IBM

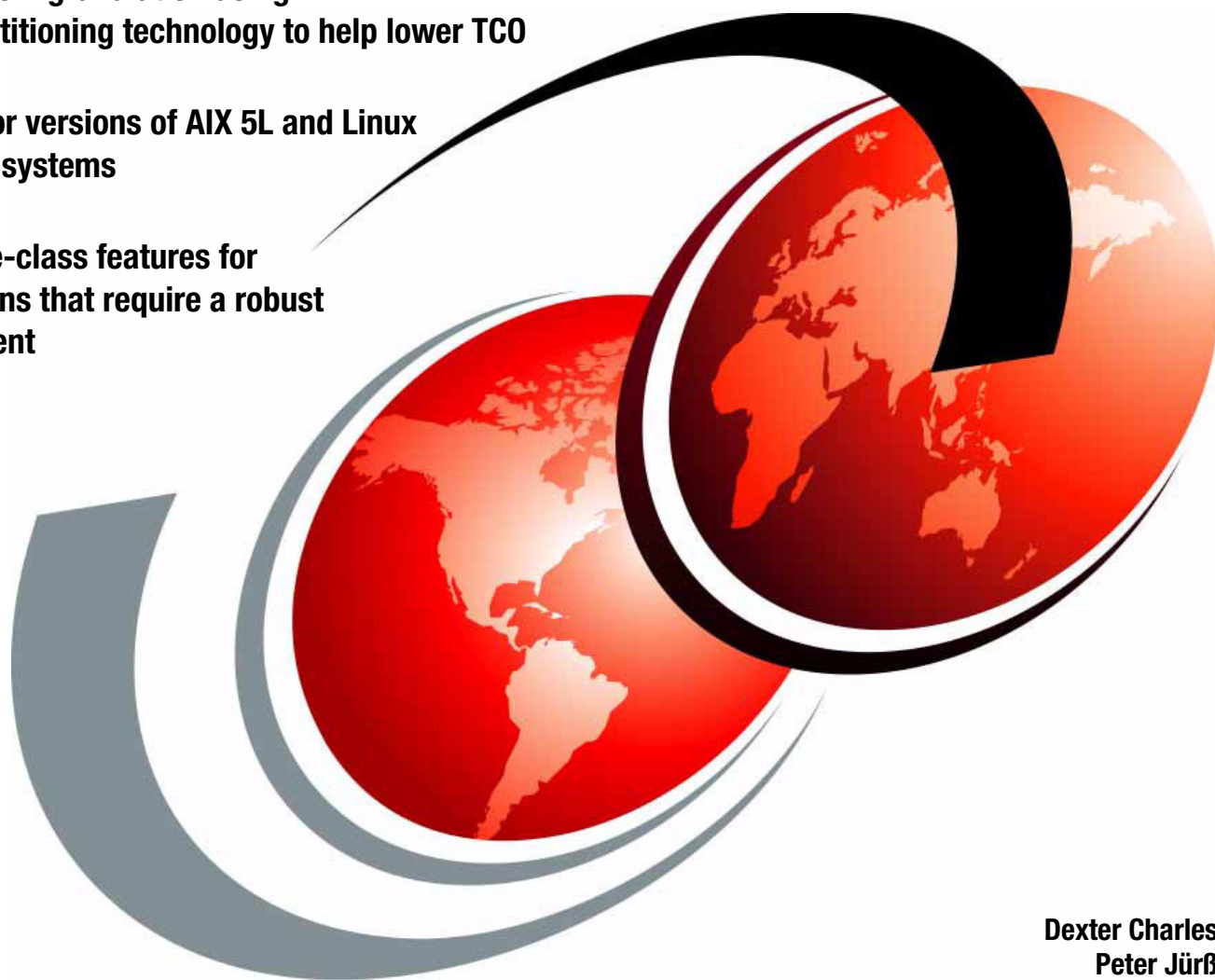# IBM @server p5 510 Technical Overview and Introduction

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

Support for versions of AIX 5L and Linux operating systems

Enterprise-class features for applications that require a robust environment

Dexter Charles
Peter Jürß

Redpaper

International Technical Support Organization

IBM @server **p5 510**
Technical Overview and Introduction

May 2005

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

First Edition (May 2005)

This edition applies to the IBM @server p5 510 and AIX 5L Version 5.3, product number 5765-G03.

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®	Enterprise Storage Server®	POWER5™
@server®	Hypervisor™	PS/2®
ibm.com®	HACMP™	Redbooks™
iSeries™	IBM®	Redbooks (logo) ™
pSeries®	Micro-Partitioning™	RS/6000®
AIX 5L™	OpenPower™	TotalStorage®
AIX®	POWER™	UltraNav®
Chipkill™	POWER4™	Virtualization Engine™
Electronic Service Agent™	POWER4+™	WebSphere®

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, 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 p5 510 AIX® 5L™ and Linux® operating system servers. It introduces major hardware offerings and discusses their prominent functions.

Professionals wishing to acquire a better understanding of IBM @server p5 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 p5-510 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 of experience 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™ systems, 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 environments.

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 redbook form found at:

> **ibm.com**/redbooks

► Send your comments in an email 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

General description

IBM @server p5 510 and IBM @server p5 510 Express rack-mount servers (9110-510) give you new tools for managing on demand business, greater application flexibility, and innovative technology—all designed to help you capitalize on the on demand business revolution. To simplify naming, both products are referred to as "p5-510 server" except where a distinction needs to be made for the Express offering.

The p5-510 Express comes in a 2U rack drawer package. It is available in a 1-way or 2-way configuration using state-of-the-art, 64-bit, copper-based and silicon on insulator (SOI) IBM POWER5™ microprocessors running at 1.5 GHz.

The p5-510 server is similar to the p5-510 Express server except that the POWER5 microprocessor operates at 1.65 GHz. The POWER5 microprocessor provides performance and reliability advances (or enhancements) over the POWER4™ architecture. Chief among the enhancements is simultaneous multi-threading, which is designed to boost performance by allowing two code streams to concurrently access the multiple processor execution units.

The p5-510 server has a base of 512 MB of main memory that can be expanded to 32 GB for faster performance and exploitation of 64-bit addressing as used in large database applications. The p5-510 server contains five internal 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 available 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. Dynamic LPAR is supported on the p5-510 server, allowing up to two logical partitions. In addition, the optional Advanced POWER Virtualization feature supports up to 20 micro-partitions using Micro-Partitioning™ technology.

IBM Cluster Systems Management (CSM) V1.4 for AIX 5L and Linux on POWER is supported on the p5-510 server.

The p5-510 server is backed by a three-year limited warranty. Check with your IBM representative for particular warranty availability in your region.

1.1 System specifications

Table 1-1 lists the general system specifications of the p5-510 server.

Table 1-1   IBM eServer p5-510 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 p5-510 server.

Table 1-2   IBM eServer p5-510 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 p5-510 server rack-mount model

Figure 1-1 shows a p5-510 server that has been removed from a rack.

Figure 1-1   The p5-510 server rack-mount server

The p5-510 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 p5-510 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 p5-510 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 p5-510 rack-mount server, including connectors, location codes, SCSI IDs, and components.
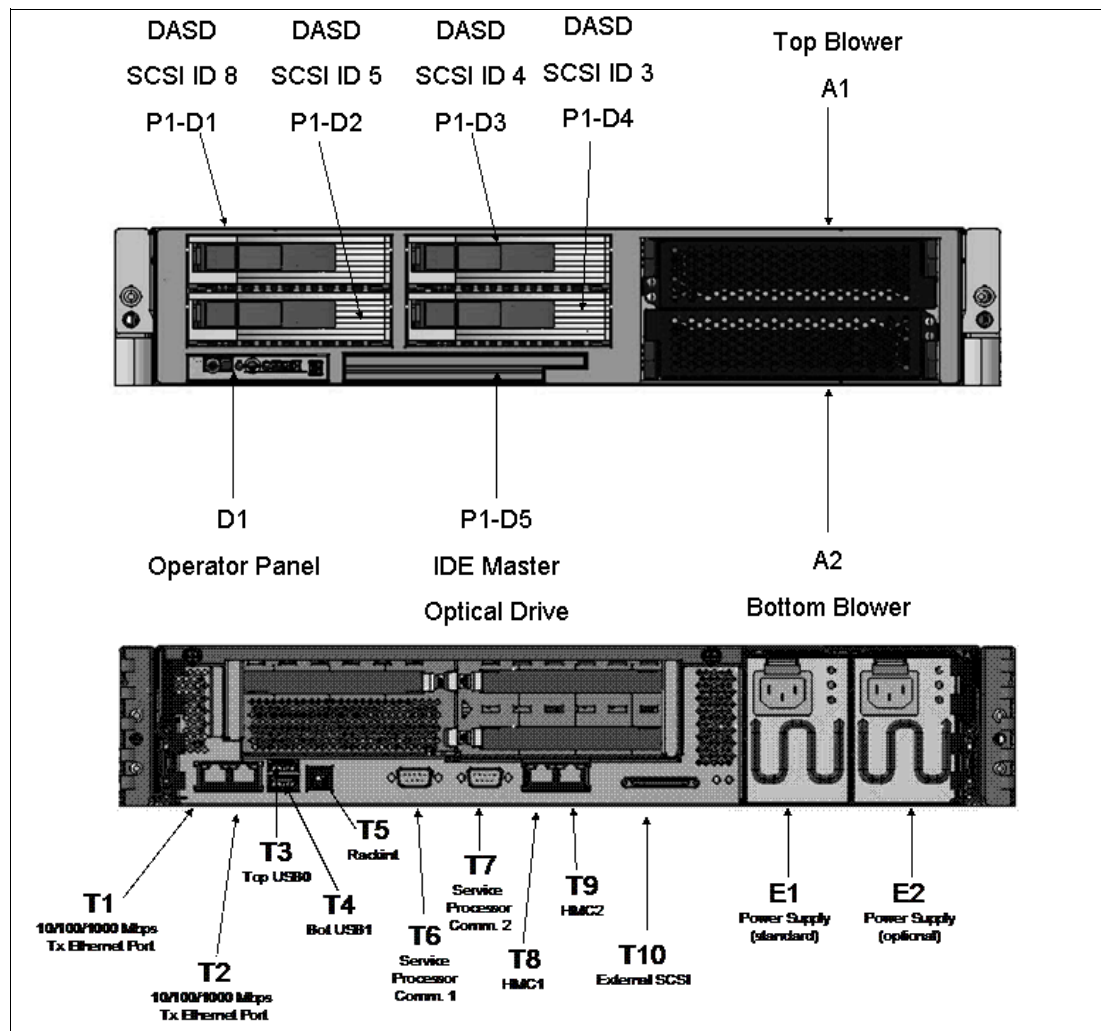
Figure 1-2   Front and rear view of the p5-510 server including location codes

1.4 Minimum and optional features

IBM @server p5 510 and p5 510 Express systems are based on a flexible, modular design featuring:

► 1-way and 2-way symmetric multiprocessor (SMP) design using one POWER5 chip packaged in a processor module soldered directly to the system planar

► 512 MB 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

► Service processor

► Hot-plug and redundant fans

The p5-510 server, supports the following integrated ports:

► Dual port 10/100/1000 Ethernet

► Dual channel Ultra320 SCSI controller (one internal and one external VHDCI LVD connector)

► Two USB ports

► Two system ports

► Two HMC ports

The p5-510 server supports AIX 5L and Linux operating systems (OS) and requires the following specific levels:

► AIX 5L Version 5.2 or Version 5.3, or later

► SUSE LINUX Enterprise Server 9 (SLES 9) for POWER, or later

► Red Hat Enterprise Linux AS (RHEL AS 3) for POWER Version 3, or later

1.4.1 Processor features

The p5-510 server features one POWER5 chip with one or two active processor cores running at 1.5 GHz or 1.65 GHz. The p5-510 server is available as a one or two way system. For a list of available processor features, refer to Table 1-3.

