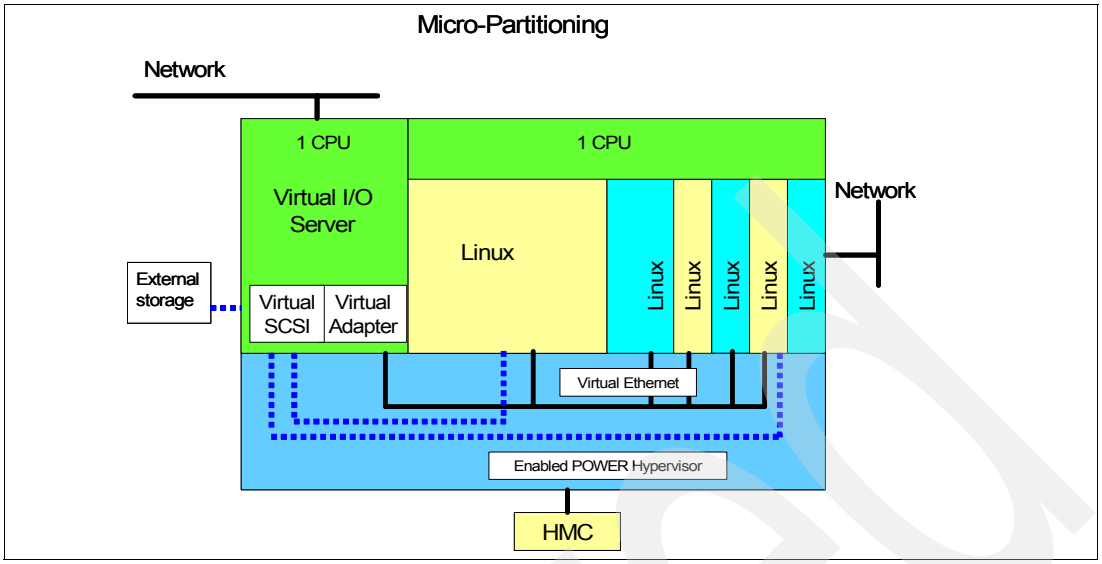


Figure 2-6 Virtual partition organization view

Since the Virtual I/O Server is an appliance, redundancy for physical devices attached to the Virtual I/O Server can be provided by using capabilities such as LVM mirroring, Multipath I/O, and IEEE 802.3ad Link Aggregation.

Note: To increase the performances of I/O-intensive applications, dedicated physical adapters should be preferred in dedicated partitions.

It is recommended that you install the Virtual I/O Server in a partition with dedicated resources to help ensure consistent performance.

The Virtual I/O Server supports logical mirroring and RAID configurations. Logical volumes created on Redundant Array of Independent Disks (RAID) or Just a Bunch Of Disks (JBOD) configurations are bootable, and the number of logical volumes is limited to the amount of storage available and architectural limits of the logical volume manager (LVM).

Two major functions are provided with the Virtual I/O Server: A shared Ethernet adapter and Virtual SCSI Server.

Shared Ethernet adapter

A 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 be assigned to the Virtual I/O Server partition.

Virtual SCSI

Access to real storage devices is implemented through the Virtual SCSI services, a part of the Virtual I/O Server partition. This is accomplished using a pair of virtual adapters, a virtual SCSI server adapter, and a virtual SCSI client adapter. The virtual SCSI server (target) adapter is responsible for executing any SCSI commands it receives. It is owned by the Virtual I/O Server partition. The virtual SCSI client adapter allows the client partition to access standard SCSI devices and LUNs that are assigned to the client partition.

All current storage device types, such as SAN, SCSI, and RAID, are supported. iSCSI and SSA are not supported.

More information on specific storage devices supported can be found at:

http://techsupport.services.ibm.com/server/virtualization/vios/documentation/datasheet.html

Note: The shared Ethernet adapter and virtual SCSI function is provided in the Virtual I/O Server that is included in the POWER Hypervisor and Virtual I/O Server feature services.

Refer to Table 2-8 on page 30 for details on requirements for configuring virtual partitions.

Note: Partition Load Manager (PLM) is not available because it does not interoperate with Linux.

2.8.2 Hardware and software guidelines for dynamic LPAR and virtualization

This section covers the main considerations regarding dynamic LPAR and virtualization.

Operating system support for dynamic LPAR and virtualization

Table 2-8 shows Linux support for dynamic LPAR and virtualization.

Table 2-8 Operating system supported function

Function	Linux SLES 9	Linux RHEL AS 3	Linux RHEL AS 4
Dynamic LPAR			
Processor	Y	N	Y
Memory	N	N	N
I/O	Y	N	Y
Virtualization			
Micro-partitions (1/10th of processor)	Y	Y	Y
Virtual Storage	Y	Y	Y
Virtual Ethernet	Y	Y	Y

Dynamic LPAR minimum requirements

The minimum resources that are needed per LPAR (not per system) are the following:

- At least one processor per partition for a dedicated processor partition or at least 1/10th of a processor when using Micro-Partitioning technology.
- At least 128 MB of physical memory per additional partition.
- At least one disk (either physical or virtual) to store the operating system.
- At least one disk adapter (either physical or virtual) or integrated adapter to access the disk.
- At least one Ethernet adapter (either physical or virtual) per partition to provide a network connection to the HMC, as well as general network access.

Note: It is recommended to use separate adapters for the management and the public LAN to protect access of your system's management functions.

- A partition must have an installation method, such as NIM, and a means of running diagnostics, such as network diagnostics.

Processor

Each LPAR requires at least one physical processor if virtualization is not used. Based on this, the maximum number of dynamic LPARs per OpenPower 710 server is two without virtualization. With the use of the IBM Advanced POWER Virtualization, the number of partitions per processor is 10.

Memory

In a partitioned environment, some of the physical memory areas are reserved by several system functions to enable partitioning in the partitioning-capable OpenPower server. You can assign unused physical memory to a partition. You do not have to specify the precise address of the assigned physical memory in the partition profile, because the system selects the resources automatically.

The Hypervisor requires memory to support the logical partitions on the server. The amount of memory required by the Hypervisor varies according to several factors. Factors influencing the Hypervisor memory requirements include the following:

- Number of logical partitions
- Partition environments of the logical partitions
- Number of physical and virtual I/O devices used by the logical partitions
- Maximum memory values given to the logical partitions

Generally, you can estimate the amount of memory required by server firmware to be approximately 8 percent of the system installed memory. The actual amount required will generally be less than 8 percent. However, there are some server models that require an absolute minimum amount of memory for server firmware, regardless of the previously mentioned considerations.

The minimum amount of physical memory for each partition is 128 MB, but in most cases the actual requirements and recommendations are between 256 MB and 512 MB for Red Hat and Novell SUSE. After that, you can assign further physical memory to partitions in increments of 16 MB. This is supported for partitions running Red Hat Enterprise Linux AS 3 (no dynamic LPAR), Red Hat Enterprise Linux AS 4, and SUSE LINUX Enterprise Server 9. There are implications on how big a partition can grow based on the amount of memory allocated initially. For partitions that are initially sized less than 256 MB, the maximum size is 16 times the initial size. For partitions initially sized 256 MB or larger, the maximum size is 64 times the initial size.

Note: For a more detailed impression of the amount of memory required by the server firmware, use the LPAR Validation Tool (LVT). Please refer to “LPAR validation tool” on page 34.

I/O

The I/O devices are assigned on a slot level to the LPARs, meaning an adapter (either physical or virtual) installed in a specific slot can only be assigned to one LPAR.

If an adapter has multiple devices, such as the 4-port Ethernet adapter or the Dual Ultra3 SCSI adapter, all devices are automatically assigned to one LPAR and cannot be shared.

Devices connected to an internal controller must be treated as a group. A group can only be assigned together to one LPAR and cannot be shared.

Therefore, the following integrated devices can be independent of each other assigned to LPARs:

- Integrated dual-channel Ultra320 SCSI controller

  All SCSI resources in the disk bays, including external disks that are connected to the external SCSI port, must be assigned together to the same LPAR. There is no requirement to assign them to a particular LPAR; in fact, they can remain unassigned if the LPAR minimum requirements are obtained using devices attached to a SCSI adapter installed in the system.

