IBM Power Systems LC921 and LC922

Technical Overview and Introduction

Ritesh Nohria

Gustavo Santos

Power Systems

Redpaper

International Technical Support Organization

IBM Power Systems LC921 and LC922: Technical Overview and Introduction

September 2018

REDP-5495-00

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

First Edition (September 2018)

This edition applies to the IBM Power Systems LC921 and LC922 servers, machine type and model numbers 9006-12P and 9006-22P.

Important: At time of publication, this book is based on a pre-GA version of a product. For the most up-to-date information regarding this product, consult the product documentation or subsequent updates of this book.

Contents

Notices

This information was developed for products and services offered in the US. This material might be available from IBM in other languages. However, you may be required to own a copy of the product or product version in that language in order to access it.

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

IBM may have patents or pending patent applications covering subject matter described in this document. The furnishing of this document does not grant you any license to these patents. You can send license inquiries, in writing, to:
IBM Director of Licensing, IBM Corporation, North Castle Drive, MD-NC119, Armonk, NY 10504-1785, US

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 jurisdictions do not allow disclaimer of express or implied warranties in certain transactions, therefore, this statement may not apply to you.

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

Any references in this information to non-IBM websites are provided for convenience only and do not in any manner serve as an endorsement of those websites. The materials at those websites are not part of the materials for this IBM product and use of those websites is at your own risk.

IBM may use or distribute any of the information you provide in any way it believes appropriate without incurring any obligation to you.

The performance data and client examples cited are presented for illustrative purposes only. Actual performance results may vary depending on specific configurations and operating conditions.

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.

Statements regarding IBM's future direction or intent are subject to change or withdrawal without notice, and represent goals and objectives only.

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 actual people or business enterprises is entirely coincidental.

COPYRIGHT LICENSE:

This information contains sample application programs in source language, which illustrate programming techniques on various operating platforms. You may copy, modify, and distribute these sample programs in any form without payment to IBM, for the purposes of developing, using, marketing or distributing application programs conforming to the application programming interface for the operating platform for which the sample programs are written. These examples have not been thoroughly tested under all conditions. IBM, therefore, cannot guarantee or imply reliability, serviceability, or function of these programs. The sample programs are provided "AS IS", without warranty of any kind. IBM shall not be liable for any damages arising out of your use of the sample programs.

Trademarks

IBM, the IBM logo, and ibm.com are trademarks or registered trademarks of International Business Machines Corporation, registered in many jurisdictions worldwide. Other product and service names might be trademarks of IBM or other companies. A current list of IBM trademarks is available on the web at "Copyright and trademark information" at http://www.ibm.com/legal/copytrade.shtml

The following terms are trademarks or registered trademarks of International Business Machines Corporation, and might also be trademarks or registered trademarks in other countries.

AIX®
BigFix®
GPFS™
IBM®
IBM Cloud™
IBM Spectrum™
IBM Spectrum Scale™
POWER®
Power Architecture®
Power Systems™
POWER9™
PowerHA®
PowerVM®
Redbooks®
Redpaper™
Redbooks (logo) ®
System Storage®
SystemMirror®

The following terms are trademarks of other companies:

Intel, Intel logo, Intel Inside logo, and Intel Centrino logo are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries.

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

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

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

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

Preface

This IBM® Redpaper™ publication is a comprehensive guide that covers the IBM Power Systems™ LC921 and LC922 (9006-12P and 9006-22P)) servers that use the current IBM POWER9™ processor-based technology and supports Linux operating systems (OSes). The objective of this paper is to introduce the offerings and their capacities and available features.

These new Linux scale-out systems provide differentiated performance, scalability, and low acquisition cost, and include the following features:

- Superior throughput and performance for high-value Linux workloads.
- Low acquisition cost through system optimization (industry-standard memory and industry-standard three-year warranty).
- Rich I/O options in the system unit. There are 12 large form factor (LFF)/small form factor (SFF) bays for 12 SAS/SATA hard disk drives (HDDs) or solid-state drives (SSDs), and four bays that are available for Non-Volatile Memory Express (NVMe) Gen3 adapters.
- Includes Trusted Platform Module (TPM) 2.0 Nuvoton NPCT650ABAWX through I2C (for secure boot and trusted boot).
- Integrated MicroSemi PM8069 SAS/SATA 16-port Internal Storage Controller Peripheral Component Interconnect Express (PCIe) 3.0 x8 with RAID 0, 1, 5, and 10 support (no write cache).
- Integrated Intel XL710 Quad Port 10 GBase-T PCIe 3.0 x8 UIO built-in local area network (LAN) (one shared management port).
- Dedicated 1 Gb Intelligent Platform Management Interface (IPMI) port.

This publication is for professionals who want to acquire a better understanding of IBM Power Systems products. The intended audience includes:

- Clients
- Sales and marketing professionals
- Technical support professionals
- IBM Business Partners
- Independent software vendors (ISVs)

Authors

This paper was produced by a team of specialists from around the world working at the International Technical Support Organization, Austin Center.

Ritesh Nohria is a Client Technical Specialist for Power Systems at IBM India. He has been with IBM since 2014. He has a total 23 years of experience in the IT field. He holds a degree in Electronics and Controls from the YMCA Institute of Engineering, Faridabad, India. He has served in various positions in his career and supported IBM mainframes, Digital VAX, Control Data Cyber mainframes, SGI Origin system, and Cray Vector systems, and worked with many Linux distributions. His areas of expertise include high-performance computing (HPC), modern data platforms, Internet of Things (IoT), and artificial intelligence (AI).

Gustavo Santos is an IBM Power Systems Consultant at IBM. He has been with IBM since 1997. He has 21 years of experience in Power Systems and UNIX. He holds a degree in Systems Engineering from Universidad Abierta Interamericana. He also has 15 years of experience working in service delivery on IBM AIX®, Virtual I/O Server (VIOS), and Hardware Management Console (HMC) for multiple accounts in the United States and Latin America. During the last 4 years, he worked as a Power Systems Consultant deploying services and training courses. His areas of expertise include Power Systems, AIX, Live Partition Mobility (LPM), IBM General Parallel File System (IBM GPFS™), UNIX, IBM Cloud™ IBM PowerVC Manager, IBM BigFix®, the IBM Spectrum™ family, and IBM PowerHA® SystemMirror®.

The project that produced this publication was managed by:

Scott Vetter, PMP
IBM Austin

Thanks to the following individuals for their contribution and support of this publication:

Ron Arroyo, Matthew Butterbaugh, Daniel Henderson, Jeanine Hinck, Ray Laning, Chris Mann, Benjamin Mashak, Stephen Mroz, Thoi Nguyen, Kanisha Patel, Justin Thaler, Julie Villarreal
IBM

Now you can become a published author, too!

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

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

ibm.com/redbooks/residencies.html

Comments welcome

Your comments are important to us!

We want our papers to be as helpful as possible. Send us your comments about this paper or other IBM Redbooks® publications in one of the following ways:

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

  ibm.com/redbooks

- Send your comments in an email to:

  redbooks@us.ibm.com

- Mail your comments to:

  IBM Corporation, International Technical Support Organization
  Dept. HYTD Mail Station P099
  2455 South Road
  Poughkeepsie, NY 12601-5400

Stay connected to IBM Redbooks

- Find us on Facebook:

  http://www.facebook.com/IBMRedbooks

- Follow us on Twitter:

  http://twitter.com/ibmredbooks

- Look for us on LinkedIn:

  http://www.linkedin.com/groups?home=&gid=2130806

- Explore new Redbooks publications, residencies, and workshops with the IBM Redbooks weekly newsletter:

  https://www.redbooks.ibm.com/Redbooks.nsf/subscribe?OpenForm

- Stay current on recent Redbooks publications with RSS Feeds:

  http://www.redbooks.ibm.com/rss.html

1

Architectural and technical description

Today, the number of sources generating data is leading to an exponential growth in the data volume. Making sense of this data, and doing it faster than the competition can, leads to an unprecedented opportunity to gain valuable insights and apply them at the best point of impact to improve your business results.

The Power LC921 server is the dense 1U server design that delivers Peripheral Component Interconnect Express (PCIe) Gen4 and Coherent Accelerator Processor Interface (CAPI) 2.0 for superior scaling efficiency.

The Power System LC922 server delivers superior performance and meets the open source workload data requirements that are being driven by the AI era or modern data platforms (for example, MariaDB, MongoDB, PHP, Hortonworks, PostgreSQL, and NoSQL). It is the server of choice for clients that want compute capabilities that are enabled by IBM POWER9 processors, scaling efficiency with the POWER9 advanced I/O, and ample storage capacity to contain the AI data volumes. Additionally, the server is simple to order and cost-optimized, and can scale from single racks to large clusters with a simple deployment.

1.1 Power LC921 hardware overview

The Power LC921 server delivers superior performance, scaling efficiency, and density that meets the requirements that are demanded by the AI era. It is the server of choice for clients that want industry-leading computation with up to two POWER9 processors in a dense 1U server design that delivers PCIe Gen4 and CAPI 2.0 for superior scaling efficiency.

The Power LC921 server, a dense server design, is shown in Figure 1-1.

Figure 1-1 IBM Power LC921 1U system

The Power LC921 server supports one or two processor sockets, offering 16-core (2.2 GHz base frequency) or 20-core (2.13 GHz base frequency) POWER9 processor-based configurations in a 19-inch rack-mount, 1U (EIA units) drawer configuration. No mixing of processors is supported.

Base processor frequencies are typically exceeded under workloads. The maximum frequency boost is 3.8 GHz, and achievable under most workload and environmental conditions.

Each chip supports 512 KB L2 cache per core and 120 MB L3 cache per chip.

The Power LC921 server provides two hot-swap and redundant power supplies and two, four, six, eight, 12, or 16 RDIMM memory slots. Supported memory features are 8 GB (#EKMA), 16 GB (#EKMF), 32 GB (#EKMG), 64 GB (#EKMD), and 128 GB (#EKME). No mixing of memory is allowed. The server supports a maximum 2 TB memory.

The Power LC921 server also offers:

- Superior throughput and performance for high-value Linux workloads.
- Low acquisition cost through system optimization (industry-standard memory and 3-year warranty).
- Powerful I/O options in the system unit, including:
  - Two PCIe G4 x16 full-height and full-length slots, CAPI 2.0-enabled. One of them is available for a full-height and full-length GPU.
  - One PCIe G4 x8 low-profile (LP) slot.
  - One PCIe G4 x8 LP slot, CAPI-2.0 enabled (internal).
- Four 3.5” large form factor (LFF)/2.5” small form factor (SFF) bays for four SAS/SATA hard disk drives (HDDs) or solid-state drives (SSDs) or 2.5” four bays available for Non-Volatile Memory Express (NVMe) Gen3 adapters.

- Two SATA Disk on Module (DOM) connectors on the system board (for optional local, captive flash devices for operating system (OS) boot).
- Integrated MicroSemi PM8069 SAS/SATA 16-port Internal Storage Controller PCIe3.0 x8 with RAID 0, 1, 5, and 10 support (no write cache):
  - The Integrated SAS/SATA controller (PM8069) does not support 4 K disks (SSD or HDD). Selecting a 4 K HDD requires an LSI host bus adapter (HBA). SSDs must be formatted by the client to 512-byte sectors for RAID.
  - The Integrated SAS/SATA controller (PM8069) does not have a self-encrypting drive (SED) feature, but can use unlocked SED drives.
- Integrated Intel XL710 Quad Port 10 GBase-T PCIe3.0 x8 UIO built-in local area network (LAN) (one shared management port).
- Includes Trusted Platform Module (TPM) 2.0 Nuvoton NPCT650ABAWX through a I2C (for secure boot and trusted boot).
- Dedicated 1 Gb Intelligent Platform Management Interface (IPMI) port.
- Two rear USB 3.0 ports.
- 19-inch rack-mount 1U configuration.
- OSes:
  - Red Hat Enterprise Linux (RHEL) 7.5 Little Endian (POWER9), or later.
  - Ubuntu Server 18.04 LTS.

1.2 Power LC922 hardware overview

The Power LC922 server supports one or two processor sockets, offering 16-core (2.91 GHz base frequency) or 20-core (2.70 GHz base frequency) or 22-core (2.6 GHz base frequency) POWER9 processor-based configurations in a 19-inch rack-mount.

Base processor frequencies are typically exceeded under workloads. The maximum frequency boost is 3.8 GHz, and achievable under most workload and environmental conditions.

It has a 2U (EIA units) drawer configuration. No mixing of processors is allowed.

The Power LC922 server provides two hot-swap and redundant power supplies, and two, four, six, eight, 12, or 16 DIMM memory slots. Supported memory features are 8 GB (#EKMA), 16 GB (#EKMF), 32 GB (#EKMG), 64 GB (#EKMD), and 128 GB (#EKME). No mixing of memory is allowed.

The Power LC922 server is shown in Figure 1-2.

Figure 1-2 IBM Power LC922 server 2U form factor

The Power LC922 server also offers:

- Superior throughput and performance for high-value Linux workloads.
- Low acquisition cost through system optimization (industry-standard 3-year warranty).
- Powerful IBM POWER9 single-chip module (SCM) processors that offer 2.9 GHz, 2.7 GHz, or 2.6 GHz performance with 16, 20, or 22 single-socket and 32, 40, or 44 fully activated cores. No mixing of processors is allowed.
- Up to 2048 GB (2 TB) of DDR4 memory.
- Rich I/O options in the system unit, including:
  - Two PCIe G4 x16 full-height and full-length slots, CAPI-2.0 enabled.
  - Three PCIe G4 x8 full-height and full-length slots, one CAPI-2.0 enabled (all physically x16).
  - One PCIe G4 x8 LP slot.
- Two full-height and full-length GPU-capable slots, but more GPUs can be added to LP slots.
- Twelve LFF/SFF bays for 12 SAS/SATA HDDs or SSDs, with four that are available for NVMe Gen3 adapters.
- Includes TPM 2.0 Nuvoton NPCT650ABAWX through I2C (for secure boot and trusted boot).
- Integrated MicroSemi PM8069 SAS/SATA 16-port Internal Storage Controller PCIe3.0 x8 with RAID 0, 1, 5, and 10 support (no write cache):
  - The Integrated SAS/SATA controller (PM8069) does not support 4 K disks (SSD or HDD). Selecting a 4 K HDD requires an LSI HBA. SSDs must be formatted by the client to 512-byte sectors for RAID.
  - The Integrated SAS/SATA controller (PM8069) does not have a SED feature, but can use unlocked SED drives.
- Two SATA DOM connectors on the system board (for optional local, captive flash devices for OS boot).
- Optional two SFF drive trays in the rear for SAS/SATA (physically blocks off 16 trays for the first CPU when installed). They are controlled only by the integrated MicroSemi PM8069.

- Integrated Intel XL710 Quad Port 10 GBase-T PCIe3.0 x8 UIO built-in LAN (one shared management port).
- Dedicated 1 Gb IPMI port.
- Two rear USB 3.0 ports.
- Rear VGA port.
- 19-inch rack-mount 2U configuration.
- OSes:
  - RHEL 7.5 Little Endian (POWER9) or later.
  - Ubuntu Server 18.04 LTS.

1.2.1 Minimum features

The Power LC921 server has by default an NVMe system board, and the Power LC922 serve has two drive backplane options, both of which are NVMe-capable. The default drive backplane directly connects SAS/SATA to the integrated MicroSemi PM8069, and the optional drive backplane includes an LSI expander (SAS3x28R) to allow connection of all 12 SAS/SATA disks to one of several LSI HBA options. The system selection determines whether the system accepts one or two processors along with the backplane.

The minimum Power LC921 or Power LC922 initial order must include one processor module, two memory DIMMs of 32 GB capacity, two power cords, rack-mounting hardware, a system software indicator, a rack integrator specify, and a Language Group Specify. Linux is the OS.

Table 1-1 shows the minimum configuration for the Power LC921 server.