Table 1-3 Available processor options

Feature code	Description
7609	1-way 1.5 GHz POWER5 Processor Card, No L3 Cache
7610	1-way 1.65 GHz POWER5 Processor Card, 36 MB L3 Cache
7611	2-way 1.65 GHz POWER5 Processor Card, 36 MB L3 Cache
7612	2-way 1.5 GHz POWER5 Processor Card, 36 MB L3 Cache

The p5-510 server POWER5 chip is mounted on the system planar and directly interfaced to the memory buffer SMI-II chips.

1.4.2 Memory features

The minimum memory requirement for the p5-510 server is 512 MB, and the maximum capacity is 32 GB. Memory DIMMs are installed into eight DIMM sockets located on the system planar. FC 4443 is the only DIMM that can be installed in pairs. The supported memory is 266 MHz DDR1 DIMMs. Table 1-4 lists the available memory features.

Table 1-4   Memory feature codes

Feature code	Description
4443	512 MB (2 x 256 MB) DIMMS, 208-pin, 266 MHz DDR SDRAM
4444	1024 MB (4 x 256 MB) DIMMS, 208-pin, 266 MHz DDR SDRAM
4447	2048 MB (4 x 512 MB) DIMMS, 208-pin, 266 MHz DDR SDRAM
4445	4096 MB (4 x 1024 MB) DIMMS, 208-pin, 266 MHz DDR SDRAM
4449	8192 MB (4 x 2048 MB) DIMMS, 208-pin, 266 MHz DDR SDRAM
4450	16384 MB (4 x 4096 MB) DIMMS, 208-pin, 266 MHz DDR SDRAM

1.4.3 Disk and media features

The p5-510 server features four disk bays and one slimline media bay. The minimum configuration requires at least one disk drive. Table 1-5 shows the disk drive feature codes that each bay can contain.

Table 1-5   Hot-swappable disk drive options

Feature code	Description
3277	36.4 GB 15,000 rpm Ultra320 SCSI hot-swappable disk drive
3274	73.4 GB 10,000 rpm Ultra320 SCSI hot-swappable disk drive
3278	73.4 GB 15,000 rpm Ultra320 SCSI hot-swappable disk drive
3275	146.8 GB 10,000 rpm Ultra320 SCSI hot-swappable disk drive
3578	300 GB 10,000 rpm Ultra320 SCSI hot-swappable disk drive

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

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

A logical partition running a supported release of the Linux operating system requires a DVD drive to provide a method to run the hardware diagnostic from the CD. Concurrent diagnostics, as provided by AIX 5L, is not available on Linux at the time of writing.

1.4.4 USB diskette drive

In some situations, an external USB 1.44 MB diskette drive for p5-510 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.

1.4.5 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 include:

- ► 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 36 for detailed information on the HMC.

Table 1-6 lists the HMC options for POWER5 processor-based systems available at the time of writing.

Table 1-6   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 Value Paks and Value Paks

If you order a p5-510 Express with a 1-way POWER5 1.5 GHz processor (FC 7609) or a 2-way POWER5 1.5 GHz processor (FC 7612), or a p5-510 server with a 1-way POWER5 1.65 GHz processor (FC 7610) or a 2-way POWER5 1.65 GHz processor (FC 7611) and a minimum configuration of memory, disks, and media device, you will qualify for a processor entitlement at no additional charge. A processor entitlement specifies the number of activated processors in the system. The number of processors, total memory, size of disk, and presence of a media device are the only features that determine if a configuration is entitled to processor entitlements at no additional charge. You will be entitled to a specially priced AIX 5L operating system license or may choose to purchase the system with no operating system. The specially priced AIX 5L operating system license is processed using a feature number of AIX 5L.

The minimum criteria for processor entitlements at no additional charge are 1 GB of memory per active processor, two 73.4 GB 10K rpm hard disk drives, and a DVD-ROM. A DVD-RAM may be substituted for the DVD-ROM.

You may make changes to the standard features as needed and still qualify for processor entitlements at no additional charge and a specially priced AIX 5L operating system license. Selection of memory or disk drives smaller than those defined as the Value Pak minimums will disqualify the order as a Value Pak.

Note: Express Value Paks and Value Paks are only available as part of an initial order.

Table 1-7 summarizes the available Express Value Pak and Value Pak configurations and minimum requirements.

Table 1-7   Express Value Pak and Value Pak configurations

Description	Processors	Memory (MB)	Disk	DVD
Express Value Pak 1	1-way, FC 7609, 1.5 GHz	1024, FC 4444	2 x 73.4 GB disk drive, FC 3274	DVD-ROM or DVD-RAM
Express Value Pak 2	2-way, FC 7612, 1.5 GHz	2048, FC 4447	2 x 73.4 GB disk drive, FC 3274	DVD-ROM or DVD-RAM
Value Pak 3	1-way, FC 7610, 1.65 GHz	1024, FC 4444	2 x 73.4 GB disk drive, FC 3274	DVD-ROM or DVD-RAM
Value Pak 4	2-way, FC 7611, 1.65 GHz	2048, FC 4447	2 x 73.4 GB disk drive, FC 3274	DVD-ROM or DVD-RAM

1.6 System racks

The Enterprise Rack Models T00 and T42 are 19-inch wide racks for general use with IBM @server p5, OpenPower™, and pSeries rack-mounted servers. The racks provide increased capacity, greater flexibility, and improved floor space utilization.

If a IBM @server p5 system is to be installed in a non-IBM rack or cabinet, you must ensure that the rack conforms to the EIA[1] standard EIA-310-D (see "OEM rack" on page 11).

> **Note:** It is the client's responsibility to ensure that the installation of the drawer in the pre-ferred rack or cabinet results in a configuration that is stable, serviceable, safe, and com-patible 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 with past and present IBM @server p5, OpenPower, and pSeries systems. It fits 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.

---

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

- 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 9), but others can fit inside the rack. See 1.6.3, "AC Power Distribution Unit and rack content" on page 8.

- 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 a p5-510 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.

- 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** at the following Web site:

> 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 specific 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 client-usable 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 amp circuit breaker.

> **Note:** The maximum number of 9110-510 systems supported on a 30 amp service (PDU derated to 24 amps) is 9 per PDU.
>
> The maximum number of 9110-510 systems 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.
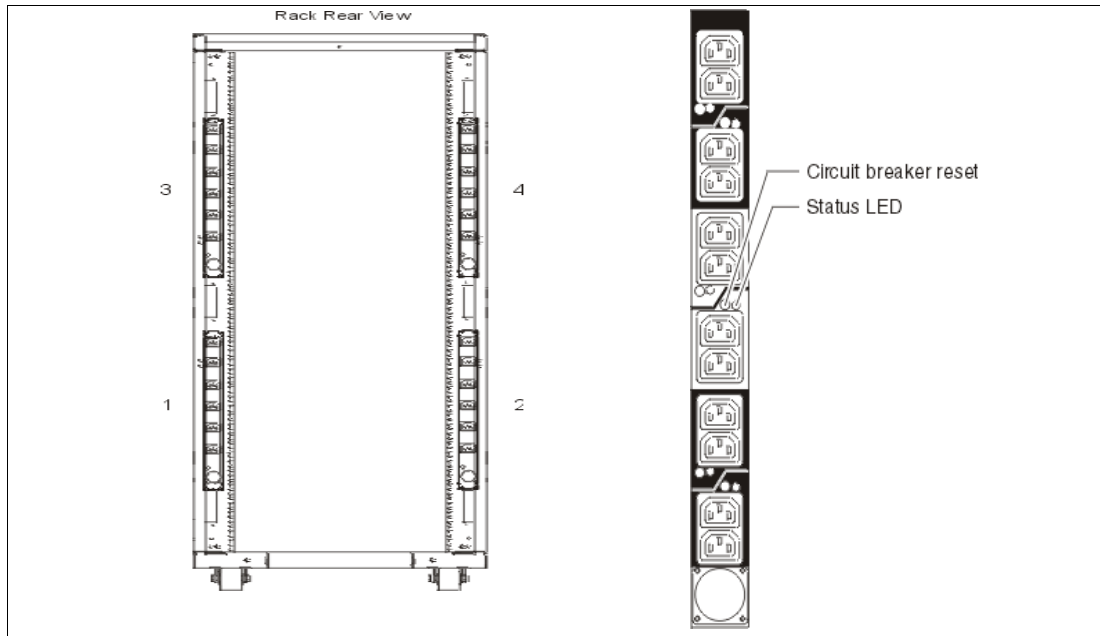
Figure 1-3   PDU placement and PDU view

1.6.4 Rack-mounting rules for p5-510 server

The p5-510 server is a 2U rack-mounted server drawer. The primary rules that should be followed when mounting the p5-510 server into a rack are:

► The p5-510 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 p5-510 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 p5-510 servers fit in the T00 rack, and 21 model p5-510 servers in T42 rack.

The 14 foot 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

This section highlights some solutions available to provide a single point of management for environments composed of multiple p5-510 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) 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[2] (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[3] 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.

---

[2] Initial Program Load
[3] Picture elements

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 p5 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 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 a p5-510 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 can be used with the Netbay LCM Switch 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 following URL:

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 some 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 36.

1.6.6 OEM rack

The p5-510 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 specification dimensions.
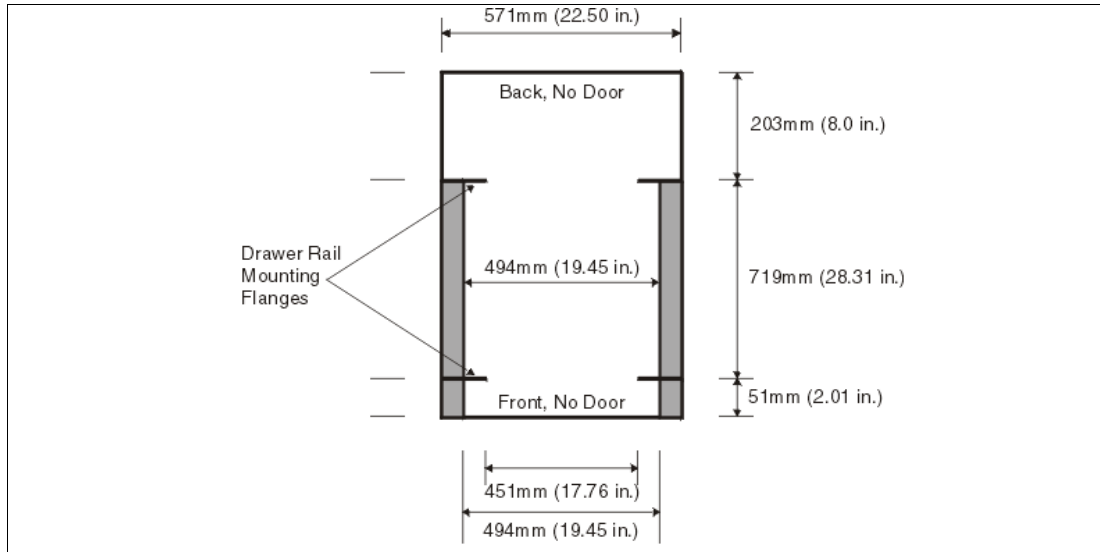
Figure 1-4   Top view non-IBM rack specification dimensions