- Media devices connected to the IDE controller

  The IDE devices can be assigned to one LPAR, independent from other resources, for example SCSI attached devices.

- Integrated dual 10/100/1000 Mbps Ethernet controller
- Integrated dual USB controller

Virtual I/O devices are also assigned to dynamic LPARs on a slot level. Each partition is capable of handling up to 256 virtual I/O slots. Therefore each partition can have up to:

- 256 virtual Ethernet adapters with each virtual Ethernet capable of being associated with up to 21 VLANs.
- 256 virtual SCSI adapters

Note: For more detailed planning of the virtual I/O slots and their requirements, use the LPAR validation tool.

Every LPAR requires disks (either physical or virtual) for the operating system. Systems with internal disks are connected to the internal SCSI port. As described previously, all integrated SCSI devices, including all internal disks, and all external devices connected to the second SCSI port of the integrated SCSI adapter, can only be assigned to the same LPAR.

Therefore, for additional LPARs without using virtualization, external disk space is necessary, which can be accomplished by using external disk subsystems. The external disk space must be attached with a separate adapter for each LPAR by using SCSI or Fibre Channel adapters, depending on the subsystem.

For additional LPARs using virtualization, the required disk drives for each partition are provided by the Virtual I/O Server partition(s). Physical disks owned by the Virtual I/O Server partition can either be exported and assigned to a client partition whole, or can be partitioned into several logical volumes. The logical volumes can then be assigned to different partitions.

For the OpenPower 710 server, additional disk space can be provided by using an external storage subsystem such as the IBM 2104 Expandable Storage Plus. For more detailed information of the available IBM disk subsystems, refer to “External disk subsystem” on page 25.

Every LPAR needs an Ethernet connection to the HMC. As described previously, all internal Ethernet devices could only be assigned to the same LPAR.

Therefore, for additional LPARs without using virtualization, an additional Ethernet adapter is necessary. As stated earlier, it is highly recommended to use separate Ethernet adapters for connection to the management LAN and public LAN. Due to the limited number of available PCI slots, it is recommended to use an additional 2-Port 10/100/1000Mbps Ethernet Adapter (FC 1983) to provide physically sperated Ethernet ports to the second LPAR, without virtualization.

Additional partitions using virtualization can implement the required Ethernet adapters as virtual Ethernet adapters. Virtual Ethernet adapters can be used for all kinds of inter-partition communication. To connect the virtual Ethernet LANs to an external network, one or more Shared Ethernet Adapters (SEAs) can be used in the Virtual I/O Server partition.

Note: When dedicated physical LPARs are a requirement, be certain to look at the total cost of the configuration. Consider other OpenPower systems offering additional dedicated devices under the covers against the cost of an external storage enclosure and the OpenPower 710 server. Or use of the POWER Hypervisor and Virtual I/O Server feature.

2.8.3 Hardware Management Console

The HMC is a dedicated workstation that provides a graphical user interface for configuring and operating IBM @server OpenPower 710 servers functioning in either non-partitioned, LPAR, or clustered environments. It is configured with a set of hardware management applications for configuring and partitioning the server. One HMC is capable of controlling multiple POWER5 processor-based systems.

Note: If the POWER Hypervisor is *not* activated on an OpenPower 710 server, a connected HMC can only provide hardware functions such as manage service utilities, perform code updates, and provide power cycle control.

Note: At the time of writing, one HMC supports up to 32 POWER5 processor-based systems and up to 256 LPARs using the HMC machine code Version 4.4.

For updates of the machine code and HMC functions and hardware prerequisites refer to the following Web page:

`https://techsupport.services.ibm.com/server/hmc/power5`

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 HMC 7310 Model C04 is a desktop model with only one integrated 10/100/1000 Mbps Ethernet port, but two additional PCI slots.
- The 7310 Model CR3 is a 1U, 19-inch rack-mountable drawer that has two native 10/100/1000 Mbps Ethernet ports and two additional PCI slots.

Note: It is possible to use virtual Ethernet and shared Ethernet adapter for an HMC connection to partitions.

The HMC connects to the OpenPower 710 server using a standard 10/100 Ethernet connection and uses HMC Port 1. A second HMC can be attached to HMC Port 2 of the OpenPower 710 server for redundancy. For more details of the possible HMC network connections refer to **Managing your server → Hardware Management Console → HMC Concepts → Types of HMC network connections** at the IBM @server Hardware Information Center at:

`http://publib.boulder.ibm.com/infocenter/eserver/v1r2s/en_US/index.htm`

Note: Both HMC Ethernet ports are only visible to the service processor.

When an HMC is connected to the OpenPower 710 server, the OpenPower 710 integrated service processor communications ports are disabled. An async adapter is required to provide serial connections if required.

Older HMC models, such as the 7315-CR2 or 7315-C03 and others, can be upgraded to support POWER5 processor-based systems.

To upgrade an existing POWER4 HMC:

- Order FC 0961 for your existing HMC. Contact your IBM Sales Representative for help.
- Call an IBM Service Center and order APAR MB00691.
- Order the CD online by selecting **Version 4.4 machine code updates** → **Order CD** → **Go** at the Hardware Management Console Support Web page at:

  `https://techsupport.services.ibm.com/server/hmc/power5`

Note: You must have an IBM ID to use this freely available service. Registration information and an online registration form can be found at the above Web page.

Table 2-9 lists the major differences between consoles.

Table 2-9 Contrast between 7310 and 7315 HMCs

Function	7310	7315
Supported platforms	POWER5	POWER4/POWER4+
HMC to service processor connection type	Ethernet (RJ45)	Serial (RS232)
Installation	Client setup	IBM System Services Representative
Licensed machine code	0961	0960

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

2.8.4 LPAR validation tool

When configuring dynamic or virtual partitions on POWER5 systems, the LPAR Validation Tool (LVT) can be used to verify system resource requirements. With the LVT, you can customize the partition design by selecting PCI slots for given adapters, specific drives to selected bays, and much more. The LVT provides a useful report that can complement the organization and validation of features required for configuration of a complex partition solution. The LVT supports IBM @server p5 and IBM @server i5 servers, iSeries™, and OpenPower systems. A proficient knowledge of LPAR design requirements, limitations, and best practice, facilities the use of this tool.

The LVT tool provides the following functions:

- Support for partitions running Linux
- Validation of dynamic LPAR design
- Validation of virtual partition design, including Virtual I/O Server and virtual clients
- Calculates unallocated memory and shared processor pool
- Calculates Hypervisor memory requirements

- Calculates number of operating system license(s) needed to support partition design
- Validates number of virtual slots required for partitions

Important: We recommend the use of the LVT to calculate Hypervisor requirements to determine memory resources needed for all partitioned and non-partitioned servers.

Figure 2-7 on page 35 shows the calculated Hypervisor memory requirements based on sample partition requirements.