Table 1-1 Power LC921 minimum configuration

Feature code	Description
#EKP6 or #EKP7	One 16-core 2.2 GHz POWER9 processor One 20-core 2.13 GHz POWER9 processor
#EKMA	Two 16 GB DIMMs, 2666 MHz, and 4 GB DDR4 DRAM
#4650	One Rack Indicator - Not Factory Integrated
#EKBB	One 1-Socket 1U LFF NVMe Fab Assembly
#2147	One Primary OS Linux
#9xxx	Language Group Specify (Select one from the announced features.)

Table 1-2 shows the minimum configuration for the Power LC922 server.

Table 1-2 IBM system LC922 minimum configuration

Feature code	Description
#EKPD or #EKPC or #EKPE	One 16-core 2.9 GHz POWER9 processor One 20-core 2.7 GHz POWER9 processor One 22-core 2.6 GHz POWER9 processor
#EKMA	Two 16 GB DIMMs, 2666 MHz, and 4 GB DDR4 DRAM

Feature code	Description
#4650	One Rack Indicator - Not Factory Integrated
#EKBF	ONe 1-Socket 2U 12x LFF/SFF 4 NVMe Direct Attach Fab Assembly
#2147	One Primary OS Linux
#9xxx	Language Group Specify (Select one from the announced features.)

Note: If a rack is needed, it must be ordered as a machine type and model (MTM) rack with initial system orders. If the rack is included on the same system order, it included in the same shipment, but in separate packing material. IBM does not offer IBM Manufacturing rack integration of the server into the rack before shipping currently.

1.3 Operating system support

The Power LC921 and Power LC922 servers support Linux, which provides a UNIX like implementation across many computer architectures.

For more information about the software that is available on Power Systems, see Linux on Power Systems.

1.3.1 Ubuntu

Ubuntu Server 18.04.5 LTS for POWER9 is supported on the servers with support for later distributions as they become available.

For more information about Ubuntu Server for Ubuntu for POWER9, see Ubuntu Server for IBM POWER.

1.3.2 Red Hat Enterprise Linux

RHEL (ppc64le) Version 7.5 pp64le is supported on the server with support for later distributions as they become available.

For more questions about this release and supported Power Systems servers, see the Red Hat Hardware Catalog.

1.4 Operating environment

The Power LC921 and Power LC922 servers can operate at nominal processor frequencies within the ASHRAE A2 envelope.

Note: ASHRAE is a global society that provides standards for environmental measures. For more information, see ASHRAE.

Figure 1-3 shows the values and different levels of ASHRAE for humidity.

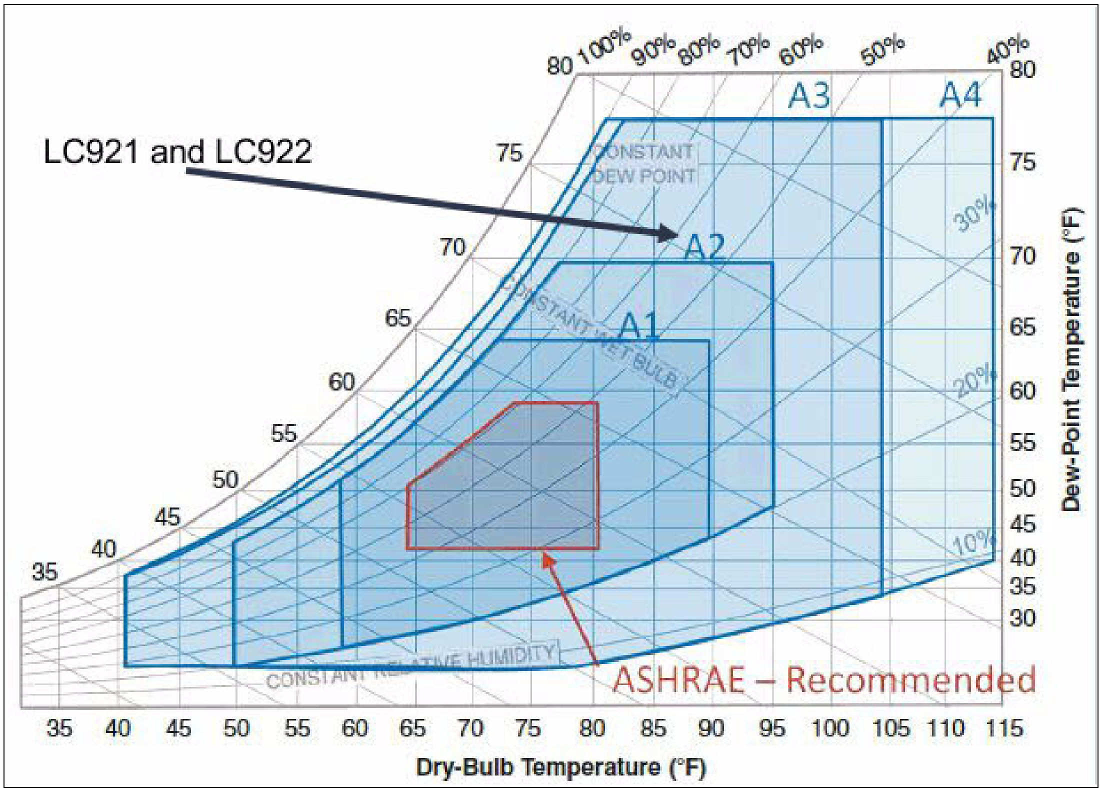


Figure 1-3 ASHRAE levels for humidity

Figure 1-4 shows levels for altitude and temperature.

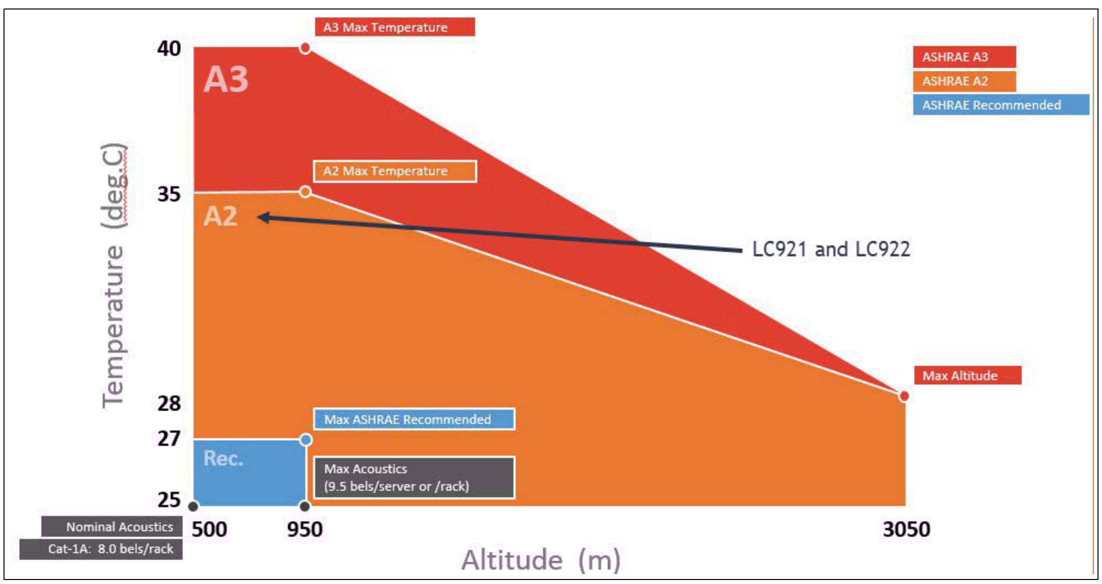


Figure 1-4 ASHRAE levels for temperature and altitude

Considerations: The presence of GPUs reduces the overall number of allowed drives in the front to 8x drives, all of which must be plugged in the bottom two rows due to thermal constraints:

- Power LC922 server with no GPU: The system can support up to twelve 3.5-inch drives in the front and two rear drives (optional). The ambient temperature must not go above 35 °C (95 °F).
- Power LC922 server with GPUs: The system can support up to eight 3.5-inch drives. They must be plugged into the bottom two rows and cannot be plugged into the top row. The ambient temperature must not go above 25 °C (77 °F).
  - In standard base systems (#EKBF and #EKBG), up to eight drives may be populated with GPUs, but only two NVMe drives are allowed.
  - In base systems with the LSI expander drive backplane (#EKBH and #EKBJ), up to eight drives may be populated with the GPUs, but only two NVMe drives are allowed.

For more information about ASHRAE A2, see ASHRAE Standards and Guidelines.

1.5 Physical package

The Power LC921 server is offered exclusively as a rack-mounted 1U server. The width, depth, height, and weight of the server are shown in Table 1-3.

Table 1-3 Physical dimensions for the Power LC921 server

Dimension	Power LC921 server
Width	441.5 mm (17.4 inches)
Depth	822 mm (32.4 inches)
Height	43 mm (1.7 inches)
Weight (maximum configuration)	13.38 kg (29.5 lbs)

The Power LC922 server is offered exclusively as a rack-mounted 2U server. The width, depth, height, and weight of the server are shown in Table 1-4.

Table 1-4 Physical dimensions for the Power LC922 server

Dimension	Power LC922 server
Width	441.5 mm (17.4 inches)
Depth	822 mm (32.4 inches)
Height	89 mm (3.5 inches)
Weight (maximum configuration)	25 kg (56 lbs)

Note: The weight of these servers does not include the weight of the HDDs.

1.6 System architecture

These servers balance processor performance, storage capacity, memory capacity, memory bandwidth, and PCIe adapter allowance to maximize price performance for big data workloads.

Figure 1-5 illustrates the block diagram of the Power LC921 server.

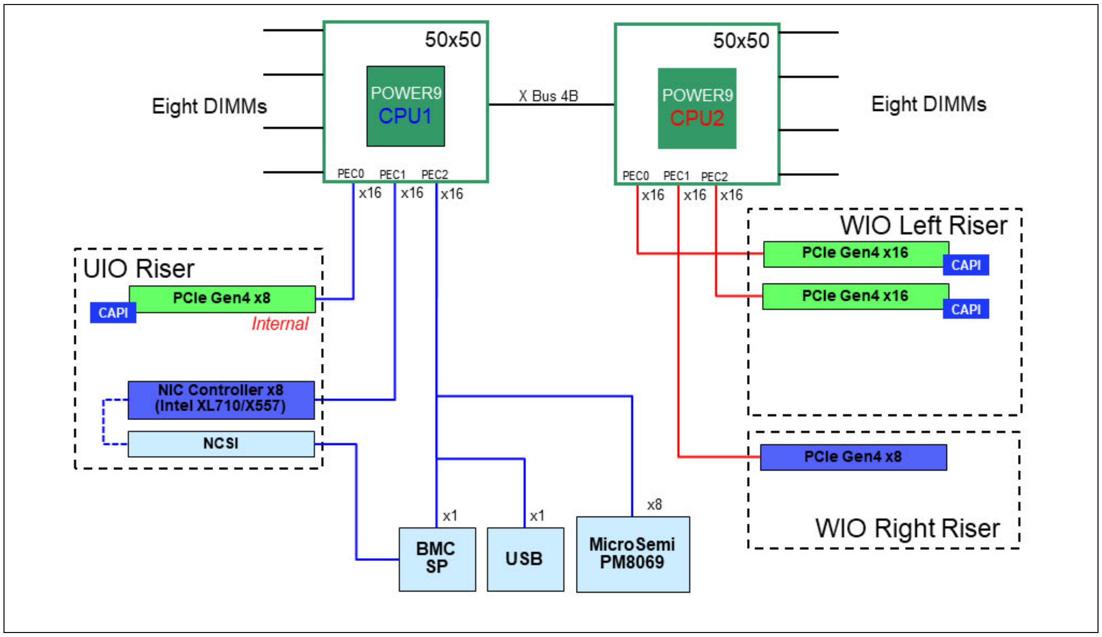


Figure 1-5 Power LC921 server logical system diagram

Figure 1-6 illustrates the block diagram of Power LC922 server.

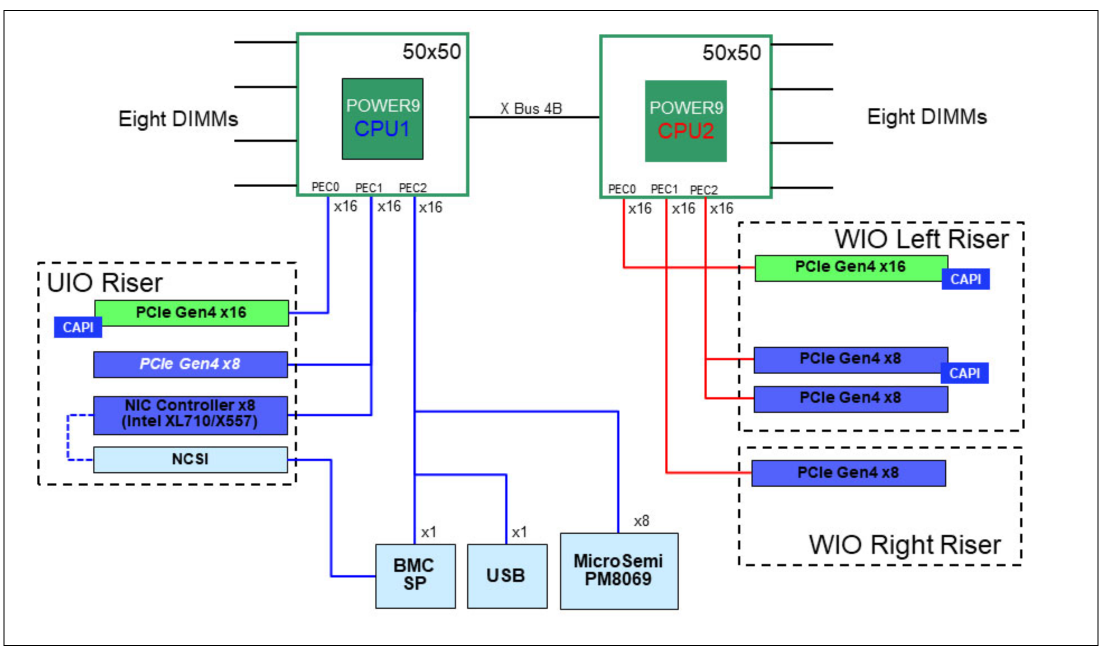


Figure 1-6 Power LC922 server logical block diagram

Bandwidths that are provided throughout the section are theoretical maximums that are used for reference.

The speeds that are shown are at an individual component level. Multiple components and application implementation are key to achieving the preferred performance. Always do the performance sizing at the application workload environment level and evaluate performance by using real-world performance measurements and production workloads.

Table 1-5 shows the feature codes (FCs) for the Power LC921 system board.

Table 1-5 Power System LC921 system board

Feature code	Description
#EKBB	1-Socket 1U LFF NVMe Fab Assembly
#EKBC	2-Socket 1U LFF NVMe Fab Assembly Base

Table 1-6 shows the FCs for the Power LC921 processors.

Table 1-6 Processor availability in the Power LC921 server

Feature code	Description
#EKP6	16-core 2.2 GHz POWER9 processor
#EKP7	20-core 2.13 GHz POWER9 processor

The number of processors in the system is determined by the base system that is selected.

Table 1-7 shows the FCs for the Power LC922 system board.

Table 1-7 Power System LC922 system board

Feature code	Description
#EKBF	1-Socket 2U 12 LFF/SFF 4 NVMe Direct Attach Fab Assembly
#EKBG	2-Socket 2U 12 LFF/SFF 4 NVMe Direct Attach Fab Assembly
#EKBH	1-Socket 2U 12 LFF/SFF 4 NVMe LSI Expander Assembly (SAS3x28R)
#EKBJ	2-Socket 2U 12 LFF/SFF 4 NVMe LSI Expander Assembly (SAS3x28R)

Table 1-8 shows the FCs for the Power LC922 processors.

Table 1-8 Processor availability in the Power System LC922 server

Feature code	Description
#EKPD	16-core 2.9 GHz POWER9 processor
#EKPC	20-core 2.7 GHz POWER9 processor
#EKPE	22-core 2.6 GHz POWER9 processor

The number of processors in the system is determined by the base system that is selected.

Note: A 1-socket board cannot be upgraded to a 2-socket board.

The Power System LC922 server is composed of a base system with an infrastructure that is determined by the number of needed processors and support for NVMe drives. The base system selection determines whether the system accepts one or two processors. The base system selection also determines the use of a default drive midplane that supports SAS and SATA drives. All base systems support s NVMe drives in four of the available slots.