► 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.
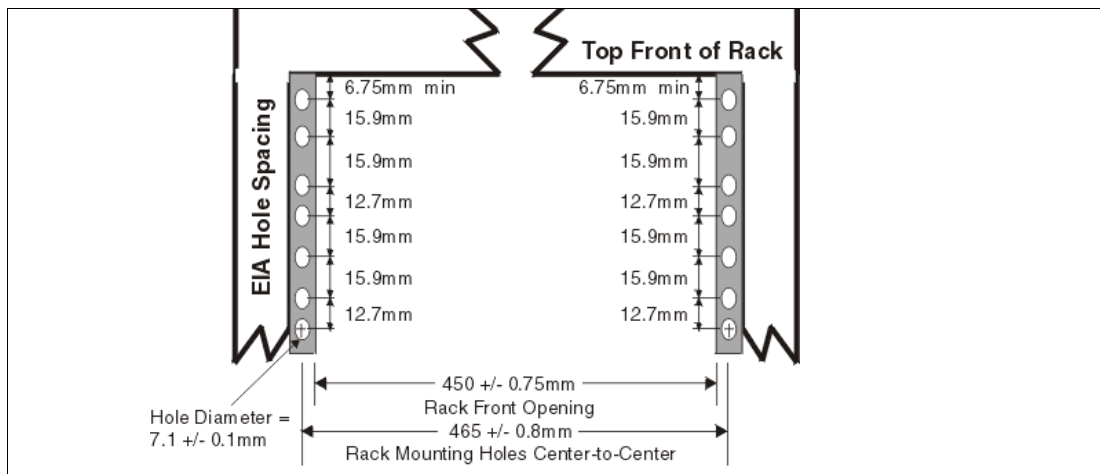
Figure 1-5   Rack specification dimensions, top front view
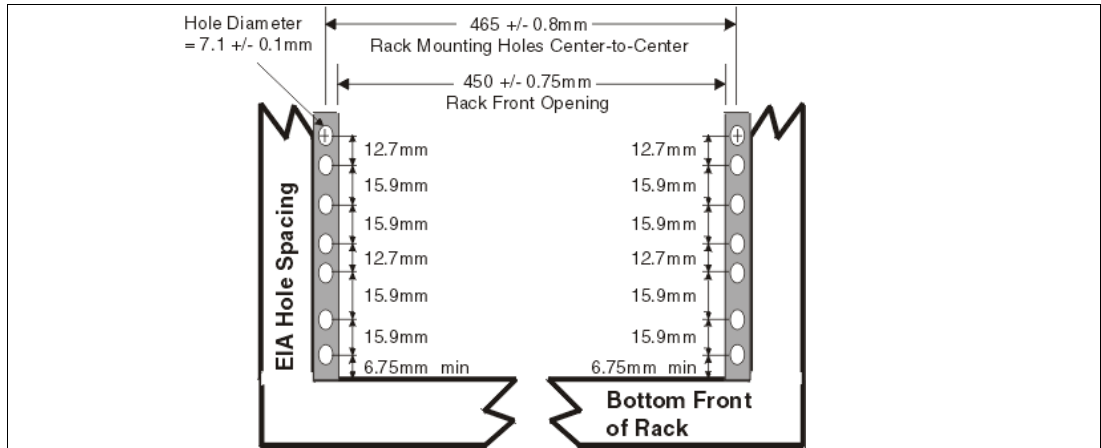
Figure 1-6 Rack specification dimensions, bottom front view

► 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 must meet the drawer power requirements, as well as the requirements of any additional products that will be connected to the same power distribution device.

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 recommend that you always obtain real-world performance measurements using production workloads.
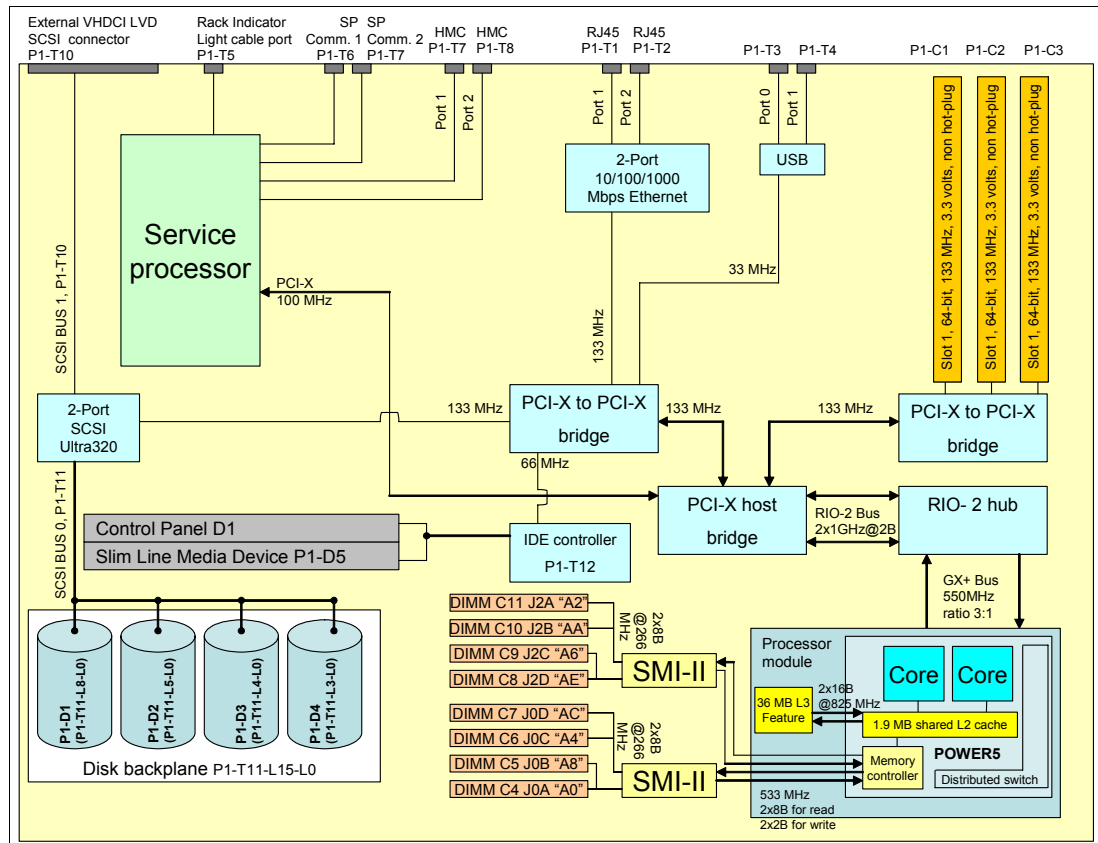
Figure 2-1   p5-510 server logic data flow

2.1 The POWER5 chip

The POWER5 chip features single and simultaneous multi-threading 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 compares the features and capabilities of POWER4 and POWER5 processors.

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; and 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. The 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 available L3 cache (not available on 1-way 1.5 GHz processor feature), with a capacity of 36 MB, operates as a back door with separate buses for reads and writes that operate at half processor speed.
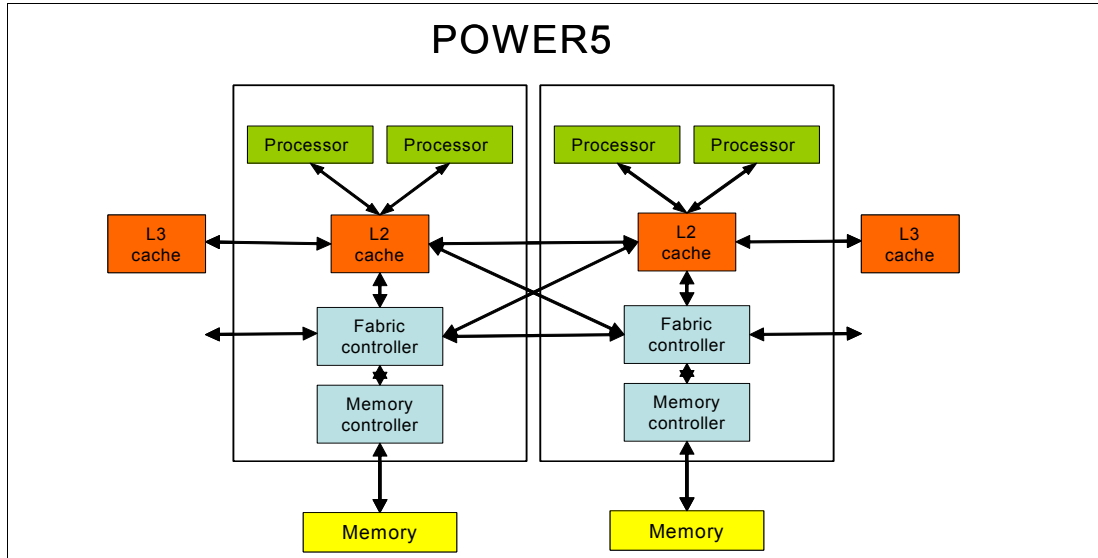
Figure 2-2   POWER5 system structure

Because of the higher transistor density of the POWER5 0.13-μm technology, it was possible to move the memory controller on chip and eliminate a chip previously needed for the memory controller function. These changes in the POWER5 processor chip 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.

▶ Support is provided 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 AIX 5L Version 5.3 or Linux operating system-based systems at an appropriate level.

IBM has documented simultaneous multi-threading performance benefit at 30 percent.

For more information, see the following Web site:

http://www.ibm.com/servers/eserver/pseries/hardware/system_perf.html

The simultaneous multi-threading policy is controlled by the operating system and is thus partition specific. AIX 5L provides the `smtctl` command that turns simultaneous multi-threading on and off without a subsequent reboot.

The `smtctl` command does not re-build the boot image. To change the simultaneous multi-threading mode of AIX 5L Version 5.3, the `bosboot` command must be used to rebuild the boot image. The boot image has been extended to include the simultaneous multi-threading indicator that controls the default simultaneous multi-threading mode.

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

Enhanced 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 reaching the threshold of L2 cache misses will be throttled to allow other threads to pass the stalled thread).

- ► Adjustable thread priority

  Adjustable thread priority 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 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 p5-510 server supports one processor module (either a 1-way or 2-way) and an integrated 36 MB L3 cache module. The 1-way 1.5 GHz version has no L3 cache. The 1-way 1.65 GHz version and all 2-way models have been packaged with the L3 cache chip into the processor module package.

> **Note:** Since the POWER5 processor modules are directly soldered to the system planar, special care must be taken in 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 p5-510 server 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 throughput 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/second.

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.
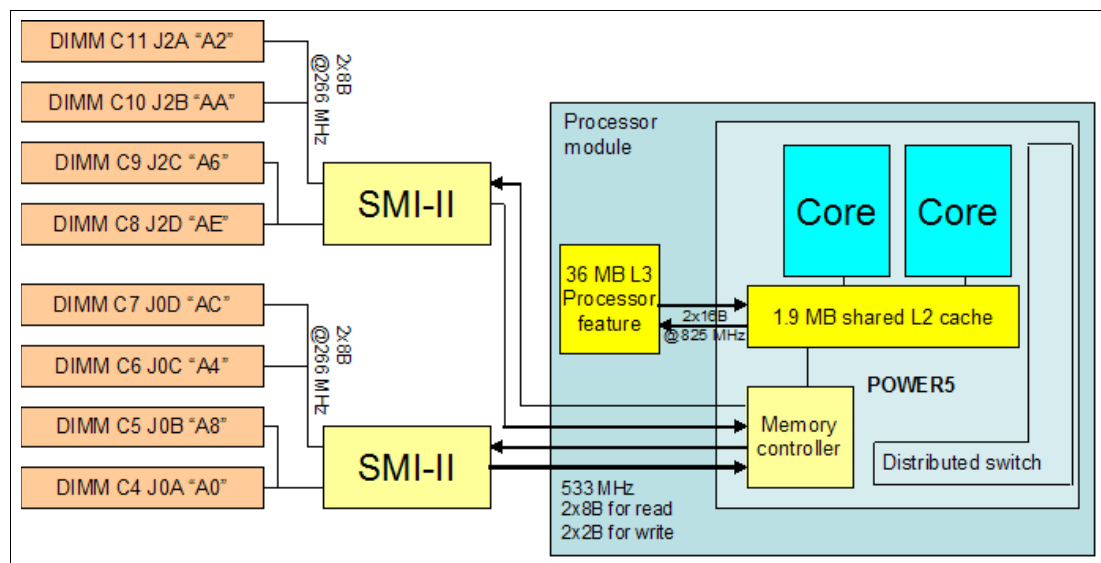

Figure 2-3   POWER5 processor core with DDR1 memory socket layout view

2.2.1 Available processor speeds

The p5-510 server operates a processor clock rate of either 1.5 GHz or 1.65 GHz.