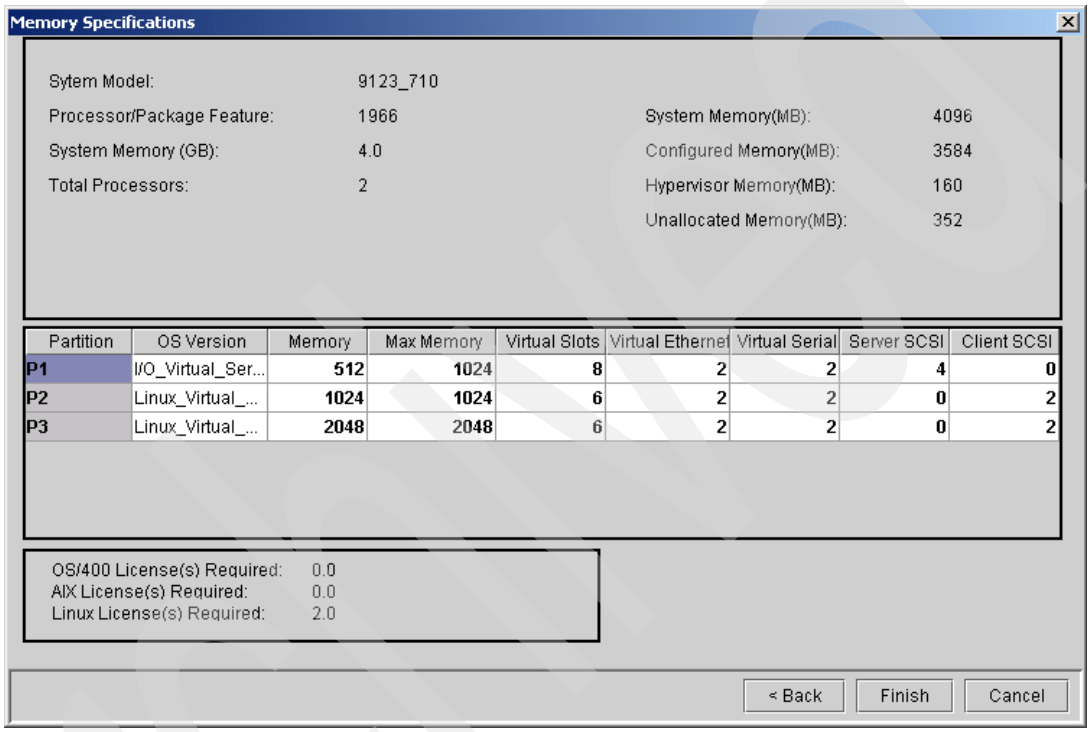


Figure 2-7 LVT screen showing Hypervisor requirements

The LVT is a standalone Java™ application that runs on a Microsoft® Windows® 95 or later workstation with 128 MB minimum of free memory.

For download and installation information, including a user's guide, visit:

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

2.8.5 Client-specific placement and eConfig

The LVT also provides the output report that is used for the Customer Specified Placement (CSP) offering. The LVT output is uploaded on the CSP site for submission to manufacturing. The CSP offering enables the placement of adapters and disks for an exact built-to-order system based on the client's specifications. Manufacturing uses the LVT output to custom build the server. The server will be shipped configured with the features placed as indicated in the LVT.lvt output file.

The server configuration must include the CSP FC 8453. This CSP feature code is selected on the Code tab in the IBM Configurator for e-business (eConfig) wizard. See Figure 2-8 on page 36 for a screen shot.

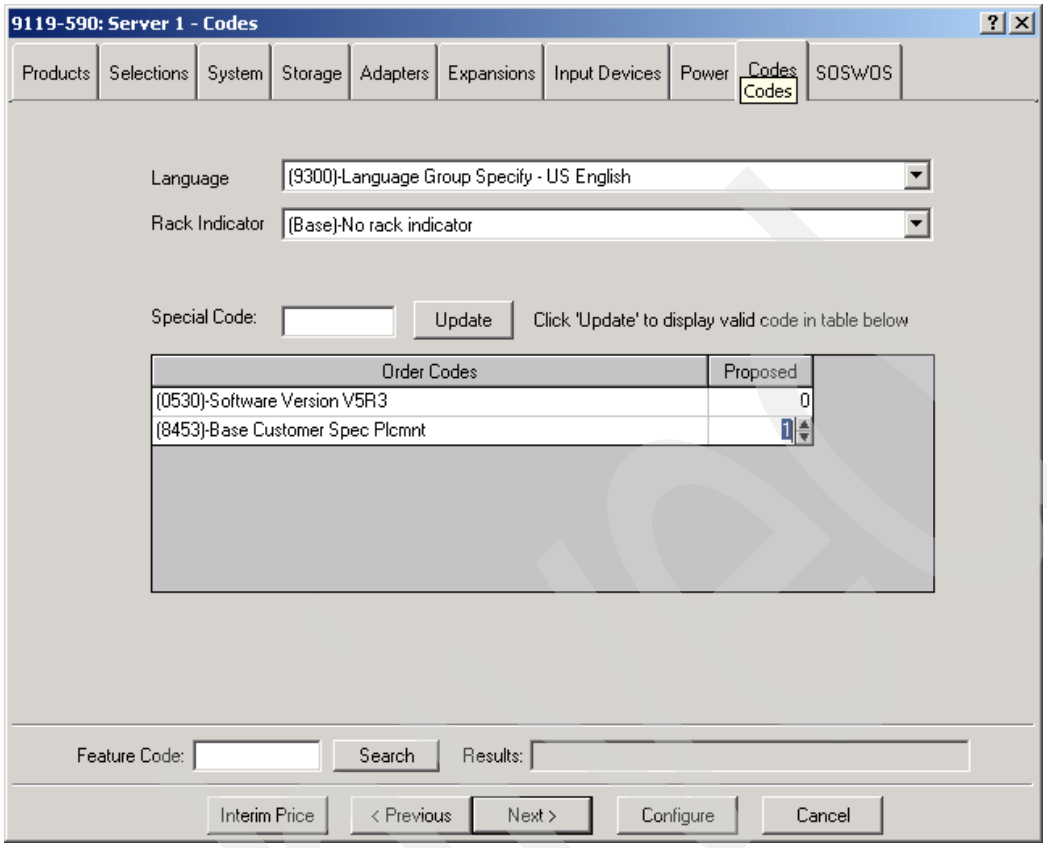


Figure 2-8 eConfig tab for CSP selection

Note: The order submitted must include the CSP feature code 8453 on the configuration, and the form submitted to the CSP site must include the.lvt output file and the order number of the system.

CSP is available on a limited number of POWER5 systems. See the following URL for details:

`http://www.ibm.com/servers/eserver/power/csp/index.html`

2.8.6 Linux

For the OpenPower 710 server, Linux distributions were available through Novell SUSE and Red Hat at the time this publication was written. The OpenPower 710 server requires the following version of Linux distributions:

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

Note: Dynamic LPAR is not supported by Red Hat Enterprise Linux AS for POWER Version 3.

Information on features and external devices supported by Linux on the OpenPower 710 server can be found at:

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

Information about SUSE LINUX Enterprise Server 9 can be found at:

http://www.novell.com/products/linuxenterpriseserver/

For information about Red Hat Enterprise Linux AS for pSeries from Red Hat. Cee:

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

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

https://www14.software.ibm.com/webapp/iwm/web/preLogin.do?source=linuxline

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

http://www.ibm.com/servers/eserver/linux/power/whitepapers/linux_overview.html

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

Cluster subscription offerings

Cluster software subscription offerings are available for OpenPower systems running SUSE LINUX Enterprise Server 9 (SLES 9) for POWER and Red Hat Enterprise Linux AS Version 3 or later (RHEL AS 3, or later).

Cluster software SLES 9 options for clustered servers

The SLES 9 cluster subscription offerings provide lower priced SLES 9 license options for clients who run clusters of one to eight servers, each of which is a 1- to 2-way capable server, when compared to the subscription price for each server individually. These offerings are available only for p5-510 and p5-510 Express, p5-520 and p5-520 Express, and OpenPower 710 server.

SLES 9 clusters require a *head node* (which must be one of the supported servers for the offering) for administration.

Cluster software RHEL AS 3 options for clustered servers

The RHEL AS Version 3 and RHEL AS Version 4 cluster subscription offerings provide lower priced RHEL AS Version 3 and RHEL AS Version 4 license options for clients who run clusters of 1–8 servers, each of which is a 1-way to 2-way capable server, when compared to the subscription price for each server individually. These offerings are available only for p5-510 and p5-510 Express, p5-520 and p5-520 Express, and OpenPower 710 server.

Note: Red Hat defines clustered servers as having the same configurations. This requirement should be considered when designing system configuration and workload requirements.

Note: The system is capable of running other operating systems, including AIX 5L, if enabled. Upon request, IBM may consider RPQs for installed systems to enable that capability.

2.9 Servicing information

The OpenPower 710 server is a client setup system and is shipped with materials to assist in the general installation of the server. The system cover also includes a quick reference Service Information label that provides graphics that can aid in identifying features and location information. The sections below provide some additional service-related information.