Figure 1-7 shows the system board overview and location of CPUs, DIMM, SATA DOM, and TPM.

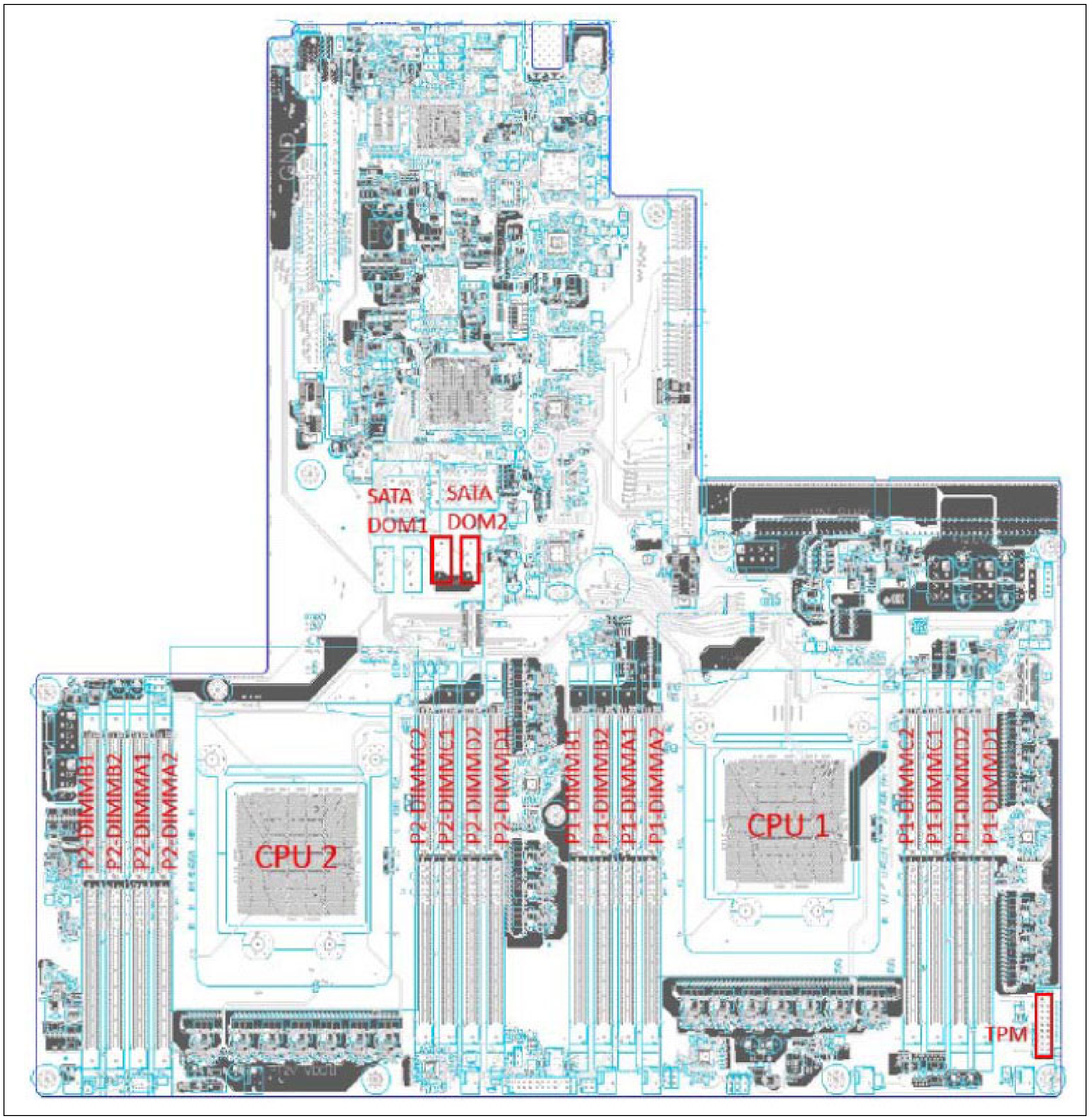


Figure 1-7   System board overview with CPU, DIMM, SATA DOM, and TPM locations

1.7 The POWER9 processor

This section introduces the current processor in the Power Systems product family and describes its main characteristics and features in general.

The POWER9 processor in the Power LC921 and LC922 servers is the most current generation of the IBM POWER® processor family.

1.7.1 POWER9 processor overview

The IBM POWER9 processor is a super-scalar symmetric multiprocessor (SMP) that is designed for use in servers and large-cluster systems. It supports a maximum SMP size of two sockets and is targeted for high CPU-consuming workloads.

The POWER9 processor offers superior cost and performance benefits. The target market segments are:

- Technical computing

  The POWER9 processor provides superior floating-point performance and high-memory bandwidth to address this market segment. It also supports off-chip floating-point acceleration.

- Cloud operating environments

  The POWER9 processor enables efficient cloud management software, enforces service-level agreements (SLAs), and provides facilities for chargeback accounting that is based on the resources that are consumed.

- Big data analytics

  The POWER9 processor with CAPI-attached large caches and on-chip accelerators provides a robust platform for analytic and big data applications.

- High-performance computing (HPC), high-performance data analysis (HPDA), and artificial intelligence (AI)

  The POWER9 processor can be connected to NVIDIA GPUs by using PCIe interfaces to increase processing capacities to fit the requirements of HPC, HPDA, and AI.

1.7.2 Processor features

The POWER9 processor is an SCM that is based on complementary metal–oxide–semiconductor (CMOS) 14-nm technology with 17 metal layers. It is optimized for cloud and data center applications. Within a 50 × 50 mm footprint, designs have four direct-attached memory channels for scale-out configurations. Each DDR4 channel supports up to 2666 Mbps for one DIMM per channel or 2133 Mbps for two DIMMs per channel. Two processors are tightly coupled through two 4 byte, 16 Gbps elastic differential interfaces (EDIs) for SMP. There are 48 lanes of PCIe Gen4 adapters at 16 Gbps.

The POWER9 processor consists of the following main components:

- Twenty-four POWER9 cores, which include both L1 instruction and data caches, shared L2 and L3 caches, and a non-cacheable unit (NCU).
- Each core has up to four threads that use simultaneous multithreading (SMT).
- On-chip accelerators.

- CAPI 2.0 allows a Field Programmable Gate Array (FPGA) or Application-Specific Integrated Circuit (ASIC) to connect coherently to the POWER9 processor-based SMP interconnect through the PCIe bus.
- On-chip: Compression, encryption, and data moves are initiated by the hypervisor, compression engine, or a Nest Memory Management Unit (NMMU) to enable user access to all accelerators.
- In-core: User invocation encryption (Advanced Encryption Standard (AES) and Secure Hash Algorithm (SHA)).

- Two memory controllers that support direct-attached DDR4 memory:
  - Supports four direct-attach memory buses.
  - Supports ×4 and ×8, 4 - 16 Gb DRAMs and 3D stacked DRAMs.
- Processor SMP interconnect:
  - Supports one inter-node SMP X-bus link.
  - Maximum two-socket SMP.
- Three PCIe Controllers (PECs) with 48 lanes of PCIe Gen4 I/O:
  - PEC0: One x16 lanes.
  - PEC1: Two x8 lanes (bifurcation).
  - PEC2: One x16 lane mode or two x8 lanes (bifurcation), or one x8 lanes and two x4 lanes (trifurcation).
  - PEC0 and PEC2 support CAPI 2.0.
- Power management.
- Pervasive interface.

From a logical perspective, the POWER9 processor consists of four main components:

- SMP interconnect (also known as an internal fabric interconnect).
- Memory subsystem.
- PCIe I/O subsystem.
- Accelerator subsystem.

Figure 1-8 shows a POWER9 processor with 24 cores.

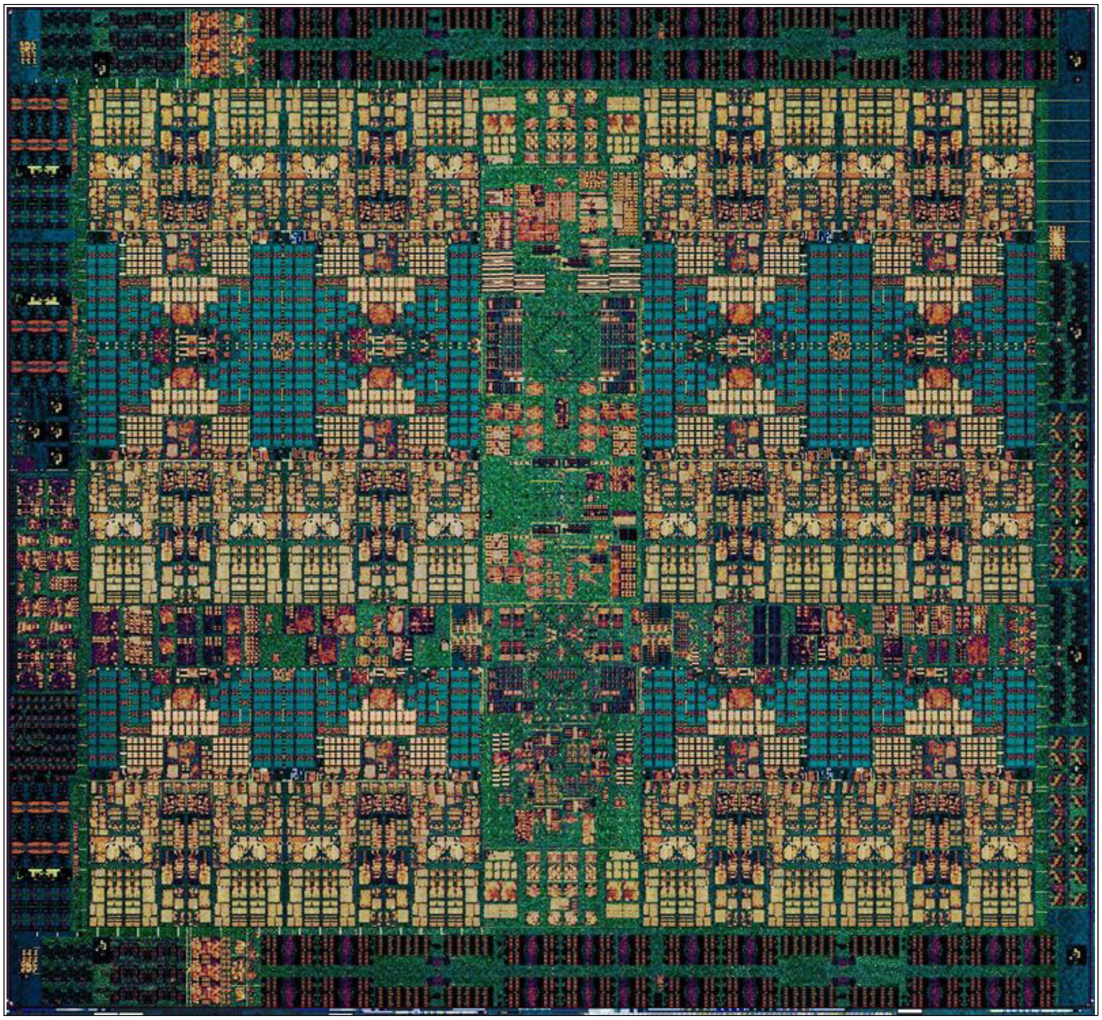


Figure 1-8 POWER9 processor with 24 cores

1.7.3 Supported technologies

The POWER9 processor supports the following technologies:

- IBM Power Instruction Set Architecture (ISA) Book I, II, and III (Version 3.0)
- Linux on IBM Power Architecture® Platform Reference
- IEEE P754-2008 for binary and decimal floating-point compliant
- Big Endian, little Endian, and strong-ordering support extension
- A 51-bit real address and a 68-bit virtual address

1.7.4 Simultaneous multithreading

The POWER9 processor has improvements in multi-core and multi-thread scaling. A significant performance opportunity comes from parallelizing workloads to enable the full potential of the microprocessor, and the large memory bandwidth. Application scaling is influenced by both multi-core and multi-thread technology.

SMT allows a single physical processor core to dispatch simultaneously instructions from more than one hardware thread context. With SMT, each POWER9 core can present four hardware threads. Because there are multiple hardware threads per physical processor core, more instructions can run at the same time.

SMT is primarily beneficial in commercial environments where the speed of an individual transaction is not as critical as the total number of transactions that are performed. SMT typically increases the throughput of workloads with large or frequently changing working sets, such as database servers and web servers.

Table 1-9 shows a comparison between the different POWER9 processors options for the Power LC921 and Power LC922 servers and the number of threads that are supported by each SMT mode.

Table 1-9 Number of threads that are supported by Power LC921 and Power LC922 servers

Cores per system	SMT 0	SMT 2	SMT 4
16	16 HTpS[a]	32 HTpS	64 HTpS
20	20 HTpS	40 HTpS	80 HTpS
22	22 HTpS	44 HTpS	88 HTpS
32	32 HTpS	64 HTpS	128 HTpS
40	40 HTpS	80 HTpS	160 HTpS
44	44 HTpS	88 HTpS	176 HTpS

a. Hardware Threads per System.

1.7.5 Processor feature codes

The Power LC921and Power LC922 servers support two processor configurations only, as shown in Table 1-10.

Table 1-10 Power LC921 supported processor features codes

Feature code	Description	Minimum	Maximum	OS support
#EKP6	16-core 2.2 GHz POWER9 processor	1	2	Linux
#EKP7	20-core 2.13 GHz POWER9 processor	1	2	Linux

A minimum of one processor module is required with up to a 16-core processor, or a 20-core processor and an extra option of a 22-core processor for Power LC922 servers. No processor activation features are used or orderable on the Power LC921 and Power LC922 servers. All processor cores are always fully activated.

Note: The mixing of different processor features on the same system is not allowed.

Table 1-11 shows the available processor FCs for the Power LC922 server.

Table 1-11 Power LC922 supported processor features codes

Feature code	Description	Minimum	Maximum	OS support
#EKPD	16-core 2.9 GHz POWER9 processor	1	2	Linux
#EKPC	20-core 2.7 GHz POWER9 processor	1	2	Linux
#EKPE	22-core 2.6 GHz POWER9 processor	1	2	Linux

1.7.6 Coherent Accelerator Processor Interface

CAPI connects a custom acceleration engine to the coherent fabric of the POWER9 chip. The hybrid solution has a simple programming paradigm while delivering performance well beyond today's I/O-attached acceleration engines.

Figure 1-9 shows the difference of non-CAPI and CAPI technology and flow of data.

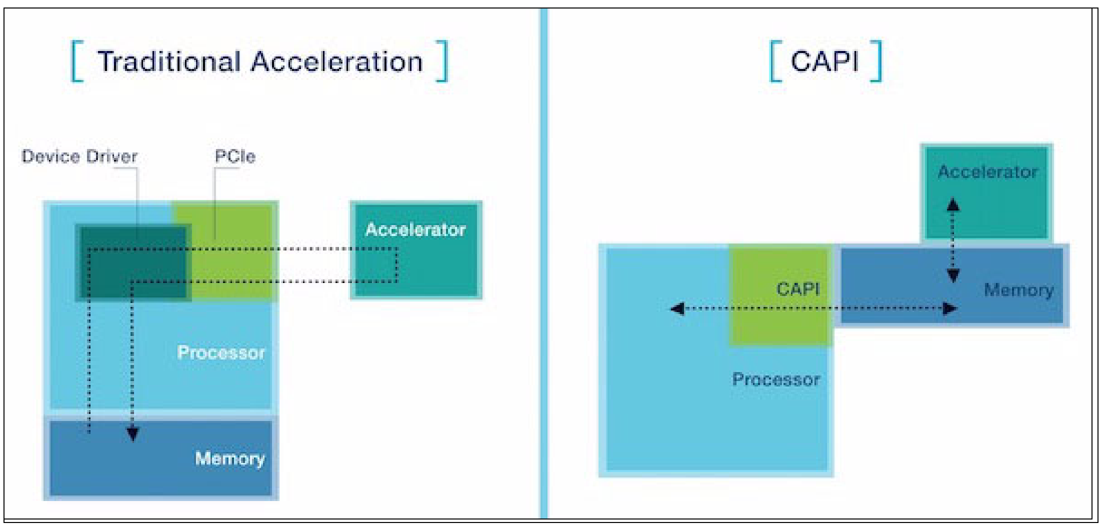


Figure 1-9 Overview of non-CAPI and CAPI technology

CAPI on POWER9 processor-based systems provides a high-performance solution for the implementation of software-specific, computation-heavy algorithms on an FPGA. This innovation can replace either application programs running on core or custom acceleration implementations that are attached through I/O. CAPI removes the complexity of the I/O subsystem, allowing an accelerator to operate as an extension of an application.