To determine the processor characteristics on a running system, use one of the commands described in this section. The first of these commands is:

```
lsattr -El procX
```

Where X is the number of the processor, for example, proc0 is the first processor in the system. The output from the command would be similar to the following (False, as used in this output, signifies that the value cannot be changed through an AIX 5L command interface):

```
type powerPC_POWER5     Processor type        False
frequency 165600000     Processor Speed       False
smt_enabled true        Processor SMT enabled False
smt_threads 2           Processor SMT threads False
state enable            Processor state       False
```

The other command is:

```
pmcycles -m
```

This command (AIX 5L) uses the performance monitor cycle counter and the processor real-time clock to measure the actual processor clock speed in MHz. The following is the output of a 2-way p5-510 server running at 1.65 GHz system and simultaneous multi-threading functionality enabled:

```
Cpu 0 runs at 1656 MHz
Cpu 1 runs at 1656 MHz
```

> **Note:** The `pmcycles` command is part of the bos.pmapi fileset. This component must be installed before using the `lslpp -l bos.pmapi` command.

2.3 Memory subsystem

The p5-510 server offers pluggable DIMMs for memory. The system planar provides eight slots for up to eight pluggable DIMMs. The minimum memory for a p5-510 server is 512 MB and 32 GB is the maximum installable memory option. Figure 2-4 on page 21 shows the memory slot and location codes.

2.3.1 Memory placement rules

The memory features available at the time of writing for the p5-510 server are listed in "Memory features" on page 5.

Memory can be pluggable in pairs or quads, as required by the feature codes. The base memory feature (512 MB) consists of two DIMMs and higher memory features consist of four DIMMs, or quads. Memory feature numbers may be mixed within a system. The DIMMs slots are accessed by first removing the PCI riser book.

When additional memory is added to a system using FC 4443, an additional feature FC 4443 must be added to the original pair to make a quad, allowing one additional quad to be added to the system. Memory is installed in the first quad in the following order: J2A, J0A, J2C, and J0C; and for the second quad, in the order J2B, J0B, J2D, and J0D. Memory must be

balanced across the DIMM quad slots. The Service Information label, located on the top cover of the system, provides memory DIMMs slot location information.

To determine how much memory is installed in a system, use the following command:

```
# lsattr -El sys0 | grep realmem
realmem     524288      Amount of usable physical memory in Kbytes Fales
```

Note: A quad must consist of a single feature (that is, be made of identical DIMMs). Mixed DIMM capacities in a quad will result in reduced RAS.
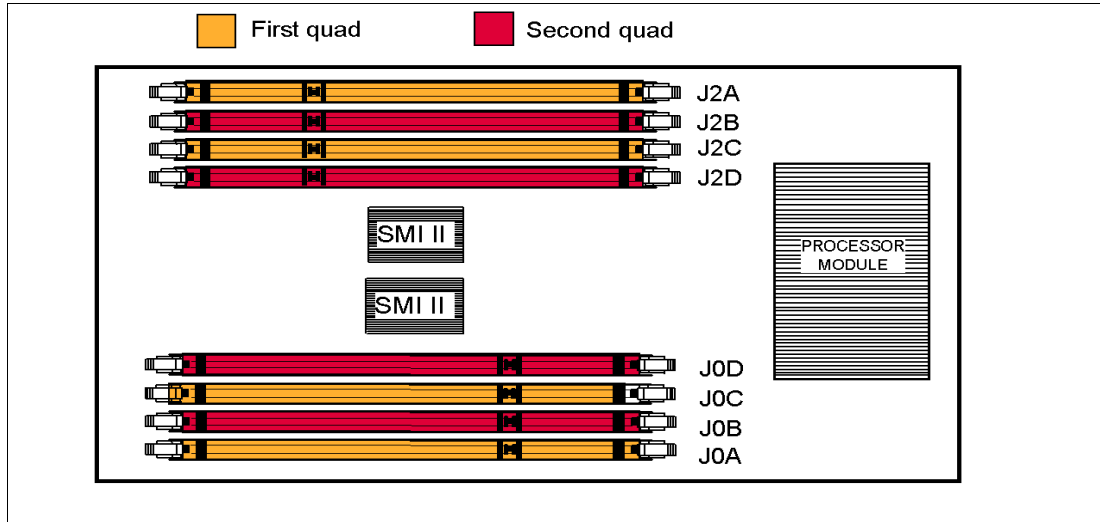
Figure 2-4   Memory placement for the p5-510 server

2.3.2 Memory restriction

OEM memory is not supported by IBM on the p5-510 server. OEM memory is not certified by IBM for use in pSeries and IBM @server p5 servers. If the p5-510 server is populated with OEM memory, you could experience unexpected and unpredictable behavior, especially when the system is using Micro-Partitioning.

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.
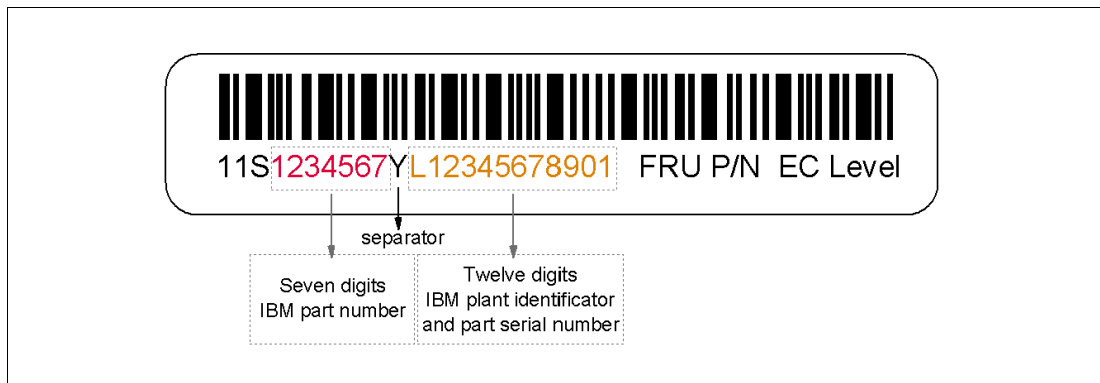
Figure 2-5   IBM memory certification label

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 for the internal I/O subsystem. 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 p5-510 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 through 3) can accept long PCI-X or PCI cards. They are all 64-bit, 3.3 volts, and 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 p5-510 servers can support existing 3.3 volt PCI adapters.

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

The PCI-X slots in the p5-510 server support Extended Error Handling (EEH). In the unlikely event of a problem, EEH-enabled adapters respond to a special data packet generated from the affected PCI-X slot hardware by calling system firmware, which is designed to examine the affected bus, allow the device driver to reset it, and continue without a system reboot.

> **Note:** As soon as a p5-510 server is connected to a Hardware Management Console, the POWER Hypervisor™ will prevent the system from using non-EEH OEM adapters. To use non-EEH OEM adapters, the system must run in full partition mode without connection to an HMC. A system previously attached to an HMC can be returned to a factory single partition state only when the service processor is reset to a factory state.

2.5.2 LAN adapters

When a p5-510 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 lists 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
4962	IBM 10/100 Mbps Ethernet PCI Adapter II	1, 2, 3	Short	3
5700	IBM Gigabit Ethernet-SX PCI-X Adapter	1, 2, 3	Short	3
5701	IBM 10/100/1000 Base-TX Ethernet PCI-X Adapter	1, 2, 3	Short	3
5706	IBM 2-Port 10/100/1000 Base-TX Ethernet PCI-X Adapter	1, 2, 3	Short	3
5707	IBM 2-Port Gigabit Ethernet-SX PCI-X Adapter	1, 2, 3	Short	3
5718	IBM 10 Gigabit Ethernet-SR PCI-X Adapter	1, 2, 3	Short	1
5719	IBM 10 Gigabit Ethernet-LR PCI-X Adapter	1, 2, 3	Short	1

2.5.3 SCSI adapters

The p5-510 server has one internal dual port Ultra320-capable SCSI adapter integrated on the system planar. Unlike the other IBM @server p5 models, the p5-510 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.

---

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

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

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

Table 2-3   Available SCSI adapters

Feature code	Adapter description	Slot priority	Size	Max
5703	PCI-X Dual Channel Ultra320 SCSI RAID Adapter	1, 2, 3	long	3
5712	PCI-X Dual Channel Ultra320 SCSI Adapter	1, 2, 3	short	3

2.5.4 Graphic accelerator

The p5-510 server supports one enhanced POWER GXT135P 2D graphic accelerators (FC 2849). The POWER GXT135P is a versatile, low-priced 2D graphics accelerator for IBM @server p5, pSeries, and OpenPower servers. 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 AIX 5L or Linux starts an X11-based graphical user interface (GUI).

2.5.5 Service processor communications ports

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

This virtual connection provides support on a non HMC-managed server to use the service processor ports. When no HMC is attached, the 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), or High Availability Cluster Multiprocessing (HACMP™) non-IP[2] heartbeat.

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

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

Table 2-4   Available serial adapter

Feature code	Adapter description	Slot priority	Size	Max
2943	8-Port Asynchronous Adapter EIA-232/RS-422, PCI bus	1, 2, 3	short	3
2944	128-Port Asynchronous Controller, PCI bus	1, 2, 3	short	3
5723	2-Port Asynchronous EIA-232 PCI Adapter	1, 2, 3	short	3

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 p5-510 server supports four bays that are

---

[2] IP stands for Internet protocol.

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 2.8.1, "Dynamic logical partitioning" on page 27 for dynamic LPAR considerations.

2.6.1 Internal media devices

The p5-510 server provides one slimline media bay for optional DVD drives. Table 2-5 lists available and supported internal media devices. Alternate methods of maintaining and servicing the system need to be available if the DVD-ROM or DVD-RAM is not ordered; an external Internet connection must be available 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
2640	IDE Slimline DVD-ROM Drive
5751	IBM 4.7 GB IDE Slimline DVD-RAM Drive

Note: If SUSE LINUX Enterprise Server 9 for POWER, or later, or Red Hat Enterprise Linux AS for POWER Version 3, or later, is being installed in the system, FC 2640 or FC 5751, or follow-ons, are required.

2.6.2 Internal hot-swappable SCSI disks

The p5-510 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 hot-swappable disk drives.

Table 2-6   Hot-swappable disk drive options

Feature code	Description
3277	36.4 GB 15,000 rpm Ultra320 SCSI hot-swappable disk drive
3274	73.4 GB 10,000 rpm Ultra320 SCSI hot-swappable disk drive
3278	73.4 GB 15,000 rpm Ultra320 SCSI hot-swappable disk drive
3275	146.8 GB 10,000 rpm Ultra320 SCSI hot-swappable disk drive
3578	300 GB 10,000 rpm Ultra320 SCSI hot-swappable disk drive

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 five second intervals). The disk drive placement priority is SCSI ID 8, 5, 4, and then 3. See Figure 1-2 on page 3 for 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, such as the AIX 5L Logical Volume Manager, when using disk hot-swap capabilities. For more information, see *Problem Solving and Troubleshooting in AIX 5L*, SG24-5496.