2.9.1 Touch point colors

Blue (IBM blue) or terra-cotta (orange) on a component indicates a touch point (for electronic parts) where you can grip the hardware to remove it from or install it into the system, open or close a latch, and so on. The touch point colors are defined as follows:

Blue Requires shut down of the system before the task can be performed; for example, removing the PCI riser book to install PCI adapters in the OpenPower 710 server.

Terra-cotta The system can remain powered on while these tasks are being performed. Keep in mind that some tasks may require some steps to be executed before performing the task. One example would be deconfiguring a physical volume in the operating system before removing the disk from the OpenPower 710 server.

Blue and Terra-cotta When both colors are present, terra-cotta takes precedence, and the rules for a terra-cotta-only touch point apply.

Important: It is important to adhere to the touch point colors on the system. Not doing so can compromise your safety and damage the system.

Important: When removing the system planar, the cable retention screws must be removed. Caution should be exercised when tightening these screws back to the chassis. Excessively tightening the screws can cause permanent hardware damage.

2.9.2 Securing system into racks

The rack-mount drawer rail kit is a unique kit designed for use with the OpenPower 710 server. No tools are required to install the OpenPower 710 server or drawer rails into the rack.

The kit has a modular design that can be adapted to accommodate various rack depths specifications. The drawer rails are equipped with thumb-releases on the sides, towards the front of the system, that allow for the system to easily slide out from its rack position for servicing.

Note: Standard safety precautions should always be exercised when installing or removing devices from racks.

To place the OpenPower 710 server in the service position:

- If necessary, open the front rack door.
- Remove the two thumbscrews (A) that secure the system unit to the rack, as shown in Figure 2-9.
- Release the rack latches (B) on both the left and right sides, as shown in Figure 2-9.

- Review the following notes, and then slowly pull the system unit out from the rack until the rails are fully extended and locked.
  - If the procedure you are performing requires you to unplug cables from the back of the system unit, do so before you pull the unit out from the rack.
  - Ensure that the cables at the rear of the system unit do not catch or bind as you pull the unit out from the rack.
  - When the rails are fully extended, the rail safety latches lock into place. This action prevents the system unit from being pulled out too far.

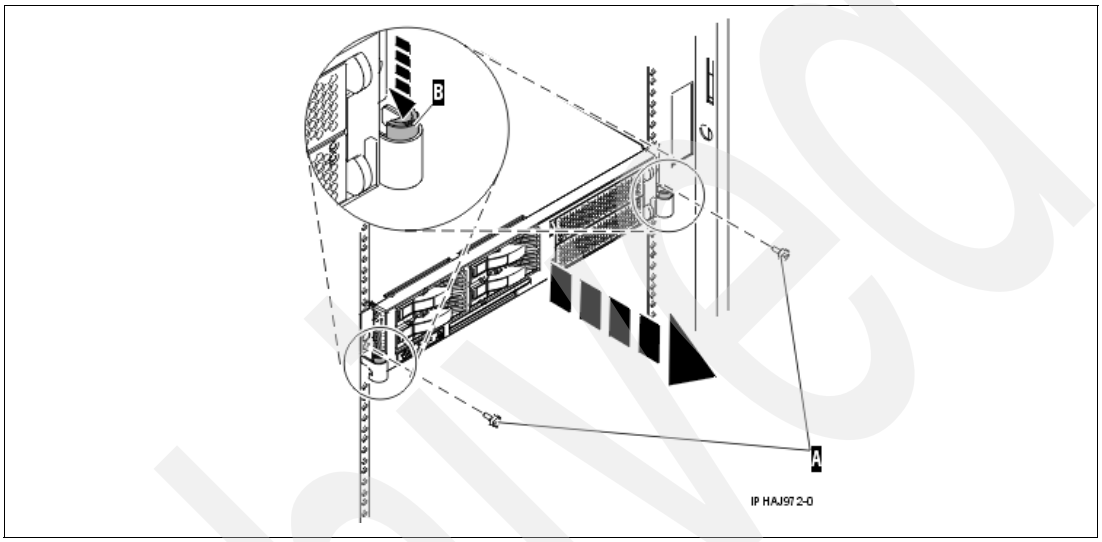


Figure 2-9 OpenPower 710 server service position

For more information see the documentation site:

http://publib16.boulder.ibm.com/infocenter/eserver/v1r2s/en_US/index.htm

2.9.3 Operator control panel

The service processor provides an interface to the control panel, which is used to display system status and diagnostic information. The OpenPower 710 server has a control panel that is a re-packaging of existing control panels, so that it fits into a smaller space. In its normal position, the control panel is seated inside the chassis, on the left of the DVD optical device if you stand in front of the server. The LCD display is invisible from the front. To read the LCD display, the client or engineer pulls the op-panel out toward the front.

Note: Systems managed by the Hardware Management Console (HMC) should use the HMC to perform control panel functions.

Accessing and storage of the control panel

To access all of the control panel's features, do the following (refer to Figure 2-10):

- Press inward on the spring-loaded tab (A) located on the right side of the control panel (B) so that it pops out slightly.
- Pull the control panel out toward the front of the system until it can be pivoted downward on its hinge.

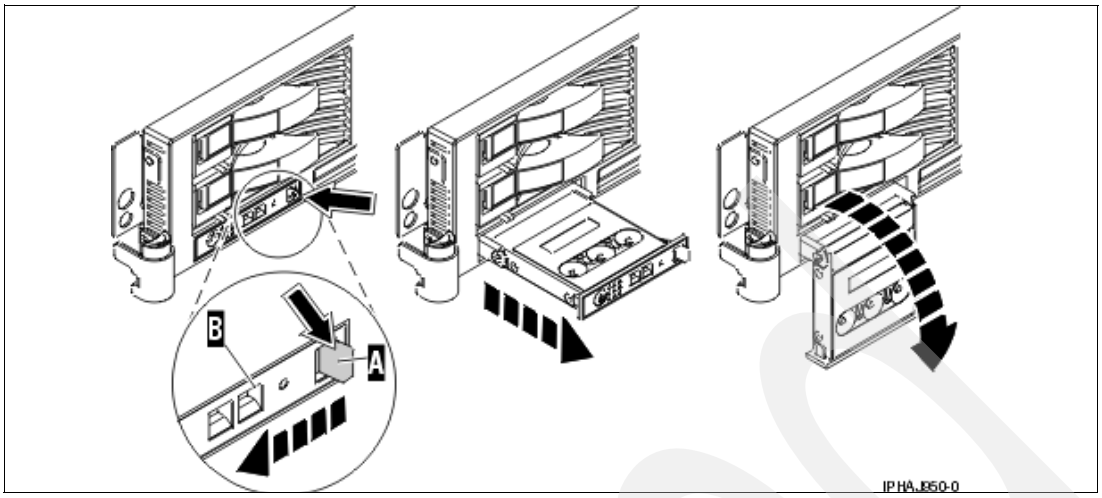


Figure 2-10 Accessing the control panel

- To move the control panel back into the device enclosure, lift the control panel up to align it with the opening and push it into place until you feel the tab lock. Refer to Figure 2-11 for details.

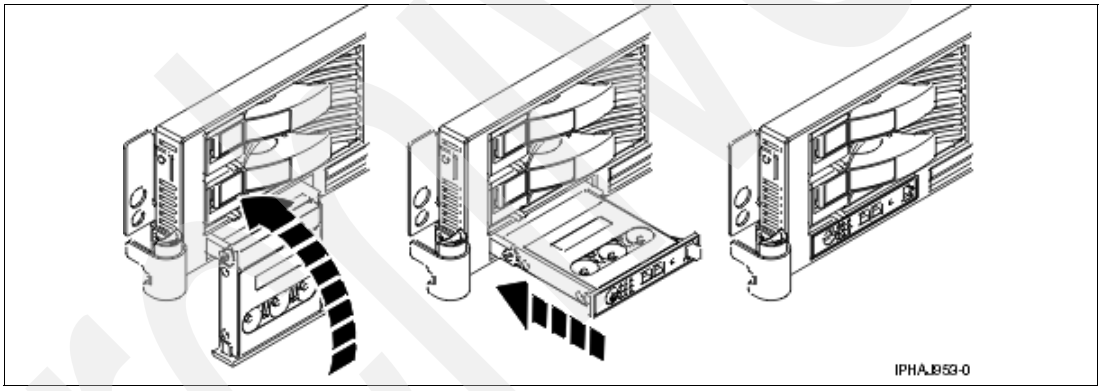


Figure 2-11 Storage of the control panel

Primary control panel functions

The primary control panel functions are defined as functions 01 to 20, including options to view and manipulate IPL modes, system operating modes, IPL speed, and IPL type.