The IBM solution enables higher system performance with a much smaller programming investment, allowing heterogeneous computing to be successful across a much broader range of applications.

Because the FPGAs are reconfigured, hardware can be specialized without the traditional costs of hardware fabrication. CAPI enables a customer-defined FPGA solution to be a peer to the POWER9 cores from a memory access (coherent) standpoint.

For more information about CAPI, see CAPI on IBM Power Systems.

1.8 Memory subsystem

The POWER9 memory subsystem consists of a POWER9 processor that supports four channels. Each memory port or channel supports single- or dual-drop industry-standard RDIMMs. The POWER9 processor connects to the DIMMs by using the standard DDR4 memory interface.

Figure 1-10 shows the location of P1's DIMMs and CPU1 and P2's DIMMs and CPU2. The DIMM information is the same for both the Power LC921 and Power LC922 servers.

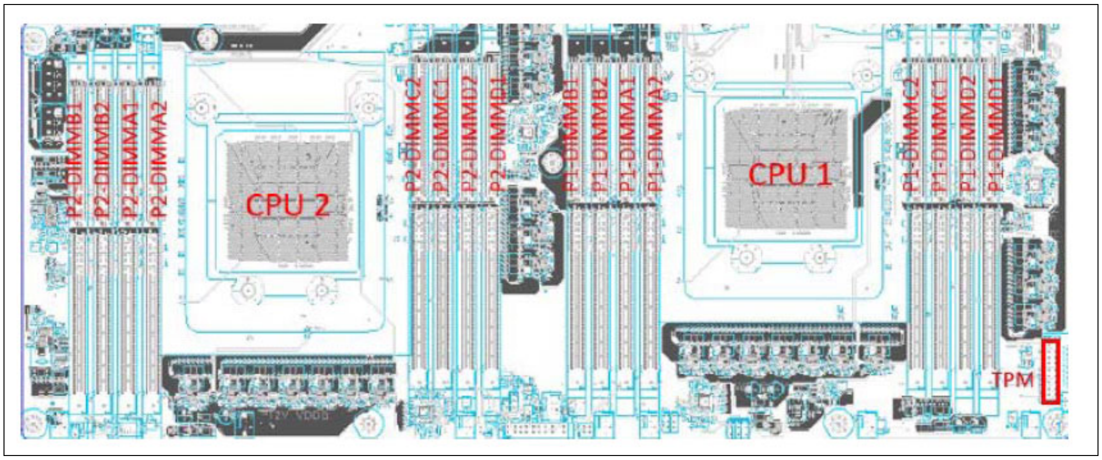


Figure 1-10 DIMM locations on Power LC921 and Power LC922 servers

Table 1-12 shows the supported DIMMs for the Power LC921 and Power LC922 servers.

Table 1-12 Memory feature codes for Power LC921 and Power LC922 servers

Feature code	Description
#EKMA	8 GB 1RX4 2666 MHz DDR4 RDIMM
#EKMF	16 GB 1RX4 2666 MHz DDR4 RDIMM
#EKMG	32 GB 2RX4 2666 MHz DDR4 RDIMM
#EKMD	64 GB 4RX4 2666 MHz DDR4 3DS RDIMM
#EKME	128 GB 8RX4 2666 MHz DDR4 3DS RDIMM

Note: Mixing of DIMMs is not allowed.

The POWER9 processor has four channels, and each channel can accommodate a maximum of two DIMMs.

Figure 1-11 shows a single DIMM per channel configuration.

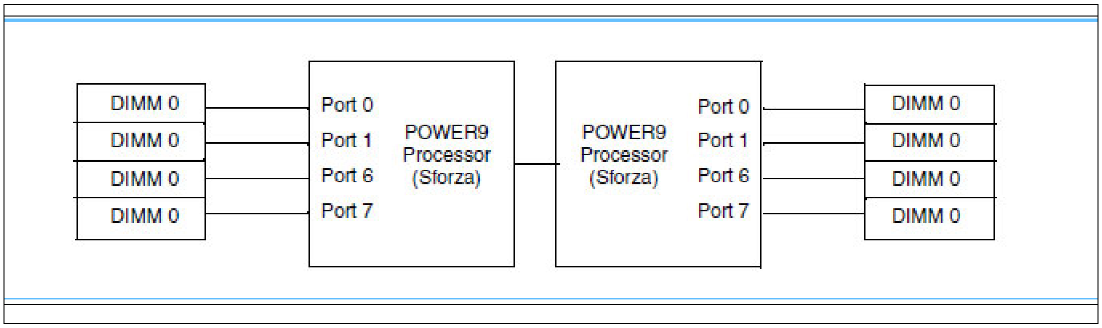


Figure 1-11 Single DIMM per channel configuration

Figure 1-12 shows a dual DIMM per channel configuration.

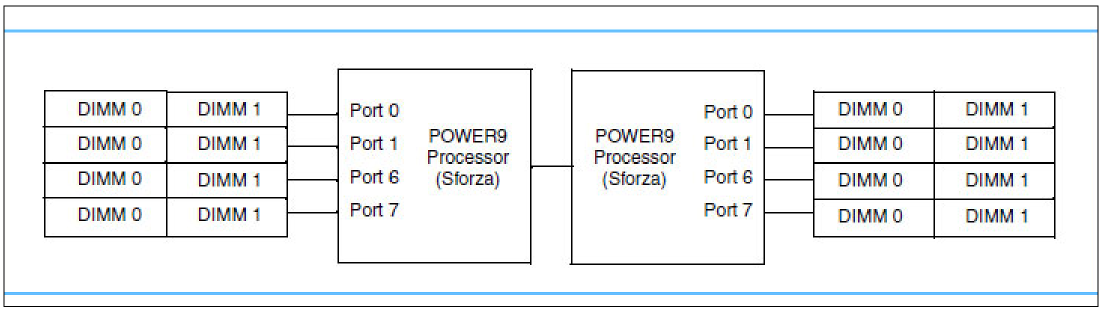


Figure 1-12 Dual DIMM per channel configuration

1.8.1 DIMM placement rules

Table 1-13 describes the basic rules of DIMMs in the Power LC921 and Power LC922 servers. The minimum memory is 32 GB and maximum is 2 TB. Follow these guidelines:

- For each DDR4 port, the far end, slot 1 must be populated first.
- DIMMs should be plugged into homogeneous groups within port pairs (A and B or C and D).

Table 1-13 DIMM placement rules

Slot location	Slot	DIMM quantity	Sequence	Comment
P1-A and P1-B	1	2	1B[a]	Minimum DIMMs that are required for both 1-socket or 2-socket servers.
P2-A and P2-B	1	2	2A[b]	
P1-C and P1-D	1	2	3	Maximum memory bandwidth for a 1-socket server
P2-C and P2-D	1	2	4A[b]	Maximum memory bandwidth for a 2-socket server

Slot location	Slot	DIMM quantity	Sequence	Comment
P1-A, P1-B, P1-C, and P1-D	2	4	5	P1 memory bandwidth reduction due to DDR4 frequency drop
P2-A, P2-B, P2-C, and P2-D	2	4	6A[b]	P2 memory BW reduction due to DDR4 frequency drop

a. B: Not allowed with an 8 GB DIMM because it does not meet the minimum memory requirement of 32 GB.
b. A: The P2 DIMMs location is the second socket, so P2-A, B, C, and D are not applicable for a 1-socket system.

Figure 1-13 shows the locations of the DIMMs.

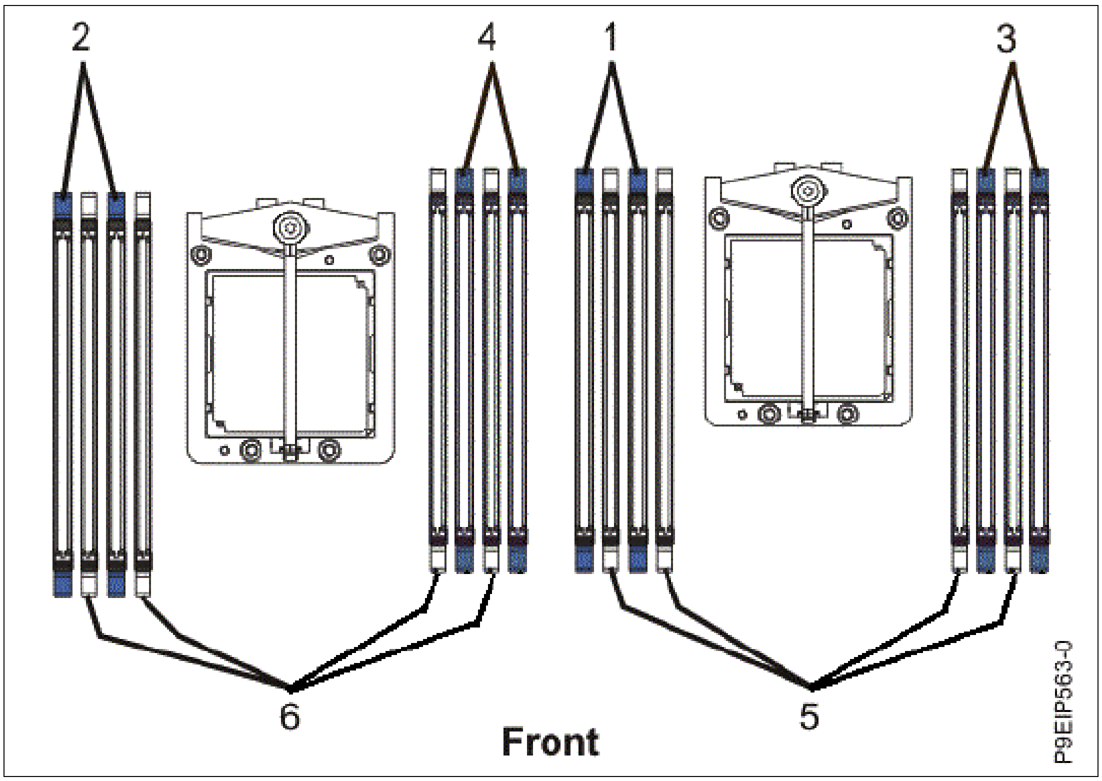


Figure 1-13 Physical mapping of the DIMM-to-plug sequence

1.8.2 Memory bandwidth

The memory bandwidth is 170 GBps peak per system in a 2-socket system with 8x RDIMMs single drop, running at 2667 Mbps (136 GBps peak memory bandwidth with all 16x RDIMMs populated, running at 2133 Mbps).

1.9 Internal I/O subsystem

The key components of the I/O subsystem are described in this section.

1.9.1 PCIe Express Controller and PCI Express

The PEC acts as a bridge between the internal processor bus and the high-speed serial (HSS) links that drive the PCI Express I/O. The PEC acts as a processor bus master on behalf of the PCI Express port, converting inbound memory read and write packets into processor bus direct memory access (DMA) traffic. The PEC also acts as a processor bus subordinate, transferring processor load and store commands to the PCI Express devices that are attached to the port.

PCIe uses a serial interface and enables point-to-point interconnections between devices by using a directly wired interface between these connection points. A single PCIe serial link is a dual-simplex connection that uses two pairs of wires, one pair for transmit and one pair for receive, and can transmit only 1 bit per cycle. These two pairs of wires are called a *lane*. A PCIe link can consist of multiple lanes. In these configurations, the connection is labeled as x1, x2, x8, x12, x16, or x32, where the number is effectively the number of lanes.

The Power LC922 and Power LC921 servers support the new PCIe Gen4 adapters, which are capable of 32 GBps simplex (64 GBps duplex) on a single x16 interface. PCIe Gen4 slots also support previous generation (Gen3 and Gen2) adapters, which operate at lower speeds according to the following rules:

- Place x1, x4, x8, and x16 speed adapters in the same size connector slots first before mixing adapter speed with connector slot size.
- Adapters with smaller link widths are allowed in larger-sized PCIe connectors, but larger width adapters are not compatible with smaller connector sizes (that is, a x16 adapter cannot go in an x8 PCIe slot connector).

> **Note:** PCIe x8 adapters use a different type of slot than PCI x16 adapters. If you attempt to force an adapter into the wrong type of slot, you might damage the adapter or the slot.
>
> All but 1x PCIe connector in the Power LC922 server is physically x8; the rest are physically x16, which can accept a x16 card, but have half the bandwidth if the slot is only electrically x8.

PCI Gen4 has double the bandwidth from previous generation PCI Gen3 adapters. The x16 PCI Gen4 adapter can handle 256 Gbps data and is ready for a 200 Gbps network connection if such a network becomes available.

Power LC922 or Power LC921 servers can support three different form factors of PCIe adapters:

- PCIe LP cards
- PCIe double-width, full-height and length cards
- PCIe full-height and full-length cards

Before adding or rearranging adapters, use the System Planning Tool to validate the new adapter configuration.

If you are installing a new feature, ensure that you have the software that is required to support the new feature and determine whether there are existing update prerequisites to install. To obtain this information, see Power Systems Prerequisites.

Each POWER9 processor has 48 PCIe lanes running at 32 Gbps full-duplex. The bandwidth formula for server is calculated as follows:

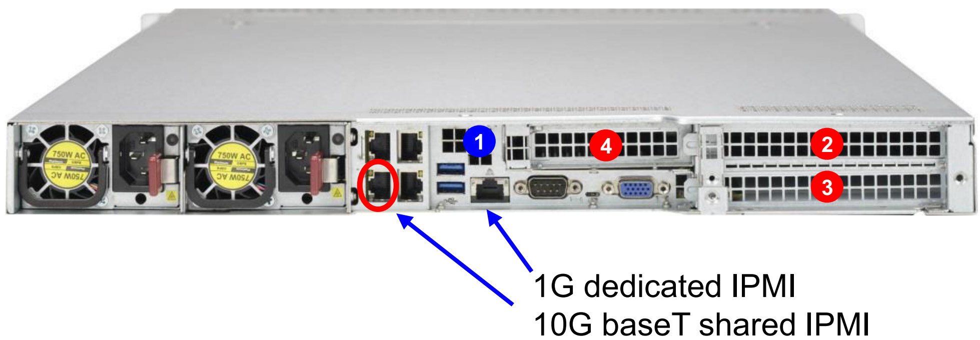
Forty-eight lanes * 2 processors * 16 Gbps * 2 (bidirectional traffic) = 3072 Gbps = 384 GBps

1.9.2 Slot configuration

The Power LC921 server has four PCIe slots, all of which are provided by the PCIe riser cards. Slot 1 is connected to CPU1. PCIe slots 2, 3 and 4 are connected to CPU2. PCIe slots 1, 2, and 3 are CAPI-enabled. PCIe slot 1 (internal only) and 4 are LP slots, and PCIe slots 2 and 3 are full-height and full-width.

Note: A double-width, full-height adapter at PCIe slot 2 also occupies PCIe slot 3.

Figure 1-14 shows the location of the PCI slots for the Power LC921 server.