> **Note:** We recommend the following procedure be used after the disk has been deconfigured, 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.

After the SCSI disk hot-swap procedure, you can expect to find `SCSI_ERR10` logged in the AIX 5L error log, with the second word of the sense data equal to `0017`. It is generated from a SCSI bus reset issued by the SES to reset all processes when a drive is inserted, and it is not an issue.

Hot-swappable disks and Linux

Hot-swappable disk drives on POWER5 systems are supported with SUSE LINUX Enterprise Server 9 for POWER, or later, and Red Hat Enterprise Linux AS for POWER Version 3, or later.

2.7 External disk subsystem

The p5-510 server has internal hot-swappable drives. Internal disks are usually used for the AIX 5L rootvg and paging space. Specific client requirements can be satisfied with the several external disk possibilities that the p5-510 server supports.

> **Note:** External I/O Drawers 7311-D11 and 7311-D20 are not supported on the p5-510 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 5712) or one Ultra3 SCSI RAID adapter (FC 5703).

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 the following models: Model DS4100, DS4300, DS4400, and DS4500. The Model DS4100 is the smallest model, and scales up to 28 TB; Model DS4500 is the largest, and scales up to 32 TB of disk storage at the time of this writing. Model DS4300 provides up to 16 bootable partitions that are attached with the Gigabit Fibre Channel adapter (FC 5716). Model DS4400 provides up to 64 bootable

partitions. In most cases, both the TotalStorage DS family and the p5-510 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 p5-510 server is also connected by using a Gigabit Fibre Channel adapter (FC 5716).

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

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

2.8 Logical partition and virtualization

Dynamic logical partition (LPAR) and virtualization increase utilization of system resources. This section provides details and configuration specifications on this topic. The virtualization discussion includes virtualization enabling technologies that are standard on the system, such as the POWER Hypervisor, and optional ones, such as the Advanced POWER Virtualization feature.

2.8.1 Dynamic logical partitioning

Introduced with the POWER4 processor product line and the AIX 5L Version 5.1 operating system, logical partitioning (LPAR) became available. This technology offered the capability to divide a pSeries system into separate logical systems, allowing each LPAR to run an operating environment on dedicated attached devices, such as processors, memory, and I/O components.

Later, dynamic LPAR increased the flexibility, allowing selected system resources, such as processors, memory, and I/O components, to be added and deleted from dedicated partitions while they are executing. AIX 5L Version 5.2 with all the necessary enhancements to enable dynamic LPAR was introduced in 2002. The ability to reconfigure dynamic LPARs encourages system administrators to dynamically redefine all available system resources to reach the optimum capacity for each defined dynamic LPAR.

For information about minimum required hardware and software, refer to "Hardware and software guidelines for dynamic LPAR and virtualization" on page 32.

2.8.2 Virtualization

With the introduction of the POWER5 processor, partitioning technology moved from a dedicated resource allocation model to a virtualized shared resource model. This section briefly discusses the key components of virtualization on IBM @server p5 servers.

More information on virtualization can be found on the following Web site:

> http://www.ibm.com/servers/eserver/about/virtualization/

POWER Hypervisor

Combined with features designed into the POWER5 processor, the POWER Hypervisor delivers functions that enable other system technologies including Micro-Partitioning, virtualized processors, IEEE VLAN, compatible virtual switch, virtual SCSI adapters, and virtual consoles. The POWER Hypervisor is a component of system firmware that is always active, regardless of system configuration.

The POWER Hypervisor does the following:

► Provides an abstraction layer between the physical hardware resources and the logical partitions using them

► Enforces partition integrity by providing a security layer between logical partitions

► Controls the dispatch of virtual processors to physical processors

► Saves and restores all processor state information during logical processor context switch

► Controls hardware I/O interrupt management facilities for logical partitions

► Provides virtual LAN channels between physical partitions that help to reduce the need for physical Ethernet adapters for inter-partition communication

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 is available with the optional Advanced POWER Virtualization feature, which is described in "Advanced POWER Virtualization feature" on page 29.

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 a POWER5 system with either AIX 5L Version 5.3 or the appropriate level of Linux and an HMC to define the virtual Ethernet devices. Virtual Ethernet does not require the purchase of any additional features or software such as the Advanced POWER Virtualization feature.

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 detects 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 provides the 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 purchase of any additional features or software such as the Advanced POWER Virtualization feature.

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 communications ports" on page 24. The virtual TTY support discussed here is configured using the HMC.

> **Note:** The POWER5 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 amount of system memory required. Use the LPAR Validation Tool for calculating the Hypervisor memory requirements.
>
> In AIX 5L Version 5.3, the `lparstat` command using the `-h` and `-H` flags will display hypervisor statistical data. Using the `-h` flag adds summary hypervisor statistics to the default `lparstat` output.

2.8.3 Advanced POWER Virtualization feature

The Advanced POWER Virtualization feature is an optional additional cost feature. This feature enables the implementation of virtual partitions on IBM @server p5 servers. Each IBM @server p5 server model has a unique feature code for this technology. For the p5-510 server, select FC 7432 to order the Advanced Virtualization feature. See Table 2-8 on page 33 for operating system support for this feature.

The Advanced POWER Virtualization feature includes:

► Firmware enablement for Micro-Partitioning:

– Supports up to 254 partitions, 1/10th of processor granularity

► Key enablement:

– All processors will be enabled for Micro-Partitioning (number of processors on system equals the number of Advanced POWER Virtualization features ordered)

► Installation image for the Virtual I/O (VIO) Server software that supports:

– Ethernet adapter sharing

– Virtual SCSI Server

- – VIO software ships on a CD
- – Software supports AIX 5L and Linux
► Partition Load Manager (AIX 5L Version 5.3 only)
- – Automated CPU and memory reconfiguration
- – Real-time partition configuration and load statistics
- – Graphical user interface
- – Software ships on a CD

For more details on Advanced POWER Virtualization and virtualization in general, see the following Web site:

> http://www.ibm.com/servers/eserver/pseries/ondemand/ve/resources.html

Note: The Advanced POWER Virtualization feature (FC 7432) is not supported on AIX 5L for POWER Version 5.2, or previous versions.

Micro-Partitioning technology

The concept of Micro-Partitioning is to allow the resource definition of a partition to allocate fractions of processors to the partition. Micro-Partitioning is only available with POWER5 systems. From an operating system perspective, a virtual processor is indistinquishable from a physical processor, unless the operating system had been enhanced to be made aware of the difference. Physical processors are abstracted into virtual processors that are available to partitions.

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 be incremented fractionally by as little as 1/100th of the processor. The shared processor partitions are dispatched and time-sliced on the physical processors under control of the POWER Hypervisor. 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 they are available. Table 2-7 lists processor partitioning information related to the p5-510 server.

Table 2-7   Processor partitioning overview of the p5-510 server

Partitioning implementation	Model 510
Processors (maximum configuration)	2
Dedicated processor partitions (maximum configuration)	2
Shared processor partitions (maximum configuration)	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 significantly lower. Also see Table 2-8 on page 33 for operating systems supported with Micro-Partitioning.

Virtual I/O Server

The Virtual I/O Server is a special purpose partition that provides virtual I/O resources to client partitions. The Virtual I/O Server owns the real resources that are 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 every partition own a dedicated network adapter, disk adapter, and disk drive.

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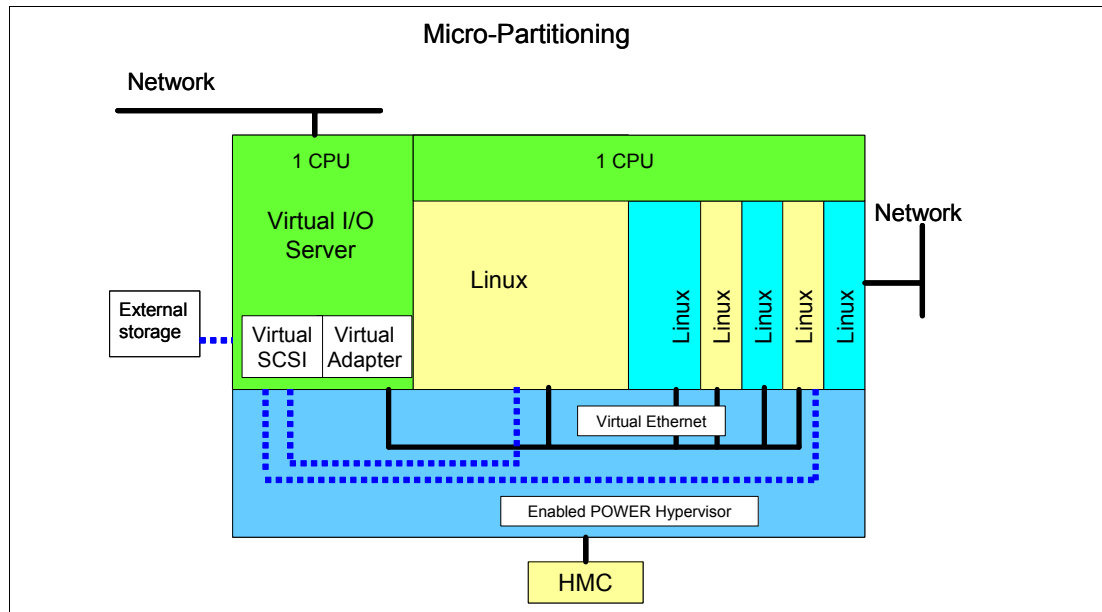

Figure 2-6   Virtual partition organization view

Since the Virtual I/O Server is an AIX 5L Version 5.3 operating system-based appliance, redundancy for physical devices attached to the Virtual I/O Server can be provided by using capabilities such as Multipath I/O and IEEE 802.3ad Link Aggregation.

Installation of the Virtual I/O Server partition is performed from a special mksysb CD that is provided to clients that order the Advanced POWER Virtualization feature. This is dedicated software only for the Virtual I/O Server operations, so the Virtual I/O Server software is only supported in Virtual I/O Server partitions.

The Virtual I/O Server can be installed by:

► Media (assigning the DVD-ROM drive to the partition and booting from the media)

► The HMC (inserting the media in the DVD-ROM drive on the HMC and using the `installios` command)

► Using Network Install Manager (NIM)

Note: To increase the performance of I/O intensive applications, dedicated physical adapters using dedicated partitions are preferred.

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 RAID or JBOD configurations are bootable, and the number of logical volumes is limited to the amount of storage available and architectural limits of the LVM.

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

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 with the Advanced POWER Virtualization feature services.

Section 2.8.4, "Hardware and software guidelines for dynamic LPAR and virtualization" provides detailed requirements for configuring virtual partitions.

Partition Load Manager

Partition Load Manager (PLM) provides automated processor and memory distribution between a dynamic LPAR and a Micro-Partitioning capable logical partition running AIX 5L. The PLM application is based on a client/server model to share system information, such as processor or memory events, across the concurrent present logical partitions.

The following events are registered on all managed partition nodes:

► Memory-pages-steal high thresholds and low thresholds

► Memory-usage high thresholds and low thresholds

► Processor-load-average high threshold and low threshold

Note: PLM is supported on AIX 5L Version 5.2 and AIX 5L Version 5.3.

2.8.4 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 lists AIX 5L and Linux support for dynamic LPAR and virtualization.