Here is a list of all available primary functions:

- Function 01 – Display selected IPL type, system operating mode, and IPL speed.
- Function 02 – Select IPL type, IPL speed override, and system operating mode.
- Function 03 – Start IPL.
- Function 04 – Lamp test.
- Function 05 - Reserved.
- Function 06 - Reserved.
- Function 07 – SPCN functions.
- Function 08 – Fast power off.
- Functions 09 to 10 – Reserved.

- Functions 11 to 19 – System reference code.
- Function 20 – System type, model, feature code, and IPL type.

All functions mentioned above are accessible either using the ASMI or the HMC or the control panel.

For detailed information about each control panel function and the available values, go to the IBM *e*server Hardware Information Center → **Service Provider Information** → **Reference Information** → **Service functions** → **Control panel functions** at:

http://publib.boulder.ibm.com/infocenter/eserver/v1r2s/en_US/index.htm

Changing system operation mode

In the unlikely event that a OpenPower 710 server is suspended and you cannot perform a power down of the CEC, use function 08 for a fast power off, if the system is not connected to a HMC.

Note: Function 08 is enabled only when the system operating mode is in Manual and the system power is on. By default the operating mode is set to Normal, which means you cannot use the white power button to power-off the system.

To change the operating mode of the system to Manual (M) do the following steps:

1. Use the Increment or Decrement buttons to scroll to function 02.
2. Press Enter to start function 02.
   - The current OS IPL type is displayed with a pointer.
   - The current system operating mode is displayed.
   - The current firmware mode is displayed.

   Note: The following current firmware modes are possible:
   - Normal (N)
   - Manual (M)

3. Use the Increment or Decrement buttons to scroll through the OS IPL types until the firmware mode is displayed with a pointer:

```
0 2 _ _ A _ _ N < _ _ _ _ _ _ _
_ _ _ _ _ _ _ _ _ _ _ _ _ P _ _
```

4. Press Enter to select the system operating mode:

```
0 2 _ _ A _ _ M < _ _ _ _ _ _ _
_ _ _ _ _ _ _ _ _ _ _ _ _ P _ _
```

5. Use the Increment or Decrement buttons to scroll through the firmware IPL modes (optional).
6. Press Enter to select the firmware IPL mode and exit function 02.

For more functions and information on the control panel refer to **Service and support** → **Reference information** → **Service function** → **Control panel function** at the IBM *e*server Hardware Information Center at:

http://publib.boulder.ibm.com/infocenter/eserver/v1r2s/en_US/index.htm

2.9.4 System firmware

Server firmware is the part of the Licensed Internal Code that enables hardware such as the service processor. Depending on your service environment, you can download, install, and manage your server firmware fixes using different interfaces and methods, including the HMC or by using functions specific to your operating system.

Note: Installing the server firmware fixes through the operating system normally is a nonconcurrent process.

Levels of firmware - Temporary and permanent side of the service processor

The service processor maintains two copies of the server firmware.

- One copy is considered the permanent or backup copy and is stored on the permanent side, sometimes referred to as the "p" side.
- The other copy is considered the installed or temporary copy and is stored on the temporary side, sometimes referred to as the "t" side. It is recommended that you start and run the server from the temporary side.
- The copy actually booted from is called the activated level, or sometimes referred to as “b”.

Note: The default value the system will boot is temporary.

The following example is the output of the **lsmcode** command for Linux, showing the firmware levels as they are displayed in the outputs:

```
system:SF220_006 (t) SF220_005 (p) SF220_006 (b)
```

When you install a server firmware fix, it is installed on the temporary side.

Note: The following points are of special interest:

- The server firmware fix is installed on the temporary side only after the existing contents of the temporary side are permanently installed on the permanent side (the service processor performs this process automatically when you install a server firmware fix).
- If you want to preserve the contents of the permanent side, you need to remove the current level of firmware (copy the contents of the permanent side to the temporary side) before you install the fix.
- However, if you get your fixes using Advanced features on the HMC interface and you indicate that you do not want the service processor to automatically accept the firmware level, the contents of the temporary side are not automatically installed on the permanent side. In this situation, you do not need to remove the current level of firmware to preserve the contents of the permanent side before you install the fix.

You might want to use the new level of firmware for a period of time to verify that it works correctly. When you are sure that the new level of firmware works correctly, you can permanently install the server firmware fix. When you permanently install a server firmware fix, you copy the temporary firmware level from the temporary side to the permanent side.

Conversely, if you decide that you do not want to keep the new level of server firmware, you can remove the current level of firmware. When you remove the current level of firmware, you copy the firmware level that is currently installed on the permanent side from the permanent side to the temporary side.

For a detailed description of firmware levels refer to the IBM @server Hardware Information Center → **Service and support** → **Customer service and support** → **Getting fixes** → **Firmware (Licensed Internal Code) fixes** → **Concepts** → **Temporary and permanent side of the service processor** at:

http://publib.boulder.ibm.com/infocenter/eserver/v1r2s/en_US/index.htm

Get server firmware fixes using an HMC

You use an HMC to manage your server and you have configured several partitions on the server. Periodically, you need to download and install fixes for your server and power subsystem firmware.

You want to use the HMC to perform this task. How you get the fix depends on whether or not the HMC or server is connected to the Internet.

- If the HMC or server is connected to the Internet:

  There are several repository locations from which you can download the fixes using the HMC. For example, you can download the fixes from your service provider's Web site or support system, from optical media that you order from your service provider, or from an FTP server on which you previously placed the fixes.

- If neither the HMC nor your server is connected to the Internet (server firmware only):

  You will need to download your new system firmware level to a CD-ROM media or FTP server.

For both of these options, you can use the interface on the HMC to install the firmware fix (from one of the repository locations or from the optical media). The Change Internal Code wizard on the HMC provides a step-by-step process for you to perform the required steps to install the fix.

1. Ensure that you have a connection to the service provider (if you have an Internet connection from the HMC or server).
2. Determine the available levels of server and power subsystem firmware.
3. Create optical media (if you do not have an Internet connection from the HMC or server).
4. Use the Change Internal Code wizard to update your server and power subsystem firmware.
5. Verify that the fix installed successfully.

For a detailed description of each task go to the IBM @server Hardware Information Center → **Service and support** → **Customer service and support** → **Getting fixes** → **Firmware (Licensed Internal Code) fixes** → **Scenarios: Firmware (Licensed Internal Code) fixes** → **Scenario: Get server firmware fixes using Task an HMC** at:

http://publib.boulder.ibm.com/infocenter/eserver/v1r2s/en_US/index.htm

Get server firmware fixes without an HMC

Periodically, you need to install fixes for your server firmware. If you do not use an HMC to manage your server, you must get your fixes through your operating system. In this situation, you can get server firmware fixes through the operating system.

You must complete the following tasks:

1. Determine the existing level of server firmware using the **lsmcode** command.
2. Determine the available levels of server firmware.
3. Get server firmware.

- If you are connected to the Internet.
- If you are not connected to the Internet.

4. Install the server firmware fix to the temporary side.
5. Verify that the server firmware fix installed successfully.
6. Install the server firmware fix permanently (optional).