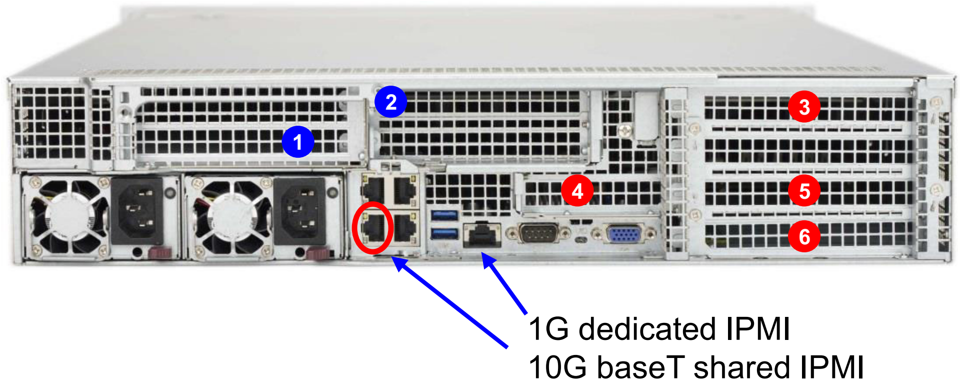


	PNOR name	Mechanical	Electrical
1	UIO slot 1	Internal low-profile (2.536”) half-length (6.6”)	x8 (CPU1-PEC0) CAPI
2	WIO slot 1	Double-width full-height (4.2”) 10.5” length	x16 (CPU2-PEC0) CAPI
3	WIO slot 2	Full-height (4.2”) 10.5” length **	x16 (CPU2-PEC2) CAPI
4	WIO-R slot	Low-profile (2.536”) half-length (6.6”)	x8 (CPU2-PEC1)
**	Slot 3 isblocked when a double-width adapter is installed in slot 2.		

Figure 1-14 Power LC921 PCIe slots: Rear view

The Power LC922 server supports up to 6 PCIe slots. PCIe slots 1 and 2 are connected to CPU1. PCIe slots 3, 4, 5, and 6 are connected to CPU2. PCIe slots 2, 3, and 6 are CAPI-enabled. PCIe slot 4 is an LP slot. PCIe slots 1, 5, and 6 are full-height slots. PCIe slots 2 and 3 are double-width, full-height slots.

Figure 1-15 shows the location of the PCI slots for the Power LC922 server.



	PNOR name	Mechanical	Electrical
1	UIO slot 2	Full height (4.2”) 10.5” length	x8 (CPU1-PEC1)
2	UIO slot 1	Double-width full-height (4.2”) 10.5” length	x16 (CPU1-PEC0) CAPI
3	WIO slot 1	Double-width full-height (4.2”) 10.5” length	x16 (CPU2-PEC0) CAPI
4	WIO-R slot	Low-profile (2.536”) half-length (6.6”)	x8 (CPU2-PEC1)
5	WIO slot 2	Full-height (4.2”) 10.5” length	x8 (CPU2-PEC2)
6	WIO slot 3	Full-height (4.2”) 10.5” length	x8 (CPU2-PEC2) CAPI

Figure 1-15 Power System LC922 PCI slots: Rear view

Note: A rear HDD tray occupies PCI slot 2, so you cannot install an adapter in PCI slot 2 while you use the rear HDD tray.

1.9.3 PCI adapters

This section describes the types and functions of the PCI adapters that are supported by the Power LC921 and Power LC922 servers.

The Power LC921 and Power LC922 servers uses the current PCIe Gen4 technology, which enables 32 GBps unidirectional and 64 GBps bidirectional bandwidth.

This section provides an overview of PCIe and bus speed and feature listings, which are segregated by function, for the supported PCIe adapters in the Power LC921 and Power LC922 servers.

Note: PCIe adapters on the Power LC921 and Power LC922 servers are not hot-pluggable.

Local area network and InfiniBand adapters

To connect the Power LC921 and Power LC922 servers to a LAN, you can use the LAN adapters that are supported in the PCIe slots of the system in addition to the standard 4-port 10 Gb BaseT Ethernet ports that are present in every system.

Figure 1-16 shows the port numbers of four 10/1/0.1 Gbps RJ45 ports that are integrated into the system.

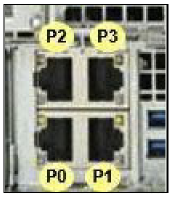


Figure 1-16 Integrated LAN port numbers

The base system has integrated quad-port 10/1/0.1 Gbps LAN ports. Various other LAN adapters are available, as shown in Table 1-14. Before selecting an adapter, see 1.9.2, “Slot configuration” on page 21.

Table 1-14 LAN and InfiniBand adapters

Feature code	Description
#EKA1	Cavium (QLogic) BCM57840 4-port SFP+ 10 Gb Ethernet Adapter PCIe3.0 x8 LP
#EKA2	Intel 82599ES Ethernet Converged Network Adapter x520-DA2 2-port 10 G/1 G SFP+ PCIe2.0 x8 LP
#EKA3	Intel 82575EB dual-port Gb Ethernet controller, PCIe x4
#EKAM	Mellanox MCX415A-CCAT ConnectX-4 EN 100 GbE single-port QSFP28 PCIe3.0 x16 LP
#EKAU	Mellanox MCX4121A-ACAT ConnectX-4 Lx EN 25 GbE dual-port SFP28 PCIe3.0 x8 LP
#EKAY	Mellanox MCX556A-EDAT ConnectX-5 VPI EDR InfiniBand 100 Gbps and 100 GbE dual-port QSFP28 PCIe4.0 x16 LP
#EKF1	Mellanox MCX414A-BCAT ConnectX-4 EN dual-port 40/56 GbE QFSP28 PCIe3.0 x8 LP
#EKFD	Mellanox MCX555A-ECAT ConnectX-5 VPI EDR InfiniBand 100 Gbps and 100 GbE single-port QSFP28 PCIe3.0 x16 LP
#EKFF	Broadcom 5719 QP 1 G (1 G/100M/10M) Network Interface Card PCIe x4 LP
#EKFH	Intel XL710 Ethernet Converged Network Adapter 4-port 10 G/1 G SFP+ PCIe3.0 x8 LP
#EKFP	Intel XL710/X557 10 GBase-T Converged Network Adapter 4-port (10 G/1 G/100M speeds) PCIe3.0 x8 LP

Table 1-15 shows that Intel network adapters require extra system memory, so when you configure the system memory, add the memory along with application sizing.

Table 1-15 Intel network adapters and the requirements for the system memory

Feature code (of optional Intel adapters)	Memory consumed per Intel adapter (in gigabytes)
#EKFP (4 ports)	15
#EKA3 (2 ports)	7.5
#EKA2 (2 ports)	7.5
#EKFH (4 ports)	15

If you use these Intel adapter options, ensure that the total system memory is large enough to meet the adapters memory requirements plus 32 GB, which is the minimum system memory.

Fibre Channel adapters

The servers support a direct or SAN connection to devices that use Fibre Channel connectivity.

Table 1-16 summarizes the available Fibre Channel adapters. All of them have LC connectors. The infrastructure that is used with these adapters determines whether you must procure LC fiber converter cables.

Table 1-16 Fiber adapters

Feature code	Description (Accessories)
#EKAF	Emulex LPE16002B-M6-O 2-port 16 Gb Fibre Channel card PCIe x8 LP
#EKAQ	QLogic QLE2692OP 2-port 16 Gb Fibre Channel adapter PCIe3 x8 LP
#EKFE	QLogic QLE2742 32 Gb Fibre Channel adapter PCIe3 x8 2-port LP

Non-Volatile Memory Express adapters

Table 1-17 shows the NVMe HBAs that may be used in the Power LC921 and Power LC922 servers.

Table 1-17 NVMe host bus adapters

Feature code	Description
#EKEG	SMC quad-port NVMe Host Bus Adapter (PEX9733) PCIe3.0 x8 LP with cables (2U server)
#EKAG	SMC quad-port NVMe Host Bus Adapter (PEX9733) PCIe3.0 x8 LP with cables (1U server)
#EKSQ	NVMe SSD; 1600 GB; AIC-HHHL; PCIe3.0 x8; 5 DWPD
#EKSR	NVMe SSD; 3200 GB; AIC-HHHL; PCIe3.0 x8; 5 DWPD

1.10 Internal storage

The Power LC921 server supports four HDDs or SSDs in the front. The Power LC922 server supports 12 HDDs in the front.

1.10.1 Power LC921 server direct attach drive backplane

The Power LC921 server has four bays of HDDs. All four slots are NVMe-enabled, as shown in Figure 1-17.

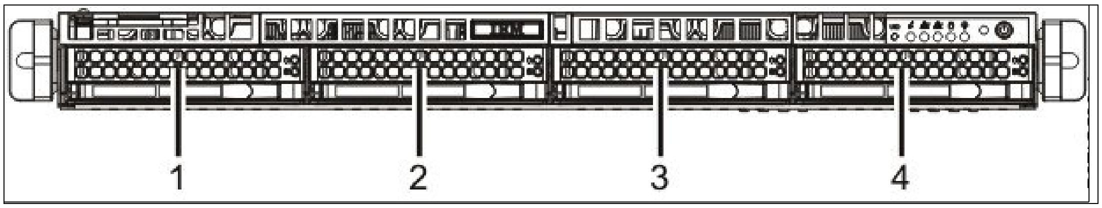


Figure 1-17 Four LFF/SFF SAS/SATA Bays (four NVMe-enabled)

1.10.2 Power LC922 server direct attach drive backplane

The Power LC922 server has an integrated PM8069 hard disk controller that supports various SAS/SATA disk options, as shown in Figure 1-18.

# of allowed SAS/SATA drives											
Integrated PM8069 controller 0 NVMe HBA + 0 NVMe drives		Integrated PM8069 controller + 1 NVMe HBA + 2 NVMe drives		Integrated PM8069 controller + 1 NVMe HBA + 4 NVMe drives		Integrated PM8069 controller + SAS/SATA HBA adapter + 0 NVMe HBA + 0 NVMe drives		Integrated PM8069 controller + SAS/SATA HBA adapter + 1 NVMe HBA + 2 NVMe drives		SAS/SATA HBA adapter + 1 NVMe HBA + 4 NVMe drives	
0 GPUs	1 or 2 GPUs	0 GPUs	1-2 GPUs	0 GPUs	1 or 2 GPUs	0 GPUs	1-2 GPUs	0 GPUs	1-2 GPUs	0 GPUs	1 or 2 GPUs
SAS/ SATA drives	SAS/ SATA drives	SAS/ SATA drives	SAS/ SATA drives	SAS/ SATA drives	N/A	4x SAS/SATA to PM8069 and 8x SAS/SATA to HBA adapter	2x SAS/SATA to PM8069 and 6x SAS/SATA to HBA adapter	2x SAS/SATA to PM8069 and 8x SAS/SATA to HBA adapter	6x SAS/SATA to HBA adapter	8x SAS/ SATA to HBA adapter	N/A
12	8	10	6	8	N/A	12	8	10	6	8	N/A

Figure 1-18 Integrated hard disk controller: SAS/SAATA drives

Figure 1-19 shows the drive location and numbering to port mapping for the Power LC922 backplane. NVMe drives are supported in positions 0 - 3 of port C through NVMe HBAs.

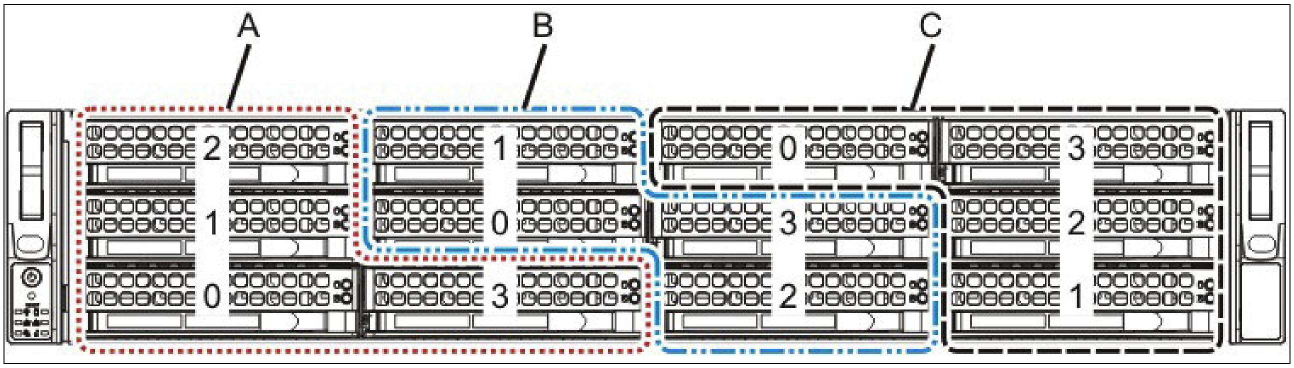


Figure 1-19 Drive port mapping

The Power LC922 serve can accommodate 12 SAS/SATA drives in the front.

Figure 1-20 shows the various SAS/SATA drive options for the server. An NVMe HBA adapter can accommodate two or four NVMe drives.

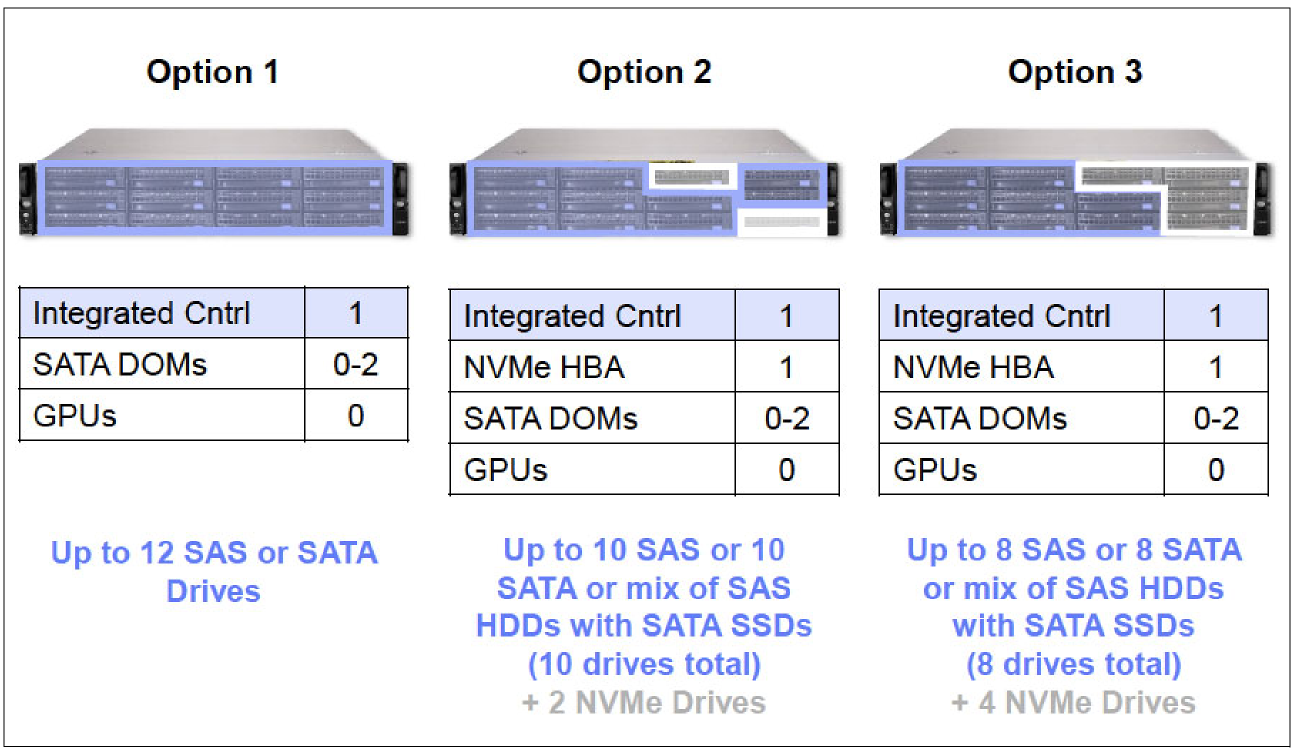


Figure 1-20 SAS/SATA drive options for the Power LC922 server

NVIDIA GPUs produce a significant amount of heat, so an efficient heat sink must be in the system, which can limit HDD placement in the server, as shown Figure 1-21.