Table 2-8   Operating system supported function

Function	AIX 5L Version 5.2	AIX 5L Version 5.3	Linux SLES 9	Linux RHEL AS 3	Linux RHEL AS 4
Dynamic LPAR					
Processor	Y	Y	Y	N	Y
Memory	Y	Y	N	N	N
I/O	Y	Y	Y	N	Y
Virtualization					
Micro-partitions (1/10th of processor)	N	Y	Y	Y	Y
Virtual Storage	N	Y	Y	Y	Y
Virtual Ethernet	N	Y	Y	Y	Y
Partition Load Manager	Y	Y	N	N	N

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 to 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 p5-510 server is two without virtualization. With the use of the Advanced POWER Virtualization feature, the number of partitions per processor is 10.

Memory
It is important to highlight that the new IBM @server p5 and OpenPower servers and their associated virtualization features have adopted an even more dynamic memory allocation policy than the previous partition-capable pSeries servers.

In a partitioned environment, some of the physical memory areas are reserved by several system functions to enable partitioning in the partitioning-capable 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 eight percent of the system installed memory. The actual amount required will generally be less than eight 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 AIX 5L, 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 AIX 5L Version 5.2 with the 5200-04 Recommended Maintenance package, AIX 5L Version 5.3, Red Hat Enterprise Linux AS 3 (no dynamic LPAR), Red Hat Enterprise Linux AS 4, and SUSE LINUX Enterprise Server 9. There are planning considerations 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 estimate of the amount of memory required by the server firmware, use the LPAR Validation Tool (LVT). Refer to "LPAR validation tool" on page 37.

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.

The following devices represent groups when assigned to an LPAR:

► 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 not using virtualization, external disk space is necessary; this 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 partitions. 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 p5-510 server, additional disk space can be provided by using an external storage subsystem, such as the IBM 2104 Expandable Storage Plus, for LPARs using virtualization or direct attachment. For more detailed information of the available IBM disk subsystems, refer to "External disk subsystem" on page 26.

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

Therefore, for additional LPARs not 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 5706) to provide physically separated 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 (SEA) 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 IBM @server p5 systems offering dedicated devices under the covers against the cost of an external storage enclosure and the p5-510 server. Or use of the IBM Advanced POWER Virtualization feature.

2.8.5 Hardware Management Console

The HMC is a dedicated workstation that provides a graphical user interface for configuring and operating IBM @server p5 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: 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 p5-510 server using standard 10/100 Ethernet connection and uses HMC Port 1. A second HMC can be attached to HMC Port 2 of the p5-510 server for redundancy. For more details on 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 p5-510 server, the p5-510 server integrated service processor communications ports are disabled. An async adapter is needed to provide additional serial connections if required, for example, for non-IP HACMP.

POWER4 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** from the Hardware Management Console Support for IBM @server i5, @server p5, pSeries, and iSeries™ 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 on the 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	Customer Setup Unit	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.

2.8.6 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 @server i5 servers, iSeries, and OpenPower systems. A proficient knowledge of LPAR design requirements, limitations, and best practices, facilitates the use of this tool.

The LVT tool does the following:

► Provides support for partitions running AIX 5L Version 5.2 and Version 5.3, and Linux

► Validates dynamic LPAR design

► Validates 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 licenses 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 and determine memory resources required for all partitioned and non-partitioned servers.

Figure 2-7 on page 38 shows the calculated Hypervisor memory requirements based on sample partition requirements.
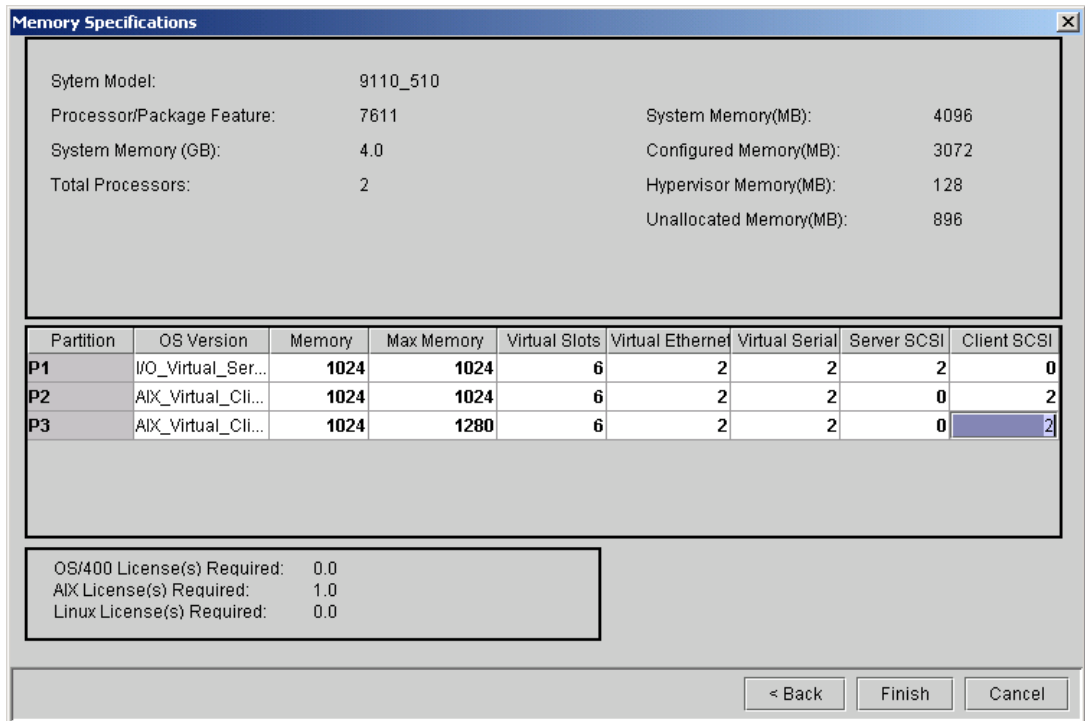
Figure 2-7   LVT screen showing Hypervisor requirements

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

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

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

2.8.7 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 a client's specifications. Manufacturing uses the LVT output to custom build the server. The server is then 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. Figure 2-8 on page 39 shows a sample screen of the eConfig wizard.
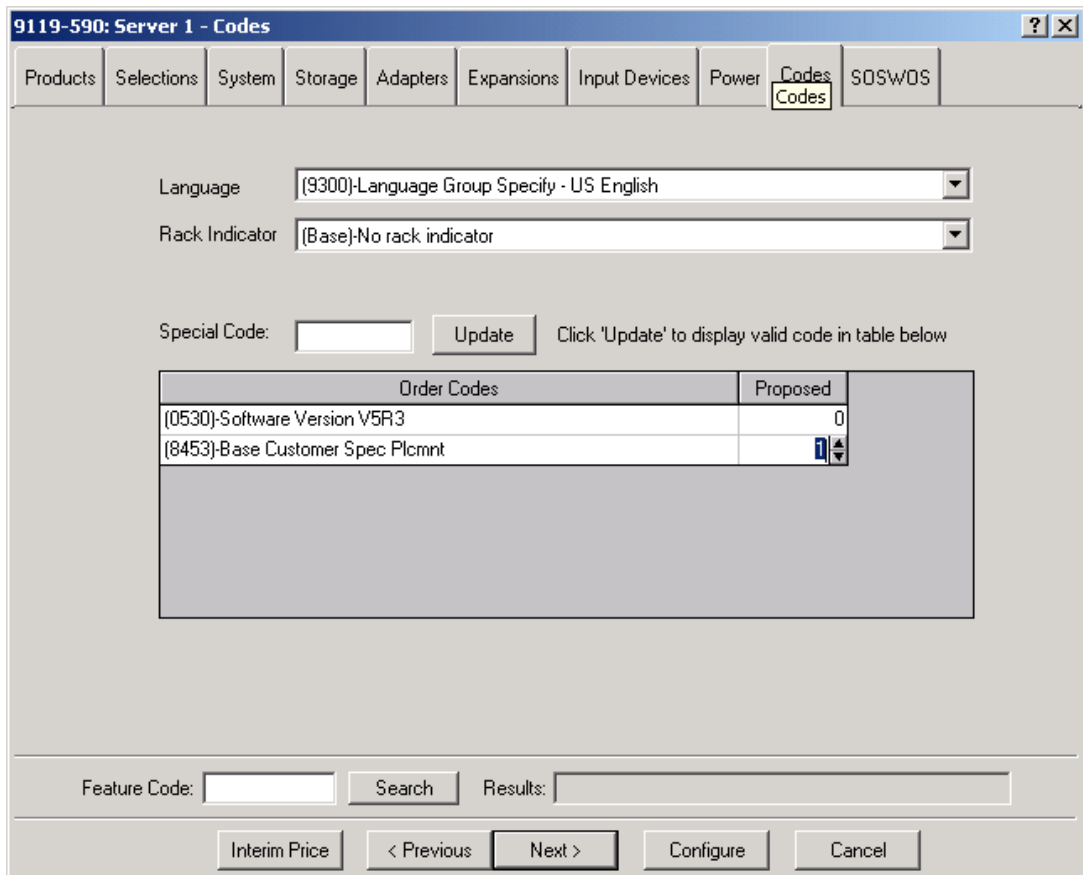
Figure 2-8   eConfig tab for CSP selection

> **Note:** An order submitted must include the CSP FC 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

2.9 Operating system support

The p5-510 server is capable of running IBM AIX 5L for POWER and supports appropriate versions of Linux. AIX 5L has been specifically developed and enhanced to exploit and support the extensive RAS features on IBM @server systems.

2.9.1 AIX 5L

If installing AIX 5L on the p5-510 server, the following minimum requirements must be met:

► AIX 5L for POWER Version 5.2 (5765-E62) with the 5200-05 Recommended Maintenance Package (APAR IY66260), dated 08/2004 (CD# LCD4-1133-04) or later.

► AIX 5L for POWER Version 5.3 (5765-G03) with the 5300-01 Recommended Maintenance Package (APAR IY62267), dated 08/2004 (CD# LCD4-7463-00) or later

> **Note:** The Advanced POWER Virtualization feature (FC 7432) is not supported on AIX 5L for POWER Version 5.2; it requires AIX 5L Version 5.3.

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

http://www.ibm.com/servers/eserver/support/pseries/aixfixes.html

Information on how to obtain the CD-ROM can be found on the Web page.

You can also get individual operating system fixes and information about obtaining AIX 5L service at this site. In AIX 5L Version 5.3, there is also the `suma` command available that helps the administrator to automate the task of checking and downloading operating system downloads. For more information about the suma command functionality refer to:

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

If you have problems downloading the latest maintenance level, ask your IBM Business Partner or IBM representative for assistance.

2.9.2 Linux

For the p5-510 server, Linux distributions are available through Novell SUSE and Red Hat at the time this publication was written. The p5-510 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:** Not all p5-510 server features available on the AIX 5L operating system are available on the Linux operating systems.

> **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 Model 510 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/

Information about Red Hat Enterprise Linux AS for pSeries from Red Hat can be found at:

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

For the latest in IBM Linux news, subscribe to the Linux Line; details are available at:

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 or later and Red Hat Enterprise Linux AS Version 3 or later (RHEL AS 3).

Cluster software SLES 9 options for clustered servers

The SLES 9 cluster subscription offerings provide lower priced SLES 9 license options for clients that 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 server and p5-510 Express, p5-520 and p5-520 Express, and OpenPower 710.

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

Cluster software RHEL AS 3 options for clustered servers

The RHEL AS Version 3 and later cluster subscription offerings provide lower priced 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 server and p5-510 Express, p5-520 and p5-520 Express, and OpenPower 710.

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

2.10 Service information

The p5-510 server is a customer setup system and is shipped with materials to assist in the general installation of the server. The system cover has a quick reference Service Information label that provides graphics that can aid in identifying features and location information. This section provides some additional service-related information.

2.10.1 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 p5-510 server.

Terra-cotta   The system can remain powered on while this task is being performed. Keep in mind that some tasks may require that other steps be performed first. One example would be deconfiguring a physical

volume in the operating system before removing the disk from the p5-510 server.