Note: To view existing levels of server firmware using the **lsmcode** command, you need to have the following service tools installed on your server:

- Linux
  - Platform Enablement Library – librtas-*xxxxx*.rpm
  - Service Aids – ppc64-utils-*xxxxx*.rpm
  - Hardware Inventory – lsvpd-*xxxxx*.rpm

    Where *xxxxx* represents a specific version of the RPM file.

  If you do not have the service tools on your server, you can download them at the following Web page:

  http://techsupport.services.ibm.com/server/lopdiags

For a detailed description of each task go to the IBM @server Hardware Information Center → **Service and support** → **Customer service and support** → **Getting fixes** → **Firmware (Licensed Internal Code) fixes** → **Scenarios: Firmware (Licensed Internal Code) fixes** → **Scenario: Get server firmware fixes without an HMC** at:

http://publib.boulder.ibm.com/infocenter/eserver/v1r2s/en_US/index.htm

2.9.5 Service processor

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

The key components include a flexible service processor-base (FSP-B) and an extender chipset (FSP-E).

Figure 2-12 shows more details.

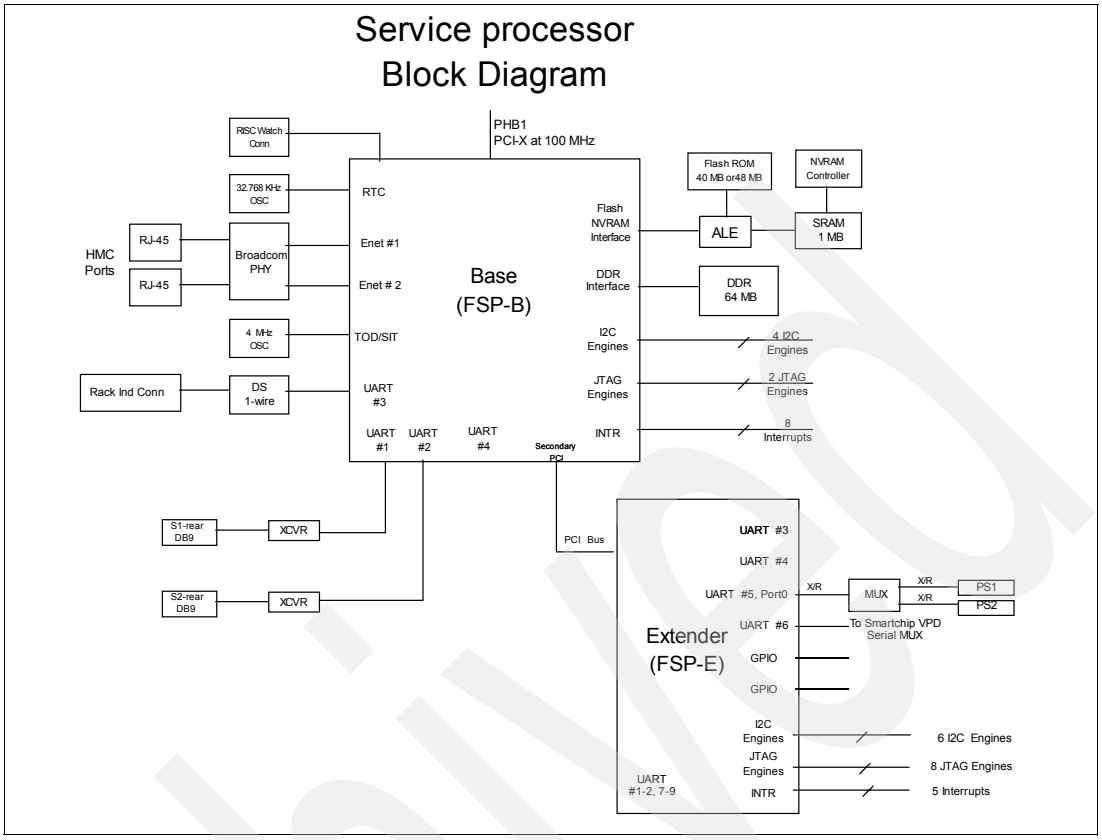


Figure 2-12 Service processor block diagram

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

FSP-B has four UART cores, which provide a full duplex serial interface. As shown in Figure 2-12, UART #1 and UART #2 are used for RS232 Serial Port #1 and RS232 Serial Port #2, respectively. UART #3 is used for Rack VPD/Light interface. UART #4 is not used.

2.9.6 Hardware management user interfaces

In the following sections we will give you a brief overview of the different OpenPower 710 server hardware management user interfaces available.

Advanced System Management Interface

The Advanced System Management Interface (ASMI) is the interface to the service processor that allows you to set flags that affect the operation of the server, such as auto power restart, and to view information about the server, such as the error log and vital product data.

This interface is accessible using a Web browser on a client system that is connected to the service processor on an Ethernet network. It can also be accessed using a terminal attached to a serial port on the server. The service processor and the ASMI are standard on all IBM ℮server i5, ℮server p5, and OpenPower servers.

You may be able to use the service processor's default settings. In that case, accessing the ASMI is not necessary.

Accessing the ASMI using a Web browser

The Web interface to the Advanced System Management Interface is accessible through Microsoft Internet Explorer 6.0, Netscape 7.1, or Opera 7.23 running on a PC or mobile computer connected to the service processor. The Web interface is available during all phases of system operation including the initial program load (IPL) and run time. However, some of the menu options in the Web interface are unavailable during IPL or run time to prevent usage or ownership conflicts if the system resources are in use during that phase.

Accessing the ASMI using an ASCII console

The Advanced System Management Interface on an ASCII console supports a subset of the functions provided by the Web interface and is available only when the system is in the platform standby state. The ASMI on an ASCII console is not available during some phases of system operation, such as the initial program load and run time.

Accessing the ASMI using a HMC

To access the Advanced System Management Interface using the Hardware Management Console, complete the following steps:

1. Ensure that the HMC is set up and configured.
2. In the navigation area, expand the managed system you want to work with.
3. Expand Service Applications and click **Service Focal Point**.
4. In the content area, click **Service Utilities**.
5. From the Service Utilities window, select the managed system you want to work with.
6. From the Selected menu on the Service Utilities window, select **Launch ASM menu**.

For more detailed information about usage of ASMI please refer to the IBM @server Hardware Information Center.

http://publib.boulder.ibm.com/infocenter/eserver/v1r2s/en_US/index.htm?info/iphau/usingsms.htm

System management services

Use the system management services (SMS) menus to view information about your system or partition, and to perform tasks such as setting a password, changing the boot list, and setting the network parameters.

To start the system management services, do the following:

1. For a server that is connected to an HMC, use the HMC to restart the server or partition.

   If the server is not connected to an HMC, stop the system, and then restart the server by pressing the power button on the control panel.
2. For a partitioned server, watch the virtual terminal window on the HMC.

   For a full server partition, watch the firmware console.
3. Look for the POST[4] indicators memory, keyboard, network, scsi, and speaker, which appear across the bottom of the screen. Press the numeric 1 key after the word keyboard appears, and before the word speaker appears.

For more detailed information about usage of SMS please refer to the IBM @server Hardware Information Center.

http://publib.boulder.ibm.com/infocenter/eserver/v1r2s/en_US/index.htm?info/iphau/usingsms.htm

[4] POST stands for Power-On-Self-Test

HMC

The Hardware Management Console is a system that controls managed systems, including IBM @server hardware, logical partitions, and Capacity on Demand. To provide flexibility and availability, there are different ways to implement HMCs, including the local HMC, remote HMC, redundant HMC, and the Web-based System Manager Remote Client.

Local HMC

This is a local HMC is any physical HMC that is directly connected to the system it manages through a private service network. An HMC in a private service network is a DHCP[5] server from which the managed system obtains the address for its firmware. Additional local HMCs in your private service network are DHCP clients.

Remote HMC

A stand-alone HMC or an HMC installed in a rack that is used to remotely access another HMC. A remote HMC may be present in an open network.

Redundant HMC

A redundant HMC manages a system that is already managed by another HMC. When two HMCs manage one system, those HMCs are peers and can be used simultaneously to manage the system. The redundant HMC in your private service network is usually a DHCP client.