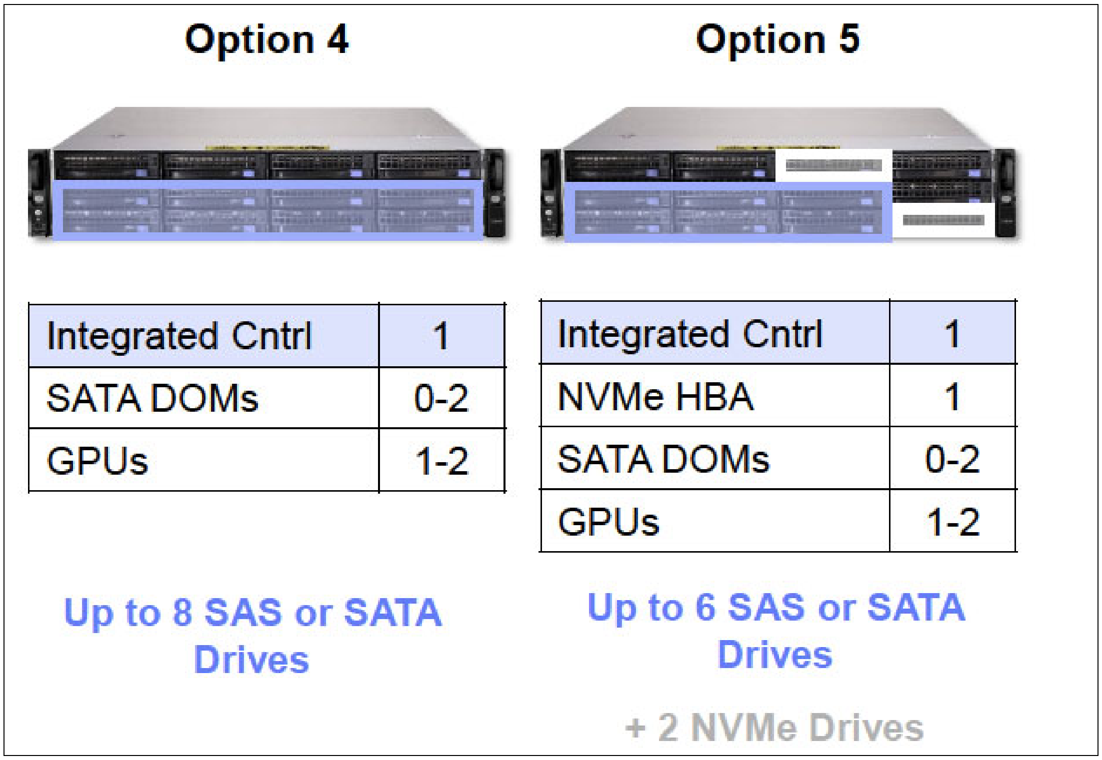


Figure 1-21 Hard disk drive placement limits with GPUs installed in the server

Figure 1-22 shows three options with two storage controllers:

- Option 6: There are no GPUs, and two controllers are present on the server. You may have a total of 12 HDDs.
- Option 7: There are GPUs, and two controllers are present on the server. You may have a maximum of eight HDDs (six HDDs with the first controller and two HDDs with the second controller).
- Option 8: There are GPUs and NVMe drives, and two controllers are present on the server. You may have a maximum of eight drives (six HDDs with the first controller and two NVMe drives with the NVMe HBA adapter).

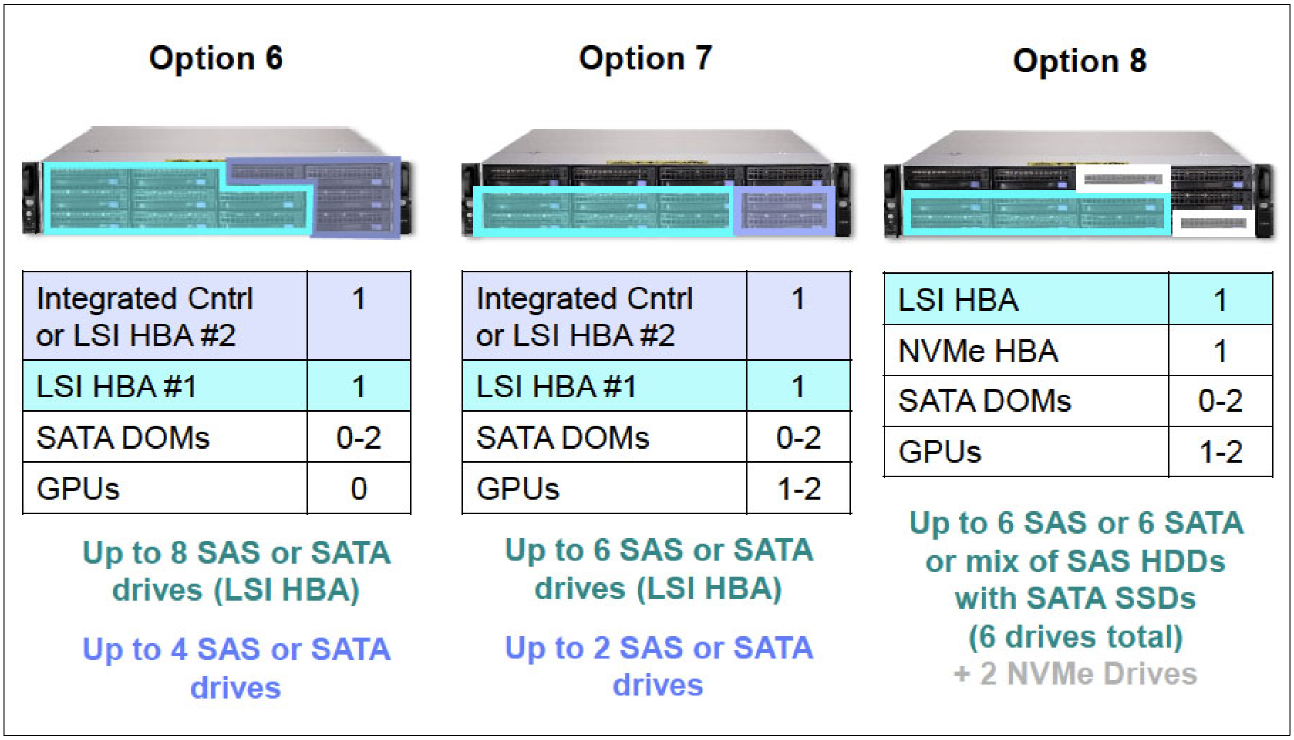


Figure 1-22 Hard disk drive options with two controllers

1.10.3 Power LC922 server: LSI drive backplane

You have the option of using an LSI drive backplane in the Power LC922 server. A maximum of 12 SAS/SATA drives can be installed when you use this backplane, but this capacity is reduced if you also install GPUs.

Table 1-18 shows the maximum number of drives with GPUs installed. The number of drives is further reduced when you use NVMe HBAs.

Notes:

- Requires an optional LSI HBA and cable kit of either #EKEA, #EKEH or #EKEB.
- This feature is required to access all 12 LFF/SFF drive bays from a single LSI HBA. An LSI HBA accommodates only eight LFF/SFF drive bays with a direct attach feature.

Table 1-18 SAS/SATA drives with GPUs

Number of allowed SAS/SATA drives			
GPUs	SAS/SATA controller and 0 - 1 NVMe PCIe HBAs		
	Zero NVMe HBAs and 0 NVMe Drives	One NVMe HBA and + 2 NVMe drives	One NVMe HBA and 4 NVMe drives
Zero	12	10	8
1 - 2	8	6	

Figure 1-23 shows the options for HDDs with an LSI HBA. There are similar options for the LSI HBA when you mix it with an NVMe HBA.

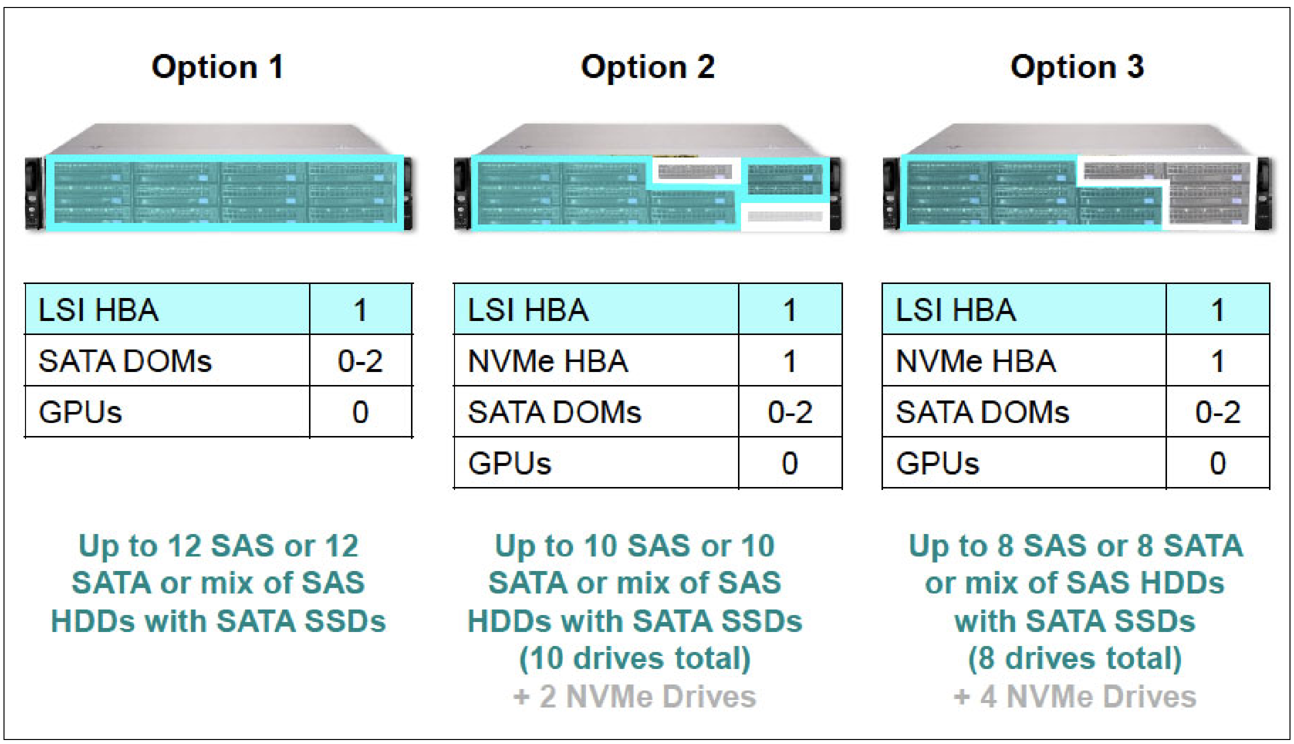


Figure 1-23 LSI storage controller and drive information without GPUs

Figure 1-24 shows the options for HDDs with an LSI HBA and NVIDIA GPUs installed.

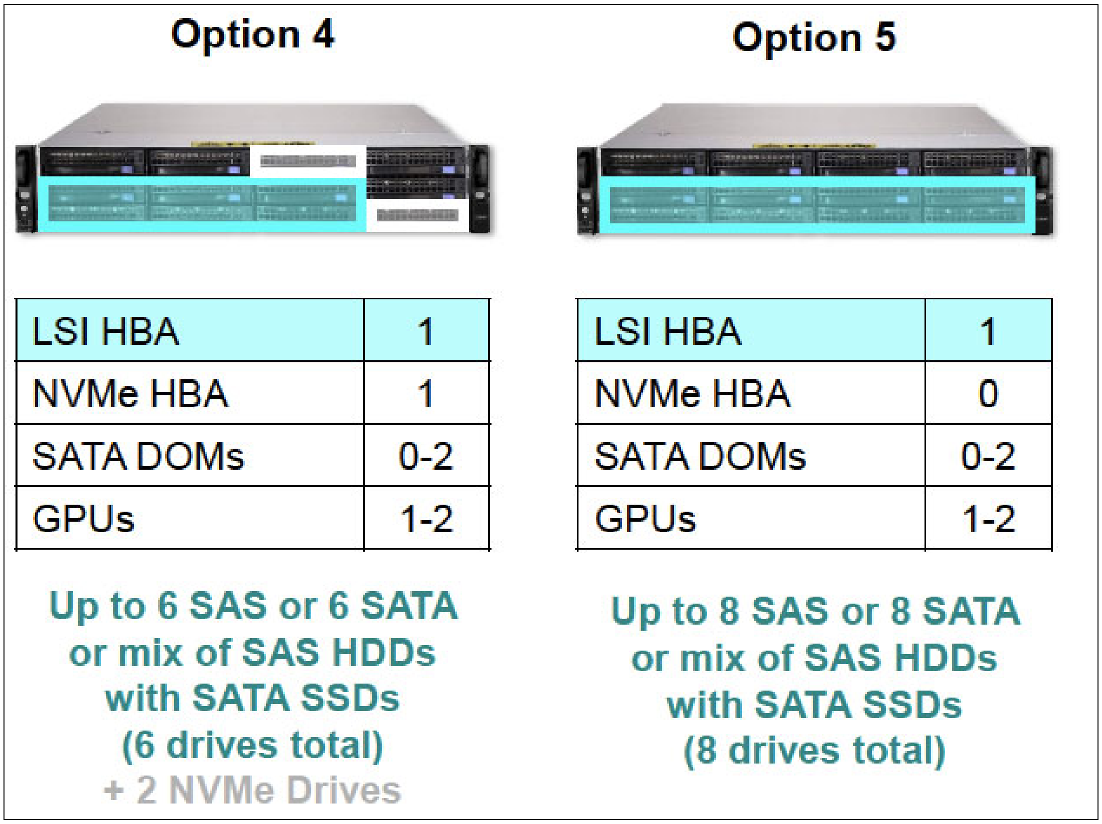


Figure 1-24 LSI storage controller and drives information with GPUs

1.10.4 Disk and media features

The Power LC921 and Power LC922 servers use the Integrated MicroSemi PM8069 SAS/SATA 16-port Internal Storage Controller PCIe3.0 x8 with RAID 0, 1, 5, and 10 support (no write cache). Note the following conditions:

- Selecting a 4 K HDD means that you must use an LSI HBA. SSDs must be formatted by the client to 512-byte sectors for RAID. The Integrated SAS/SATA controller (PM8069) does not support RAID 4 K devices (SSD or HDD). This setup is automatically checked when you place the order and is for informational purposes only.
- An integrated SAS/SATA controller (PM8069) does not have a SED feature. SED drives can be used by the PM0869, but it cannot unlock the encryption feature of those drives.

For 4 K HDDs, use the storage controllers that are shown in Table 1-19 (Power LC921) and Table 1-20 (Power LC922).

Table 1-19 Power LC921 storage controllers

Feature code	Description
#EKAA	Broadcom (LSI) MegaRAID 9361-8i SAS3 Controller with eight internal ports (1 GB Cache) PCIe3.0 x8 LP with cables
#EKAB	SMC AOC-K-S3008L-L8i 12 Gbps SAS3/RAID 0, 1, and 10 PCIe3.0 x8 LP with cables
#EKAH	Broadcom (LSI) MegaRAID 9361-8i SAS3 Controller with eight internal ports (2 GB Cache) PCIe3.0 x8 LP with cables

Table 1-20 Power LC922 storage controllers

Feature code	Description
#EKEA	Broadcom (LSI) MegaRAID 9361-8i SAS3 Controller with eight internal ports (1 GB Cache) PCIe3.0 x8 LP with cables
#EKEB	SMC AOC-K-S3008L-L8i 12 Gbps SAS3/RAID 0, 1, and 10 PCIe3.0 x8 LP with cables
#EKEH	Broadcom (LSI) MegaRAID 9361-8i SAS3 Controller with eight internal ports (2 GB Cache) PCIe3.0 x8 LP with cables

#EKAH and #EKEH are electronically identical. #EKAH is used in a 1U server and #EKEH is used in a 2U server, and this is also true for #EKAA/#EKEA and #EKAB/#EKEB. Only the cables are different for 1U and 2U, which forces different FCs for the kits.

Note: The SED drive requires one LSI MegaRAID HBA feature (#EKAA or #EKAH) along with either #EKWB (the software license) or #EKWC (the LSI hardware key) to unlock the LSI SafeStore feature. The SED feature cannot be used with LSI HBA #EKAB or #EKEB or with no adapter.

Table 1-21 shows a list of external HBA options to connect JBOD drawers.