Blue and Terra-cotta Terra-cotta takes precedence over this color combination, 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.

2.10.2 Securing a system into a rack

The *optional* rack-mount drawer rail kit is a unique kit designed for use with the p5-510 server. No tools are required to install the p5-510 server or drawer rails into the rack.

The kit has a modular design that can be adapted to accommodate various rack depth 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 p5-510 server in the service position:

1. If necessary, open the front rack door.

2. Remove the two thumbscrews **A** that secure the system unit to the rack as shown in Figure 2-9.

3. Release the rack latches **B** on both the left and right sides as shown in the same figure.

4. 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.
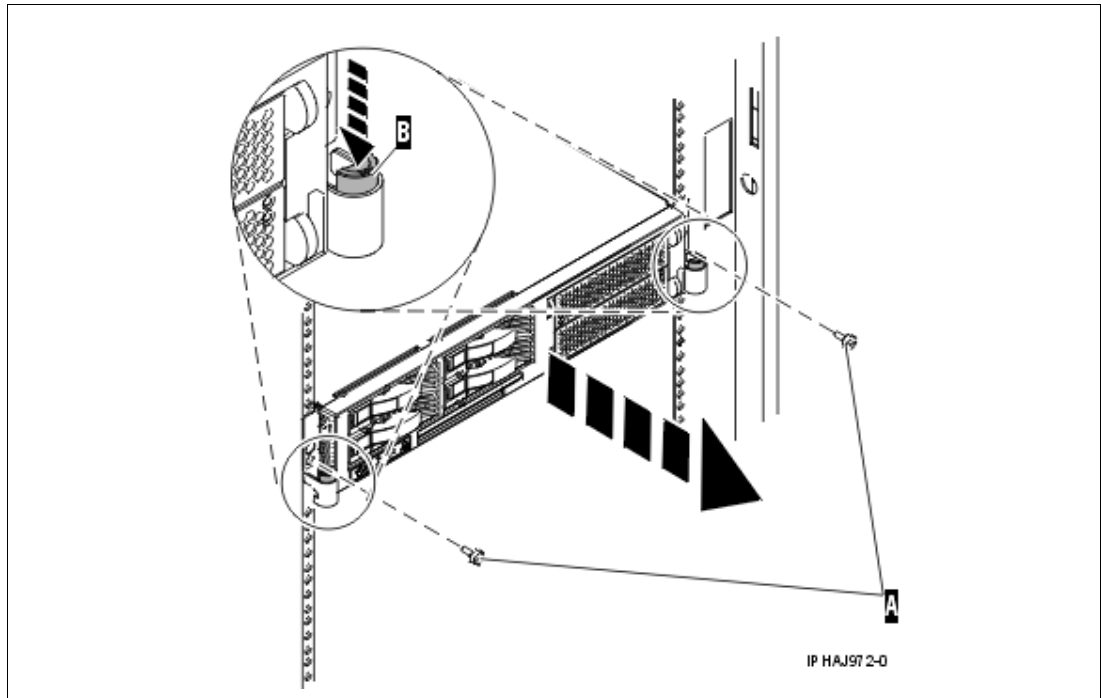
Figure 2-9   p5-510 server service position

For more information see the documentation site:

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

2.10.3 Operator control panel

The service processor provides an interface to the control panel, which is used to display system status and diagnostic information. The p5-510 server has a control panel packaged so that it fits into a smaller space. Under 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 p5-510 server control panel

To access all of the control panel's features, perform 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.

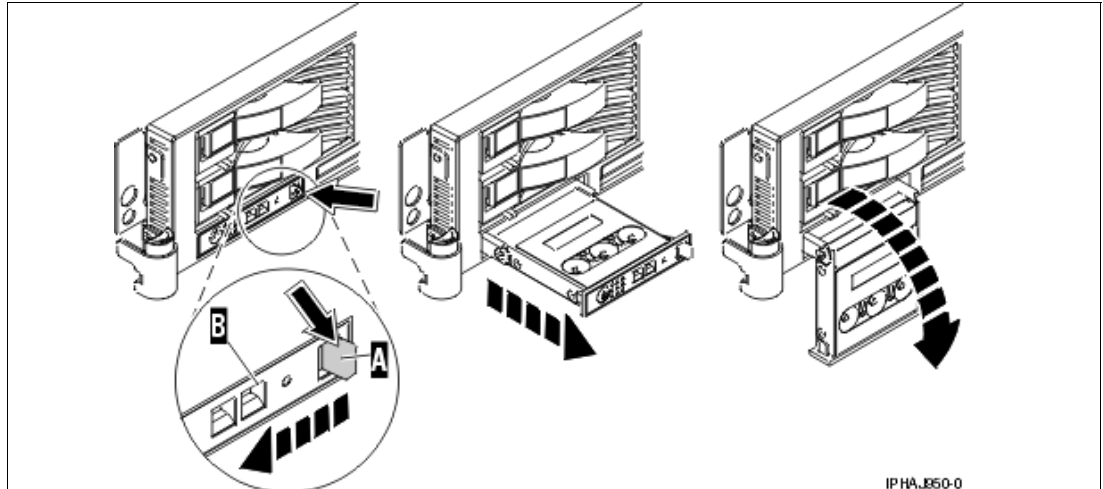► Pull the control panel out toward the front of the system until it can be pivoted downward on its hinge.

Figure 2-10   Accessing to the p5-510 server 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.
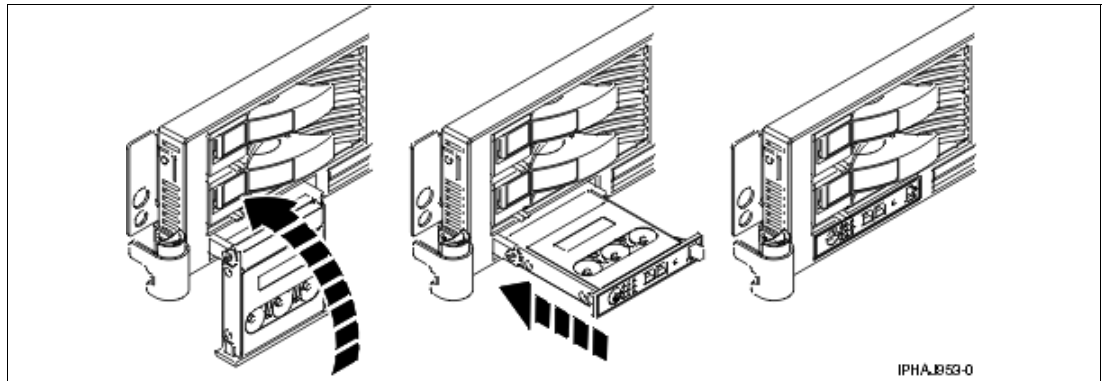
Figure 2-11   Restoring the p5-510 server control panel

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's 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 the functions mentioned are accessible either using the ASMI, the HMC, or the control panel.

For detailed information about each control panel function and the available values, select **Service Provider Information** → **Reference Information** → **Service functions** → **Control panel functions** from the IBM @server Hardware Information Center Web site at:

Changing system operation mode

In the very unlikely event that a p5-510 server cannot perform a power down of the CEC, use function 08 for a fast power off, if the system is not connected to an HMC.

> **Note:** Function 08 is enabled only when the system operating mode is 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:

- ► Use the Increment or Decrement buttons to scroll to function 02.
- ► 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)

- ► 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 _ _
  ```
- ► Press Enter to select the system operating mode.

  ```
  0 2 _ _ A _ _ M < _ _ _ _ _ _
  _ _ _ _ _ _ _ _ _ _ _ _ P _ _
  ```
- ► Use the Increment or Decrement buttons to scroll through the firmware IPL modes (optional).
- ► Press Enter to select the firmware IPL mode and exit function 02.

For more information about the control panel and functions, refer to **Service and support** → **Reference information** → **Service function** → **Control panel function** at the IBM *eserver* Hardware Information Center at:

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

2.10.4 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 sides 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, sometimes referred to as "b".

> **Note:** The default value, from which the system will boot, is temporary.

The following examples are the output of the **lsmcode** command for AIX 5L and Linux, showing the firmware levels as they are displayed in the outputs.

► AIX 5L:

```
The current permanent system firmware image is SF220_005.
The current temporary system firmware image is SF220_006.
The system is currently booted from the temporary image.
```

► Linux:

```
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 select **Service and support** → **Customer service and support** → **Getting fixes** → **Firmware (Licensed Internal Code) fixes** → **Concepts** → **Temporary and permanent side of the service processor** from the IBM *℮server* Hardware Information Center Web site at:

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

Get server firmware fixes using an HMC

Let's say 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 procedure to install the fix.

1. Ensure 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 and select **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**. The Web site address is:

> 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 regardless of whether your operating system is AIX 5L or Linux.

To do this, 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 the server firmware.

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:

► AIX 5L

You must have AIX 5L diagnostics installed on your server to perform this task. AIX 5L diagnostics are installed when you install the AIX 5L operating system on your server. However, it is possible to deselect the diagnostics. Therefore, you need to ensure that the online AIX 5L diagnostics are installed before proceeding with this task.

► Linux

– Platform Enablement Library – librtas-*nnnnn*.rpm

– Service Aids – ppc64-utils-*nnnnn*.rpm

– Hardware Inventory – lsvpd-*nnnnn*.rpm

  where *nnnnn* 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 and select **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**. The Web site address is:

2.10.5 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). The service processor is illustrated in Figure 2-12.
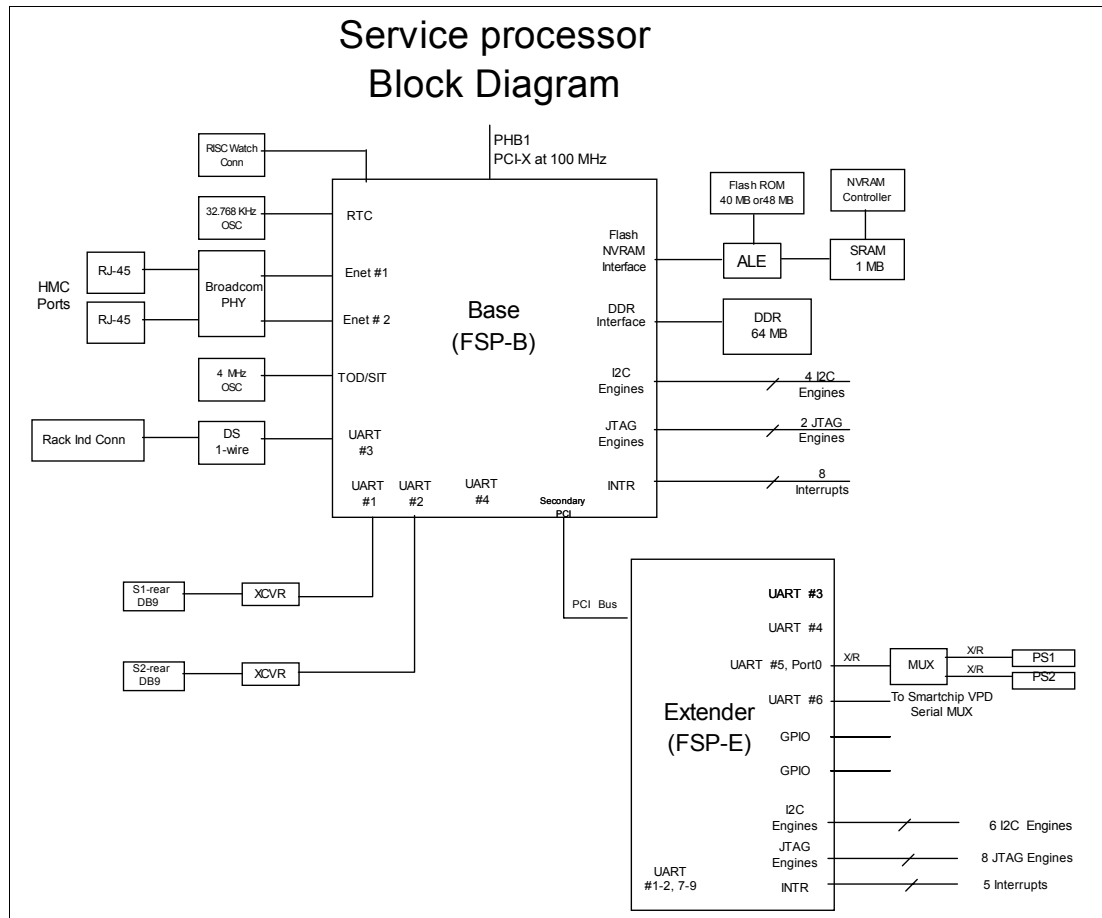
Figure 2-12   Service processor block diagram

The PPC405 core is a 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.