Web-based System Manager remote client

The Web-based System Manager Remote Client is an application that is usually installed on a PC. You can then use this PC to access other HMCs remotely. Web-based System Manager Remote Clients can be present in private and open networks. You can perform most management tasks using the Web-based System Manager Remote Client.

The remote HMC and the Web-based System Manager Remote Client allow you the flexibility to access your managed systems (including HMCs) from multiple locations using multiple HMCs.

For more detailed information about usage of the HMC please refer to the IBM @server Hardware Information Center.

http://publib.boulder.ibm.com/infocenter/eserver/v1r2s/en_US/index.htm?info/iphau/usingsms.htm

Open Firmware

An IBM @server p5 and OpenPower servers have one instance of Open Firmware, both when in the partitioned environment and when running as a full system partition. Open Firmware has access to all devices and data in the system. Open Firmware is started when the system goes through a power-on reset. Open Firmware, which runs in addition to the Hypervisor in a partitioned environment, runs in two modes: Global and partition. Each mode of Open Firmware shares the same firmware binary that is stored in the flash memory.

In a partitioned environment, Open Firmware runs on top of the global Open Firmware instance. The partition Open Firmware is started when a partition is activated. Each partition has its own instance of Open Firmware and has access to all the devices assigned to that partition. However, each instance of Open Firmware has no access to devices outside of the partition in which it runs. Partition firmware resides within the partition memory and is replaced when Linux takes control. Partition firmware is needed only for the time that is necessary to load Linux into the partition system memory.

[5] DHCP stands for Dynamic Host Control Protocol.

The global Open Firmware environment includes the partition manager component. That component is an application in the global Open Firmware that establishes partitions and their corresponding resources (such as CPU, memory, and I/O slots), which are defined in partition profiles. The partition manager manages the operational partitioning transactions. It responds to commands from the service processor external command interface that originate in the application that is running on the HMC.

The ASMI can be accessed during boot time. Use the ASMI to boot to an Open Firmware prompt.

For more information on Open Firmware refer to *Partitioning Implementations for IBM* e*server p5 Servers,* SG24-7039-02, at:

http://www.redbooks.ibm.com/redpieces/abstracts/SG247039.html?Open

3

Reliability, availability, and serviceability

The following chapter provides more detailed information on IBM @server OpenPower 710 server reliability, availability, and serviceability features.

3.1 Reliability, fault tolerance, and data integrity

The reliability of the OpenPower 710 server starts with components, devices, and subsystems that are designed to be fault-tolerant. During the design and development process, subsystems go through rigorous verification and integration testing processes. During system manufacturing, systems go through a thorough testing process designed to help ensure the highest level of product quality.

- The OpenPower 710 server L3 cache and system memory offers ECC (error checking and correcting) fault-tolerant features. ECC is designed to correct environmentally induced, single-bit, intermittent memory failures and single-bit hard failures. With ECC, the likelihood of memory failures will be substantially reduced.
- ECC also provides double-bit memory error detection that helps protect data integrity in the event of a double-bit memory failure.
- System memory also provides 4-bit packet error detection that helps to protect data integrity in the event of a DRAM chip failure.
- The system bus, I/O bus, and PCI buses are designed with parity error detection.
- Linux supports disk mirroring (RAID 1). This is supported in software using the md driver. Some of the hardware RAID adapters supported under Linux also support mirroring.
- The Journaled File System maintains file system consistency and reduces the likelihood of data loss when the system is abnormally halted due to a power failure.

3.1.1 Memory error correction extensions

The OpenPower 710 server uses Error Checking and Correcting (ECC) circuitry for memory reliability, fault tolerance, and integrity.

- Memory has single-error-correct and double-error-detect ECC circuitry designed to correct single-bit memory failures. The *double-bit* detection is designed to help maintain data integrity by detecting and reporting multiple errors beyond what the ECC circuitry can correct.
- The memory chips are organized such that the failure of any specific memory module only affects a single-bit within an ECC word (*bit-scattering)*, thus allowing for error correction and continued operation in the presence of a complete chip failure (Chipkill™ recovery).
- The memory also utilizes memory scrubbing and thresholding to determine when spare memory modules, within each bank of memory, if available, should be used to replace ones that have exceeded their threshold value (*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.

3.1.2 Redundancy for array self-healing

Although the most likely failure event in a processor is a soft single-bit error in one of its caches, there are other events that can occur, and they need to be distinguished from one another.

- For the L1, L2, and L3 caches and their directories, hardware and firmware keep track of whether permanent errors are being corrected beyond a threshold. If this threshold is exceeded, a deferred repair error log is created. Additional run-time availability actions, such as CPU vary off[1] or L3 cache line delete, are also initiated.

[1] This RAS function is only available for a Linux operating system running the 2.6 kernel.

- L1 and L2 caches and L2 and L3 directories on the POWER5 chip are manufactured with spare bits in their arrays that can be accessed via programmable steering logic to replace faulty bits in the respective arrays. This is analogous to the redundant bit-steering employed in main storage as a mechanism that is designed to help avoid physical repair, and is also implemented in POWER5 systems. The steering logic is activated during processor initialization and is initiated by the built-in self-test (BIST) at power-on time.
- L3 cache redundancy is implemented at the cache line level. Exceeding correctable error thresholds while running causes a dynamic L3 cache line delete function to be invoked.

3.1.3 Service processor

The service processor included in the OpenPower 710 server is designed for an immediate means to diagnose, check status, and sense operational conditions of a remote system, even when the main processor is inoperable.

- The service processor enables firmware and operating system surveillance, several remote power controls, environmental monitoring (only critical errors are supported under Linux), reset, boot features, remote maintenance, and diagnostic activities, including console mirroring.
- The service processor can place calls to report surveillance failures, critical environmental faults, and critical processing faults.

For more detailed information on the service processor refer to 2.9.5, “Service processor” on page 44.

3.1.4 Fault monitoring functions

The following are a few of the fault monitoring systems included with an OpenPower 710 server.

- BIST and power-on self-test (POST) check the processor, L3 cache, memory, and associated hardware required for proper booting of the operating system every time the system is powered on. If a noncritical error is detected or if the errors occur in the resources that can be removed from the system configuration, the booting process is designed to proceed to completion. The errors are logged in the system nonvolatile RAM (NVRAM).
- Disk drive fault tracking can alert the system administrator of an impending disk failure before it impacts client operation.
- The Linux log (where hardware and software failures are recorded and analyzed by the Error Log Analysis (ELA) routine) warns the system administrator about the causes of system problems. This also enables service representatives to bring along probable replacement hardware components when a service call is placed, thus minimizing system repair time.

3.1.5 Mutual surveillance

The service processor monitors the operation of the POWER Hypervisor firmware during the boot process and watches for loss of control during system operation. It also allows the POWER Hypervisor to monitor service processor activity.

The service processor can take appropriate action, including calling for service, when it detects the POWER Hypervisor firmware has lost control. Likewise, the POWER Hypervisor can request a service processor repair action if necessary.

3.1.6 First Failure Data Capture

Diagnosing problems in a computer is a critical requirement for autonomic computing. The first step to producing a computer that truly has the ability to self-heal is to create a highly accurate way to identify and isolate hardware errors. IBM has implemented a server design that builds in hardware error-check stations that capture and help to identify error conditions within the server. Each of these checkers is viewed as a diagnostic probe into the server, and, when coupled with extensive diagnostic firmware routines, allows quick and accurate assessment of hardware error conditions at run-time.

- First Failure Data Capture (FFDC) check stations are carefully positioned within the server logic and data paths to help ensure that potential errors can be quickly identified and accurately tracked to an individual field replaceable unit (FRU).
- These checkers are collected in a series of Fault Isolation Registers, where they can be accessed by the service processor.
- All communication between the service processor and monitored components is accomplished *out of band*. That is, operation of the error-detection mechanism is transparent to an operating system. This entire structure is *below the architecture* and is not seen, nor accessed, by system-level activities.

3.1.7 Environmental monitoring functions

The following are some of the environmental monitoring functions available for an OpenPower 710 server.