Table 1-21 HBA options to connect JBOD drawers

Feature code	Description
#EKED	Broadcom (LSI) 9300-8E SAS-3 HBA PCIe3.0 x8 LP
#EKGC	Broadcom (LSI) 9305-16E SAS-3 HBA PCIe3.0 x8 LP

Table 1-22 provides a list of MES options:

- If no drive is ordered, the front drive bay is filled with #EKT1.
- All drive features include the correct drive carrier, but customers who damage carriers or want to install their own 2.5” (SFF) drives must order MES drive carriers.

Table 1-22 MES options

Feature code	Description
#EKT0	Tool-less 3.5" to 2.5" Converter Drive Tray
#EKT1	Hot-swap 3.5" HDD Tray with Hollowed Dummy
#EKT2	Tool-less U.2 NVMe 3.5" Drive Tray

Table 1-23 and Table 1-24 shows the various options for SATA/SAS HDD/SSD drives.

Table 1-23 SATA controllers that are based on HDDs/SSDs

Feature code	Description
#EKS1	240 GB 2.5-inch SATA SSD 0.78-DWPD SED
#EKSE	480 GB 2.5-inch SATA SSD 3.5-DWPD SED
#EKSF	960 GB 2.5-inch SATA SSD 3.5-DWPD SED
#EKS5	1.9 TB 2.5-inch SATA SSD 0.78-DWPD SED
#EKSG	1.92 TB 2.5-inch SATA SSD 3.5-DWPD SED
#EKS4	3.8 TB 2.5-inch SATA SSD 0.78-DWPD SED
#EKS3	960 GB 2.5-inch SATA SSD 0.6-DWPD NON-SED
#EKDA	2 TB 3.5-inch SATA 6 Gbps 512-Byte HDD NON-SED WrtCache
#EKDB	4 TB 3.5-inch SATA 6 Gbps 512 ByteHDD NON-SED WrtCache
#EKDD	8 TB 3.5-inch SATA 6 Gbps 512 ByteHDD NON-SED WrtCache
#EKDG	10 TB 3.5-inch SATA 6 Gbps 512-Byte HDD NON-SED WrtCache

Table 1-24 SAS Controller-based HDDs/SSDs

Feature code	Description
#EKS6	3.8 TB 2.5-inch SAS SSD 1-DWPD NON-SED
#EKS7	1.9 TB 2.5-inch SAS SSD 1-DWPD NON-SED
#EKS8	960 GB 2.5-inch SAS SSD 3-DWPD NON-SED
#EKS9	1.92 TB 2.5-inch SAS SSD 3-DWPD NON-SED
#EKDQ	1.2 TB 2.5-inch 10000 RPM SAS 12 Gbps 512e Byte HDD NON-SED WrtCache
#EKDR	1.8 TB 2.5-inch 10000 RPM SAS 12 Gbps 512e Byte HDD NON-SED WrtCache
#EKD5	HDD; 4 TB; 3.5"; 7200 RPM; SAS; 12 Gbps; 4 K; SED; WrtCache
#EKD6	HDD; 8 TB; 3.5"; 7200 RPM; SAS; 12 Gbps; 4 K; SED; WrtCache

Feature code	Description
#EKD1	HDD; 2 TB; 3.5"; 7200 RPM; SAS; 12 Gbps; 512e; Non-SED; WrtCache
#EKD2	HDD; 4 TB; 3.5"; 7200 RPM; SAS; 12 Gbps; 512e; Non-SED; WrtCache
#EKD4	HDD; 8 TB; 3.5"; 7200 RPM; SAS; 12 Gbps; 512e; Non-SED; WrtCache
#EKD8	HDD; 10 TB; 3.5"; 7200 RPM; SAS; 12 Gbps; 512e; Non-SED; WrtCache
#EKDA	HDD; 2 TB; 3.5"; 7200 RPM; SATA; 6 Gbps; 512e; Non-SED; WrtCache
#EKDB	HDD; 4 TB; 3.5"; 7200 RPM; SATA; 6 Gbps; 512e; Non-SED; WrtCache
#EKDD	HDD; 8 TB; 3.5"; 7200 RPM; SATA; 6 Gbps; 512e; Non-SED; WrtCache
#EKDG	HDD; 10 TB; 3.5"; 7200 RPM; SATA; 6 Gbps; 512e; Non-SED; WrtCache

Two SATA USB SuperDOMs (SATA DOMs) of 64 GB or 128 GB capacities can be used to boot. You use them to boot an OS or to perform other read-intensive operations only.

Figure 1-25 shows the location of a SuperDOM.

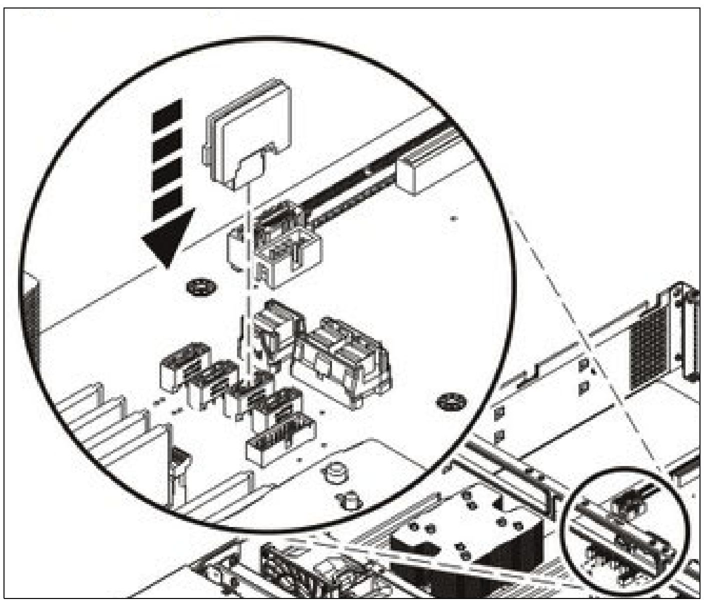


Figure 1-25 SuperDOM location

Table 1-25 shows the SATA DOM options.

Table 1-25 SATA Disk on Module

Feature code	Description
#EKSK	128 GB SATA DOM SuperDOM
#EKSL	64 GB SATA DOM SuperDOM

1.10.5 Power LC922 optional 2 x SFF rear SAS/SATA drives

The Power LC922 server can accommodate two SAS/SATA drives at the rear of the server.

Figure 1-26 shows the location of rear drive tray. It is intended for use by the OS. It occupies PCI slot 2, which is attached to CPU 1. You must order #EKSD to obtain an SFF rear tray.

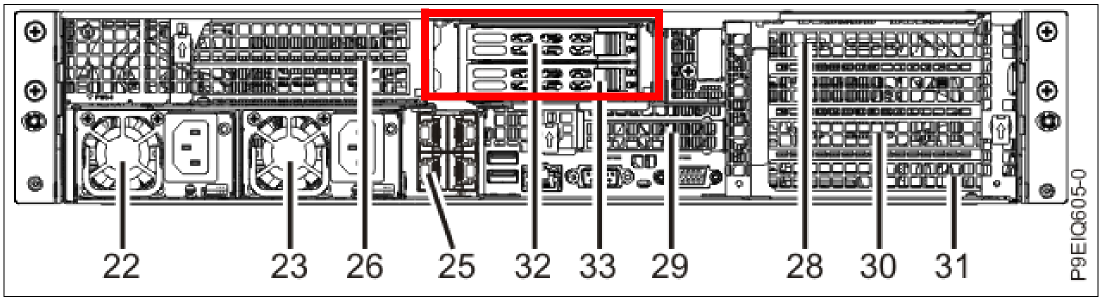


Figure 1-26 Optional SAS/SATA hot-swap rear drives

> **Note:** A single-Socket system that uses a rear drive tray has only one PCI slot for an adapter.
>
> The two SFF trays in the rear are controlled only by the integrated MicroSemi PM8069.

1.11 System ports

The rears of the Power LC921 and Power LC922 servers have a VGA port, a serial port, an IPMI or BMC port, two USB 3.0 ports, and a quad-port 10/1/0.1 Gbps RJ45 port. They are all integrated into the system board. Both servers support redundant power supply.

Figure 1-27 shows the system ports in the rear view of the Power LC921 server.

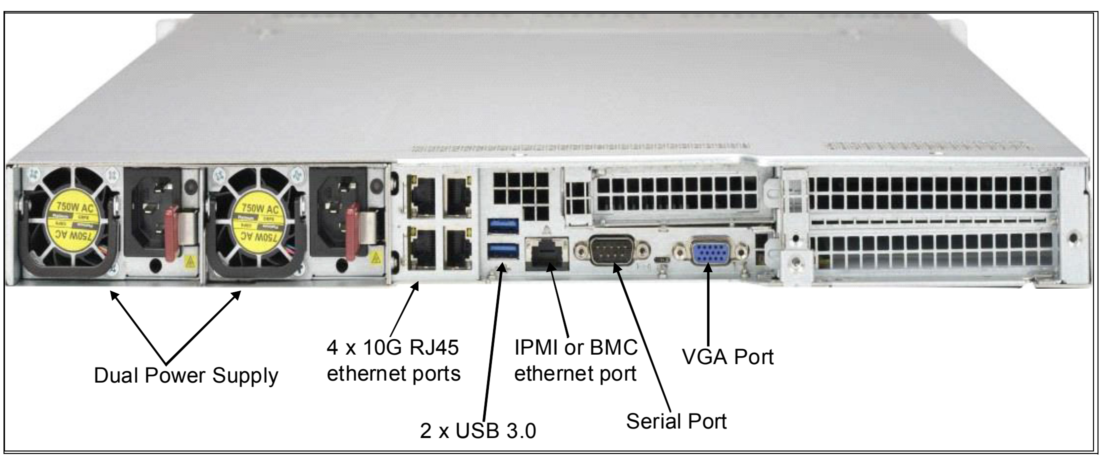


Figure 1-27 System ports of the Power LC921 server

Figure 1-28 shows the system ports in the rear view of the Power LC922 server.

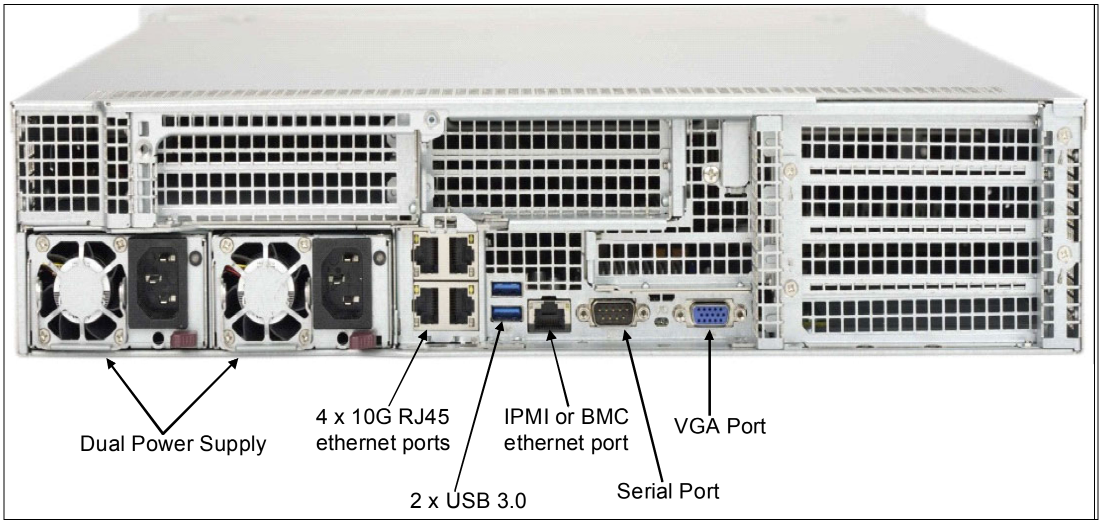


Figure 1-28 System ports of the Power LC922 server

These ports are available by default on both servers.

1.12 External I/O subsystems

The Power LC921 and Power LC922 servers do not support I/O drawers.

1.13 IBM Systems Storage

The IBM System Storage® disk systems products and offerings provide storage solutions with superior value for all levels of business, from entry-level to high-end storage systems. For more information about the various offerings, see IBM Storage.

The following section highlights a few of the offerings.

1.13.1 IBM Flash Storage

The next generation of IBM Flash Storage delivers the extreme performance and efficiency you need. It has a new pay-as-you-go option to reduce your costs and scale on demand. For more information about the hardware and software, see IBM Flash Storage.

1.13.2 Software-defined storage

Software-defined storage (SDS) manages data growth and enables multi-cloud flexibility by providing an agile, scalable, and operations-friendly infrastructure. For more information, see Software-defined storage.

1.13.3 IBM Hybrid Storage

Use IBM Hybrid Storage to optimize your mix of storage media, including flash storage, to achieve the best balance of performance and economics. For more information, see IBM Hybrid Storage.

1.13.4 IBM Storage Area Networks

IBM SAN Storage offers a comprehensive portfolio of Fibre Channel SAN switches that support your virtualization, cloud, and big data requirements. For more information, see IBM SAN Storage.

2

Management and virtualization

As you look for ways to maximize the return on your IT infrastructure investments, virtualization workloads become an attractive proposition.

The IBM Power System LC921 and IBM Power System LC922 servers are an excellent choice for clients that want to run their big data, Java, open source, and industry applications on a platform that is optimized for data and Linux.

This chapter identifies and clarifies the tools that are available for managing Linux on Power servers.

2.1 Main management components overview

Figure 2-1 shows the logical management flow of a Linux on Power server.

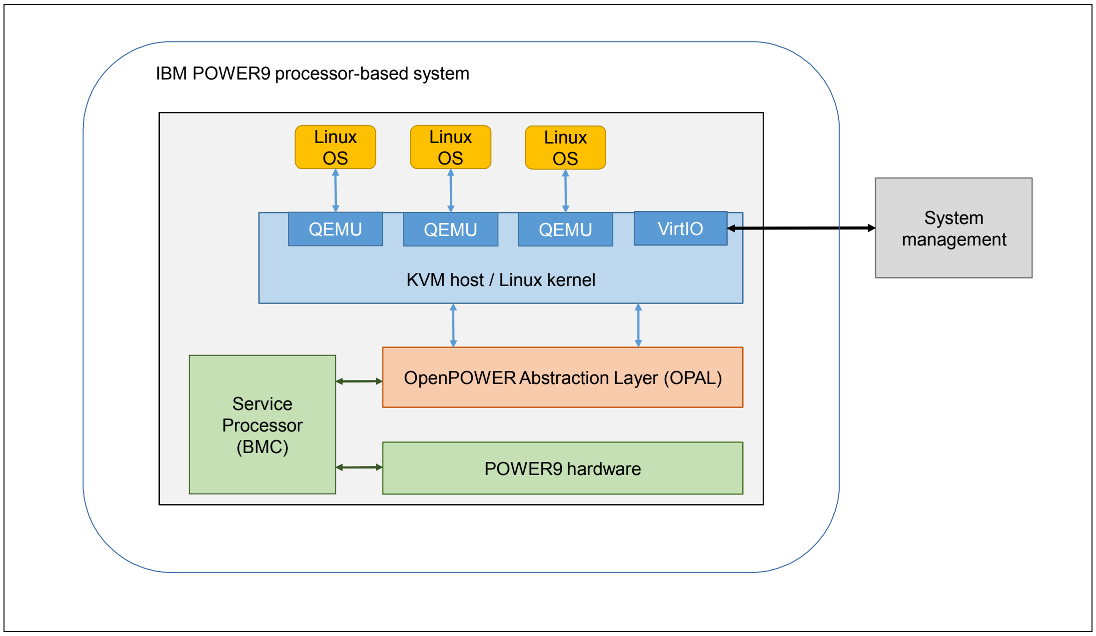


Figure 2-1 Logical diagram of a Linux on Power server

The service processor, or baseboard management controller (BMC), uses Kernel-based Virtual Machine (KVM) for virtual machines (VMs) and provides robust error detection and self-healing functions that are built in to the POWER9 processor and memory buffer modules.