2.10.6 Hardware management user interfaces

In this section we provide a brief overview of the different p5-510 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 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 an 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 refer to the IBM @server Hardware Information Center at:

> 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:

- ► 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.

- ► For a partitioned server, watch the virtual terminal window on the HMC.

  For a full server partition, watch the firmware console.

- ► Look for the POST[4] indicators memory, keyboard, network, scsi, 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 refer to the IBM @server Hardware Information Center at.

> `http://publib.boulder.ibm.com/infocenter/eserver/v1r2s/en_US/index.htm?info/iphau/usings`
> `ms.htm`

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

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 remote HMC is 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 refer to the IBM @server Hardware Information Center.

---

[4] POST stands for Power-On-Self-Test
[5] DHCP stands for Dynamic Host Control Protocol

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 AIX 5L takes control. Partition firmware is needed only for the time that is necessary to load AIX 5L into the partition system memory.

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 or using the ASMI and selecting the boot to Open Firmware prompt.

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

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

3

Reliability, availability, and serviceability

This chapter provides detailed information about IBM ℮server p5 510 reliability, availability and serviceability (RAS) features. It describes several features that are available when using AIX 5L. Support of these features using Linux can vary.

3.1 Reliability, fault tolerance, and data integrity

The reliability of the p5-510 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.

Among the features that provide fault tolerance and ensure data integrity are the following:

- ▶ The p5-510 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.
- ▶ Disk mirroring and disk controller duplexing are also provided by the AIX 5L operating system. 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 PCI extended error handling

In the past, PCI bus parity errors caused a global machine check interrupt, which eventually required a system reboot to continue. In the POWER5 systems, new I/O drawer hardware, system firmware, and AIX 5L interaction have been designed to allow transparent recovery of intermittent PCI bus parity errors and graceful transition to the I/O device available state in the case of a permanent parity error in the PCI bus. This mechanism is called PCI extended error handling (EEH).

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

Note: This RAS function is not supported under Linux.

3.1.2 Memory error correction extensions

The p5-510 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.

3.1.3 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.

► 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.

3.1.4 Service processor

The service processor included in the p5-510 server is designed to provide 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 Chapter 2.10.5, "Service processor" on page 49.

3.1.5 Fault monitoring functions

Among the fault monitoring systems included with a p5-510 server are the following:

► Built-in self-test (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).

---

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

- Disk drive fault tracking can alert the system administrator of an impending disk failure before it impacts client operation.
- The AIX 5L or 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 IBM service representatives to bring along probable replacement hardware components when a service call is placed, thus minimizing system repair time.

3.1.6 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.

3.1.7 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 easily be accessed by the service processor.
- All communication between the SP and the FIR 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.

3.1.8 Environmental monitoring functions

Among the environmental monitoring functions available for the p5-510 server are the following:

- Temperature monitoring increases the fan speed rotation when ambient temperature is above the normal operating range.
- Temperature monitoring warns the system administrator of potential environmentally 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.

3.1.9 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 AIX 5L or Linux operating system.

► Error Log Analysis (ELA) 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.

3.1.10 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 p5-510 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 p5-510 server allows clients to replace service parts (Customer Replaceable Unit) if they want to. The p5-510 server has incorporated LEDs that will indicate the parts needing to be replaced.

► The p5-510 server allows support personnel to remotely log into a system to review error logs and perform remote maintenance. The p5-510 server 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, and Online Diagnostics.

► Online Diagnostics, when installed, are resident with AIX 5L on the disk or system. They can be booted in single-user mode (service mode), run in maintenance mode, or run concurrently (concurrent mode) with other applications. They have access to the AIX 5L Error Log and the AIX 5L Configuration Data.

  – Service mode allows checking of system devices and features.

  – Concurrent mode allows the normal system functions to continue while selected resources are being checked.

  – Maintenance mode allows checking of most system resources.

▶ 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

Service Agent is available at no additional charge. When installed on an IBM ⓔ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 report the results to IBM Service automatically

Note: Because the 9110-510 system has an optional DVD-ROM (FC 2640) and DVD-RAM (FC 5751), alternate methods for maintaining and servicing the system need to be available if the DVD-ROM or DVD-RAM is not ordered; an external Internet connection must be available to maintain or update system microcode to the latest required level.

3.2.2 Online customer support

Online customer support (OCS) for hardware problem reporting may be performed via remote login by IBM ⓔserver specialists. The Electronic Service Agent™ software can also be used for this capability.

AIX 5L support offerings will be under AIXSERV and Electronic Service Agent.

Note: This RAS function is not supported under Linux.

3.3 IBM ⓔserver Cluster 1600

Today's IT infrastructure requires that systems meet increasing demands, while offering the flexibility and manageability to rapidly develop and deploy new services. IBM clustering hardware and software provide the 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

IBM ⓔserver Cluster 1600 is a POWER processor-based AIX 5L and Linux cluster targeting scientific and technical computing, large-scale databases, and workload consolidation. IBM Cluster Systems Management (CSM) is designed to provide a robust, powerful, and

centralized way to manage a large number of POWER5 processor-based systems all from one single point-of-control. CSM can help lower the overall cost of IT ownership by helping to simplify the tasks of installing, operating, and maintaining clusters of servers. CSM can provide one consistent interface for managing both AIX 5L and Linux nodes (physical systems or logical partitions), with capabilities for remote parallel network install, remote hardware control, and distributed command execution.

Cluster Systems Management V1.4 for AIX 5L and Linux on POWER is supported on the p5-510 server. For hardware control, a Hardware Management Console (HMC) is required. Additionally, the p5-510 server is added to the hardware models supported with the pSeries cluster 1600 running CSM.

Information regarding the IBM @server Cluster 1600, HMC control, cluster building block servers, and cluster software available can be found at:

   http://www-1.ibm.com/servers/eserver/clusters/hardware/1600.ht

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

- ► *IBM @server pSeries 670 and pSeries 690 System Handbook*, SG24-7040
- ► *The Complete Partitioning Guide for IBM pSeries Servers*, SG24-7039
- ► *Managing AIX Server Farms*, SG24-6606
- ► *Practical Guide for SAN with pSeries*, SG24-6050
- ► *Problem Solving and Troubleshooting in AIX 5L*, SG24-5496
- ► *Understanding IBM @server pSeries Performance and Sizing*, SG24-4810
- ► *Advance POWER Virtualization on IBM @server p5 Servers,* SG24-7940
- ► *IBM @server p5 520 Technical Overview and Introduction*, REDP-9111-01

Other publications

These publications are also relevant as further information sources:

- ► *7014 Series Model T00 and T42 Rack Installation and Service Guide*, SA38-0577, contains information regarding the 7014 Model T00 and T42 Rack, in which this server can be installed.
- ► *7316-TF3 17-Inch Flat Panel Rack-Mounted Monitor and Keyboard Installation and Maintenance Guide*, SA38-0643, contains information regarding the 7316-TF3 Flat Panel Display, which can be installed in your rack to manage your system units.
- ► *RS/6000 and @server pSeries Adapters, Devices, and Cable Information for Multiple Bus Systems*, SA38-0516, contains information about adapters, devices, and cables for your system. This manual is intended to supplement the service information found in the *Diagnostic Information for Multiple Bus Systems* documentation.
- ► *RS/6000 and pSeries PCI Adapter Placement Reference*, SA38-0538, contains information regarding slot restrictions for adapters that can be used in this system.
- ► *System Unit Safety Information*, SA23-2652, contains translations of safety information used throughout the system documentation.
- ► *IBM @server Planning*, SA38-0508, contains site and planning information, including power and environment specification.

Online resources

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

► AIX 5L operating system maintenance packages downloads

  http://www.ibm.com/servers/eserver/support/pseries/aixfixes.html

► IBM @server p5, pSeries, OpenPower and IBM RS/6000 Performance Report

  http://www.ibm.com/servers/eserver/pseries/hardware/system_perf.html

► IBM TotalStorage Expandable Storage Plus

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

► IBM TotalStorage Mid-range Disk Systems

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

► IBM TotalStorage Enterprise disk storage

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

► IBM Virtualization Engine™

  http://www.ibm.com/servers/eserver/about/virtualization/

► Advanced POWER Virtualization on IBM @server p5

  http://www.ibm.com/servers/eserver/pseries/ondemand/ve/resources.html

► Virtual I/O Server supported environments

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

► Hardware Management Console support information

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

► The LVT is a PC based tool intended assist you in logical partitioning

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

► Customer Specified Placement and LPAR delivery

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

► SUMA on AIX 5L

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

► Linux on IBM @server p5 and pSeries

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

► SUSE LINUX Enterprise Server 9

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

► Red Hat Enterprise Linux details

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

► IBM @server Linux on POWER overview

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

► IBM @server Cluster 1600

  http://www.ibm.com/servers/eserver/clusters/hardware/1600.html

- Autonomic computing on IBM @server pSeries servers

  http://www.ibm.com/autonomic/index.shtml

- Copper circuitry

  http://www.ibm.com/chips/technology/technologies/copper/

- IBM @server p5 AIX 5L Support for Micro-Partitioning and Simultaneous Multi-threading White Paper

  http://www.ibm.com/servers/aix/whitepapers/aix_support.pdf

- Hardware documentation

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

- IBM @server Information Center

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

- IBM @server pSeries support

  http://www.ibm.com/servers/eserver/support/pseries/index.html

- IBM @server support: Tips for AIX administrators

  http://techsupport.services.ibm.com/server/aix.srchBroker

- IBM Linux news: Subscribe to the Linux Line

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

- IBM online sales manual

  http://www.ibmlink.ibm.com

- Linux for IBM @server pSeries

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

- Microcode Discovery Service

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

- POWER4 system micro architecture, comprehensively described in the IBM Journal of Research and Development, Vol 46 No.1 January 2002

  http://www.research.ibm.com/journal/rd46-1.html

- SCSI T10 Technical Committee

  http://www.t10.org

- Silicon-on-insulator (SOI) technology

  http://www.ibm.com/chips/technology/technologies/soi/

- Microcode Downloads for IBM @server i5, OpenPower, p5, pSeries, and RS/6000 Systems

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

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 *e*server p5 510 Technical Overview and Introduction

Redpaper

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

Support for versions of AIX 5L and Linux operating systems

Enterprise-class features for applications that require a robust environment

This document is a comprehensive guide covering the IBM *e*server p5 510 AIX and Linux OS servers. It introduces major hardware offerings and discusses their prominent functions.

Professionals wishing to acquire a better understanding of IBM *e*server p5 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 *e*server documentation by providing a desktop reference that offers a detailed technical description of the p5-510 server.

This publication does not replace the latest IBM *e*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.