- Temperature monitoring increases the fan speed rotation when ambient temperature is above the normal operating range.
- Temperature monitoring warns the system administrator of potential environmental-related problems (for example, air conditioning and air circulation around the system) so that appropriate corrective actions can be taken before a critical failure threshold is reached. It also performs an orderly system shutdown when the operating temperature exceeds the critical level.
- Fan speed monitoring provides a warning and an orderly system shutdown when the speed is out of the operational specification.
- Voltage monitoring provides a warning and an orderly system shutdown when the voltages are out of the operational specification.

3.1.8 Error handling and reporting

In the unlikely event of system hardware or environmentally induced failure, the system run-time error capture capability systematically analyzes the hardware error signature to determine the cause of failure.

- The analysis will be stored in the system NVRAM. When the system can be successfully rebooted either manually or automatically, the error will be reported to the Linux operating system.
- Error Log Analysis can be used to display the failure cause and the physical location of failing hardware.
- With the integrated service processor, the system has the ability to automatically send out an alert via phone line to a pager or call for service in the event of critical system failure. A hardware fault will also turn on the two Attention Indicators (one located on the front of the system unit and the other on the rear of the system) to alert the user of an internal hardware problem. The indicator may also be turned on by the operator as a tool to allow

system identification. For identification, the indicators will flash, whereas the indicator will be on solid when an error condition occurs.

3.1.9 Availability enhancement functions

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

3.2 Serviceability

The OpenPower 710 server is designed for client setup of the machine and for subsequent addition of most features (adapters/devices). For a fee, IBM Service can perform the installation.

- The OpenPower 710 server allows clients to replace service parts (Customer Replaceable Unit) if they want to. The OpenPower 710 server has incorporated LEDs that will indicate the parts needing to be replaced.
- The OpenPower 710 server allows support personnel to remotely log into a system to review error logs and perform remote maintenance. The OpenPower 710 service processor enables the analysis of a system that will not boot.
- The diagnostics consist of Stand-alone Diagnostics, which are loaded from the DVD-ROM drive.
- The System Management Services (SMS) error log is accessible from the SMS menu for tests performed through SMS programs. For results of service processor tests, access the error log from the service processor menu.

3.2.1 Service Agent

The Service Agent is available at no additional charge. When installed on an IBM *e*server system, the Service Agent can enhance IBM's ability to provide the system with maintenance service.

The Service Agent:

- Monitors and analyzes system errors and, if needed, can automatically place a service call to IBM without client intervention
- Can help reduce the effect of business disruptions due to unplanned system outages and failures
- Performs problem analysis on a subset of hardware-related problems and, with client authorization, can automatically report the results to IBM Service

3.3 Clustering and high availability

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 building blocks, with availability, scalability, security, and single-point-of-management control, to satisfy these needs. The advantages of clusters are:

- Large-capacity data and transaction volumes, including support of mixed workloads
- Scale-up (add processors) or scale-out (add servers) without downtime

- Single point-of-control for distributed and clustered server management
- Simplified use of IT resources
- Designed for 24x7 access to data applications
- Business continuity in the event of disaster

Cluster Systems Management (CSM) V1.4.1 for Linux on POWER is supported on the OpenPower 710 servers running SUSE LINUX Enterprise Server Version 9 or later and Red Hat Enterprise Server Version 3 or later. For hardware control, a Hardware Management Console (HMC) is required.

For details on IBM cluster software for Linux see the following URL:

http://www-1.ibm.com/servers/eserver/clusters/library/linux.html

The Beowulf clustering technology and other open source and some commercial products can be used to cluster POWER systems running Linux to provide compute or high-availability clusters.

For details on Beowulf see the following URL:

http://www.beowulf.org

Myricom has the Myrinet switch available for Linux on IBM @server p5, pSeries, and OpenPower systems. The Linux distributions that support the respective technologies also support the Myrinet switch—specifically SLES 8 systems, SLES 9 systems, and RHEL AS 3 systems. It can be used as a high-speed interconnect to cluster systems of IBM @server p5, pSeries, and OpenPower machines running Linux. Gigabit or 10/100 Ethernet connections can also be used.

For details on Myricom see the following URL:

http://www.myri.com/

In the high-availability arena, IBM provides Tivoli® System Automation as one solution. This product is based on technology from IBM's mainframe z/OS® and AIX 5L high-availability products. Other third-party and open source solutions are also available.

For details on Tivoli System Automation see the following URL:

http://www-306.ibm.com/software/tivoli/products/sys-auto-linux/

For details on operating systems supported with Tivoli System Automation, see the following URL:

http://www-306.ibm.com/software/tivoli/products/sys-auto-linux/platforms.html

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 on 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.

- *Partitioning Implementations for IBM @server p5 Servers,* SG24-7039-02
- *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 @server pSeries Clusters*, SG24-7014
- *Understanding IBM @server pSeries Performance and Sizing*, SG24-4810
- *Advance POWER Virtualization on IBM @server p5 Servers,* SG24-7940
- *IBM @server OpenPower 720 Technical Overview and Introduction*, REDP-1965-01

Online resources

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

- IBM @server Hardware Information Center:
  http://publib16.boulder.ibm.com/infocenter/eserver/v1r2s/en_US/index.htm
- IBM Offering Information, including IBM Announcements and Sales Manuals:
  http://www-306.ibm.com/common/ssi/OIX.wss
- IBM Microcode download Web page:
  http://techsupport.services.ibm.com/server/mdownload
- IBM @server pSeries and RS/6000 microcode update:
  http://techsupport.services.ibm.com/server/mdownload2/download.html
- IBM TotalStorage Web page:
  http://www.ibm.com/servers/storage/
- The IBM TotalStorage DS4000 Storage server family:
  http://www.ibm.com/servers/storage/disk/ds4000/index.html
- IBM TotalStorage Enterprise Storage Server:
  http://www.ibm.com/servers/storage/disk/enterprise/ds_family.html

- IBM 2104 Expandable Storage Plus:

  http://www.ibm.com/servers/storage/disk/expplus/index.html
- POWER Hypervisor firmware enablement code download page:

  http://www.ibm.com/servers/eserver/openpower/cod
- Virtual I/O Server supported environment:

  http://techsupport.services.ibm.com/server/virtualization/vios/documentation/datasheet.html
- HMC support Web page:

  https://techsupport.services.ibm.com/server/hmc/power5
- IBM Service Tools for Linux download page:

  http://techsupport.services.ibm.com/server/lopdiags
- Supported features and external devices on the OpenPower 710 server:

  http://www.ibm.com/servers/eserver/pseries/linux/
- Information about SUSE LINUX Enterprise Server 9 can be found at:

  http://www.novell.com/products/linuxenterpriseserver/
- Information about Red Hat Enterprise Linux AS can be found at:

  http://www.redhat.com/software/rhel/details/
- For the latest in IBM Linux news, subscribe to the Linux Line. See:

  https://www14.software.ibm.com/webapp/iwm/web/preLogin.do?source=linuxline
- Many of the features described in this document are operating system dependant and may not be available on Linux. For more information, see:

  http://www.ibm.com/servers/eserver/linux/power/whitepapers/linux_overview
- LPAR validation tool

  http://www.ibm.com/servers/eserver/iseries/lpar/systemdesign.htm
- Client specific placement

  http://www.ibm.com/servers/eserver/power/csp/index.html
- For details on IBM cluster software for Linux see the following URL:

  http://www-1.ibm.com/servers/eserver/clusters/library/linux.html
- For details on Tivoli System Automation see the following URL:

  http://www-306.ibm.com/software/tivoli/products/sys-auto-linux/
- For details on Myricom see the following URL:

  http://www.myri.com/
- For details on Beowulf see the following URL:

  http://www.beowulf.org

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 710
Technical Overview and Introduction

Redpaper

Finer system granulation using Micro-Partitioning technology to help lower TCO

Outstanding performance based on IBM POWER5 processor technology

From Web servers to integrated cluster solutions

This document is a comprehensive guide covering the IBM @server OpenPower 710 server. 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:

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