The OpenPOWER Abstraction Layer (OPAL) is the system firmware in the stack of POWER9 processor-based Linux on Power servers.

The KVM Hypervisor technology offers key capabilities that can help you consolidate and simplify your IT environment. Quick Emulator (QEMU) is a generic and open source machine emulator and virtualizer that hosts the VMs on a KVM hypervisor. It is the software that manages and monitors the VMs.

KVM hosts can be managed by open source Linux tools that use the libvirt API, such as the Kimchi point-to-point administration tool and IBM Cloud PowerVC Manager.

IBM Cloud PowerVC Manager delivers easy-to-use advanced virtualization management capabilities that are virtualized by KVM. IBM PowerVC manages KVM VMs within a resource pool and enables the capture, deployment, and inventory of VM images.

2.2 Service processor

The service processor, or BMC, is the primary controlling mechanism for autonomous sensor monitoring and event logging features on the Power LC921 and Power LC922 servers.

BMC supports the Intelligent Platform Management Interface (IPMI V2.0) and Data Center Management Interface (DCMI V1.5) for system monitoring and management.

BMC monitors the operation of the firmware during the boot process and monitors the hypervisor for termination. Firmware code updates can be managed by the BMC and IPMI interfaces.

2.2.1 OpenPOWER Abstraction Layer

On Linux on Power, the OPAL firmware provides a hypervisor interface to the underlying hardware. OPAL firmware allows KVM to use the VirtIO API. The VirtIO API specifies an independent interface between the VMs and the service processor.

The VirtIO API is a high-performance API that para-virtualized devices use to gain speed and efficiency. VirtIO para-virtualized devices are especially useful for guest operating systems (OSes) that run I/O heavy tasks and applications.

For the Power LC921 and Power LC922 servers, OPAL Bare Metal is the only available system firmware.

2.2.2 Intelligent Platform Management Interface

The IPMI is an open standard for monitoring, logging, recovery, inventory, and control of hardware that is implemented independently of the main CPU, BIOS, and OS. It is the default console to use when you configure a KVM Host. The Power LC921 and Power LC922 servers provide one 10M/100/1000M baseT IPMI port.

The *ipmitool* is a utility for managing and configuring devices that support IPMI. It provides a simple command-line interface (CLI) to the service processor. You can install the ipmitool from the Linux distribution packages in your workstation or another server (preferably on the same network as the installed server). For example, to do the installation in Ubuntu, run the following command:

```
$ sudo apt-get install ipmitool
```

To connect to your system with IPMI, you must know the IP address of the server and have a valid password. To turn on the server with ipmitool, complete the following steps:

1. Open a terminal program.
2. Turn on your server by running the following command:
   ```
   ipmitool -I lanplus -H fsp_ip_address -P ipmi_password power on
   ```
3. Activate your IPMI console by running the following command:
   ```
   ipmitool -I lanplus -H fsp_ip_address -P ipmi_password sol activate
   ```

For more information about configuring Linux on a Linux on Power server, see IBM Knowledge Center.

2.2.3 Petitboot bootloader

Petitboot is a kexec-based bootloader that is used by POWER9 processor-based systems that are configured with KVM.

After the system turns on, the Petitboot bootloader scans local boot devices and network interfaces to find boot options that are available to the system. Petitboot returns a list of boot options that are available to the system.

If you are using a static IP or if you did not provide boot arguments in your network boot server, you must provide the details to Petitboot. You can configure Petitboot to find your boot server by following the instructions that are found at IBM Knowledge Center.

You can edit Petitboot configuration options, change the amount of time before Petitboot automatically starts, and so on, by following the instructions found at IBM Knowledge Center.

After you start the Linux installer, the installer wizard walks you through the steps to set up disk options, your root password, time zones, and so on.

You can read more about the Petitboot bootloader program at IBM Knowledge Center.

2.3 IBM Cloud PowerVC Manager

The following list shows the new functional enhancements that are provided by IBM Cloud PowerVC Manager V1.4.0:

- Database as a Service (DBaaS) technology preview.

  DBaaS refers to database instances that are deployed in a cloud where the user does not need to configure or maintain the database instance or its environment, and can access and use the database just like any other service. IBM PowerVC uses the OpenStack Trove component to deliver these capabilities.

- Software-defined infrastructure (SDI).

  With an SDI, you may virtualize nearly every aspect of your environment: compute resources, networking, and storage. You can use a combination of software-defined storage (SDS) and software-defined networking (SDN) to implement the virtualization level you want. The new IBM Cloud PowerVC Manager for SDI product bundles IBM Spectrum Scale™ and IBM Cloud PowerVC Manager, which makes SDI easier to implement.

- Software-defined storage.

  SDS (converged infrastructure) was previously a technology preview. This support is now included in the official product release for AIX and Linux environments. On systems that are managed by IBM PowerVM® NovaLink, you can use SDS. SDS requires the SDN capability. Together, these two technologies allow administrators to have a predefined node type that has network, storage, and compute resources all within the server. Then, as you increase the number of servers in your environment, the infrastructure elastically and automatically grows. For more information, see 1.13.2, “Software-defined storage” on page 35.

- Change VM ownership from the user interface.

  Administrators and project managers can change the owner of a VM from the VM's details page.

- Capture a live VM.

  You can capture running workloads to avoid disruptions in your environment without first enabling the VM for live capture. If the VM is running when it is captured, a point-in-time snapshot is created, such as would be created if a server were immediately turned off. For more permanent captures, such as for creating a master image, stop the VM first.

- Use SAN-backed clusters.

  When using SDS, you can now use SAN-backed or local disk-backed clusters. All disks must be SAN-backed or all disks must be local-backed within each cluster.

- Flexible SAN zone names.

  You can use IBM PowerVC to use a template to customize SAN zone names. To set up this template, use the following command:

  ```
  powervc-config storage fc-zone zone-name-template
  ```

- Specify multiple IP address ranges and edit IP address ranges.

  When adding a network, you can now specify multiple IP address ranges, and you can edit the IP address ranges of existing networks.

- Additional project quotas.

  These new quotas are supported: Volume Backup (GB), Volume Backups, and Volume Groups.

- Set the Physical Page Table (PPT) ratio.

  The PPT ratio controls the size of the page table that is used by the hypervisor when performing Live Partition Mobility (LPM). The larger the PPT, the more entries are available for use by the partition for mapping virtual addresses to physical real addresses. If the PPT is set too small and the partition is running workloads, performance in the partition can decline during LPM. This setting is helpful for SAP HANA workloads. You can set the PPT ratio on a compute template by using either the user interface or the Flavors API.

- Create or attach multiple volumes.

  You can create more than one new volume of the same size and under the same storage template on the Create Volume page. You can also attach multiple volumes in a single operation.

- User interface updates.

  Along with a new look and feel, some items on the user interface were moved, renamed, or have different behaviors:

  - Links to fabrics and networks were removed from the home page.
  - The Verify environment button was moved to a new tab on the home page. This tab is called Environment Checker. This tab also displays the most recent Environment Checker results.
  - The Configuration page was reorganized.
  - The IP Addresses tab on the Network page shows only addresses that are locked or in use.
  - The storage count in the navigation pane indicates the number of registered storage providers instead of volumes.

- IBM POWER9 processor-based servers support.

  IBM PowerVC now supports POWER9 processor-based hardware (requires interim fix 1 and NovaLink 1.0.0.9).

- Support of OpenPOWER and other systems running KVM on Power.

  IBM PowerVC now supports KVM on Power running Ubuntu 16.04. KVM-based systems such as OpenPOWER can be managed from the same installation of IBM PowerVC that is managing the existing PowerVM infrastructure. This support builds on the SDN and the new SDS features.

3

Reliability, availability, and serviceability

The Power System LC922 and Power System LC921 servers bring POWER9 processor-based and memory reliability, availability, and serviceability (RAS) functions to a cloud data center, with open source Linux technology supplying the operating system (OS) and virtualization. The OpenPOWER Abstraction Layer (OPAL) firmware provides a hypervisor and OS-independent layer that uses the error detection and self-healing functions that are built into the POWER9 processor.

The processor address paths and data paths, the control logic, state machines, and computational units are protected with parity or error correction code (ECC). The processor core soft errors or intermittent errors are recovered with processor instruction retry. Unrecoverable errors are reported as machine check (MC) errors, and errors that affect the integrity of data lead to a system checkstop.

3.1 Processor and cache reliability

The Level 1 (L1) data and instruction caches in each processor core are parity-protected. Data is stored through to L2 immediately. L1 caches have a retry capability for intermittent errors and a cache set delete mechanism for handling solid failures. The L2 and L3 caches in the POWER9 processor are protected with double-bit detect, single-bit correct ECC.

In addition, a threshold of correctable errors that is detected on cache lines can result in the data in the cache lines being purged and the cache lines removed from further access without requiring a restart. An uncorrectable error that is detected in these caches can trigger a purge and deletion of cache lines, which does not impact the current operation if the cache lines contained data that was unmodified from what was stored in system memory.

The memory subsystem has proactive memory scrubbing to help prevent the accumulation of multiple single-bit errors. The ECC scheme can correct the complete failure of any one memory module within an ECC word. After marking the module as unusable, the ECC logic can still correct single symbol (two adjacent bit) errors. An uncorrectable error of data of any layer of cache up to the main memory is marked to prevent usage of fault data. The processor's memory controller has retry capabilities for certain fetch and store faults.

3.1.1 L3 cache line delete

The L3 cache is protected by ECC and Special Uncorrectable Error (SUE) handling. The L3 cache also incorporates technology to handle memory cell errors through a special cache line delete algorithm.

During the central electronics complex initial program load (IPL), if a solid error is detected during L3 initialization, a full L3 cache line is deleted. During system run time, a correctable error is reported as a recoverable error to the service processor. If an individual cache line reaches its predictive error threshold, it is dynamically deleted.

The state of the L3 cache line deletion is maintained in a “deallocation record” and persists through system IPL, which ensures that cache lines “varied offline” by the server remain offline if the server restarts. Therefore, these “error prone” lines cannot cause system operational problems. The server can dynamically delete up to several multiple cache lines. If this threshold is reached, the L3 is marked for persistent deconfiguration on subsequent system restarts until repaired.

3.1.2 Special Uncorrectable Error handling

SUE handling prevents an uncorrectable error in memory or cache from immediately causing the system to terminate. Rather, the system tags the data and determines whether it will ever be used again. If the error is irrelevant, it will not force a checkstop. If the data is used, termination may be limited to the program/kernel or hypervisor owning the data, or a “freeze” occurs to the I/O adapters that are controlled by an I/O hub controller if data would be transferred to an I/O device.

3.1.3 Thermal management and current/voltage monitoring

The On Chip Controller (OCC) monitors various temperature sensors in the processor module, memory modules, and environmental temperature sensors, and steers the throttling of processor cores and memory channels if the temperature rises over thresholds that are defined by the design. The power supplies have their own independent thermal sensors and monitoring. Power supplies and voltage regulator modules (VRMs) monitor over-voltage, under-voltage, and over-current conditions. They report into a power good tree that is monitored by the service processor.

3.1.4 GPU throttling

When the temperature of the GPUs is too high, the GPUs throttle performance to improve cooling.

3.1.5 PCI extended error handling

PCI extended error handling (EEH)-enabled adapters respond to a special data packet that is generated from the affected PCI slot hardware by calling system firmware, which examines the affected bus, allows the device driver to reset it, and continues without a system restart. For Linux, EEH support extends to most of the frequently used devices, although some third-party PCI devices might not provide EEH support.

3.1.6 Chassis policy after input power loss and autorestart after system checkstop

The boot parameter **`chassis policy`** controls whether the server returns to the previous state or starts axiomatically after an input power loss. The system automatically restarts after a system checkstop, and the system management software decides whether to use the server with potentially fewer resources.

3.2 Serviceability

The server is designed for system installation and setup, feature installation and removal, proactive maintenance, and corrective repair that is performed by the client. Warranty Service Upgrades are offered for On Site Repair (OSR) by an IBM System Services Representative (IBM SSR) or authorized warranty service provider.

IBM Knowledge Center provides current documentation to service effectively the system:

- *Quick Install Guide*
- *User's Guide*
- *Troubleshooting Guide*
- *Boot Configuration Guide*

The documentation can be downloaded in PDF format or used online with an internet connection at IBM Knowledge Center (Power LC921 server) and IBM Knowledge Center (Poewr LC922 server).

3.2.1 Service processor

The service processor supports Intelligent Platform Management Interface (IPMI 2.0), Data Center Management Interface (DCMI 1.5), and Simple Network Management Protocol (SNMP V2 and V3) for system monitoring and management. The service processor provides platform system functions such as power on/off, power sequencing, power fault monitoring, power reporting, fan/thermal control, fault monitoring, vital product data (VPD) inventory collection, Serial over LAN (SOL), Service Indicator LED management, code update, and event reporting through system event logs (SELs).

All SELs can be retrieved either directly from the service processor or from the host OS (Linux). The service processor monitors the operation of the firmware during the boot process and also monitors the hypervisor for termination. Firmware code update is supported through the service processor and IPMI interface. The firmware image can be updated or flashed regardless of its current state.

3.2.2 Service interface

The service interface enables the client and the support personnel to communicate with the service support applications in a server by connecting directly or remotely through a web browser or command-line interface (CLI). It provides access to various service applications and available actions. The service interface enables client and support personnel to manage system resources, inventory, and service information in an efficient and effective way.

Different service interfaces are used, depending on the state of the system and its operating environment. The primary service interfaces are:

- Service processor: Ethernet service network with IPMI V2.0, or a systems management GUI through a web browser
- Service indicator LEDs: System attention and system identification (front and back)
- Host OS: CLI

The primary service applications are:

- SELs
- OS event logs
- Sensor status GUI LEDs for problem determination when next to the system

3.2.3 Concurrent maintenance

The following components can be replaced without turning off the server:

- Hard disk drives (HDDs)
- Redundant hot-plug power supplies

3.2.4 Error handling and reporting

If there is a system hardware failure or environmentally induced failure, the system error capture capability systematically analyzes the hardware error signature to determine the cause of failure. The processor and memory recoverable errors are handled through Processor Runtime Diagnostics (PRD) in the OPAL layer, which generates a SEL. An extended SEL (eSEL) is associated with each SEL. It contains additional First Failure Data Capture (FFDC) information for use by the support structure.

For system checkstop errors, the OCC collects Failure Information Register (FIR) data and saves it in nonvolatile memory. PRD analyzes the data upon restart and creates a SEL and eSEL. The host Linux OS can monitor the SELs on the service processor through the IPMI tool. Hardware and firmware failures are recorded in the SELs and can be retrieved through the IPMI interface. The system can report errors that are associated with Peripheral Component Interconnect Express (PCIe) adapters/devices through the host OS.

A

Server racks and energy management

This appendix provides information about the racking options and energy management-related concepts that are available for the IBM Power System LC921 and IBM Power System LC922 servers.

IBM server racks

The Power LC921 server mounts in the 42U Slim Rack (7965-94Y) along with square hole industry-standard racks. These racks are built to the 19-inch EIA 310D standard.

Order information: Power LC921 servers cannot be integrated into these racks during the manufacturing process, and are not orderable together with servers. If the Power LC921 server and any of the supported IBM racks are ordered together, they are shipped at the same time in the same shipment, but in separate packing material. IBM does not offer integration of the server into the rack before shipping.

If a system is installed in a rack or cabinet that is not an IBM rack, ensure that the rack meets the requirements that are described in "OEM racks" on page 53.

An IBM Business Partner can offer installation of these servers into a rack before delivery to a clients site, or at the client's site.

Responsibility: The client is responsible for ensuring that the installation of the drawer in the preferred rack or cabinet results in a configuration that is stable, serviceable, safe, and compatible with the drawer requirements for power, cooling, cable management, weight, and rail security.

Note: Round hole interposers are automatically included in new systems that fit into IBM branded T42 and 7965-S42 racks.

IBM 42U SlimRack 7965-94Y

The 2.0-meter (79-inch) Model 7965-94Y is compatible with all Power Systems servers. It provides an excellent 19-inch rack enclosure for your data center. Its 600 mm (23.6 in.) width combined with its 1100 mm (43.3 in.) depth plus its 42 EIA enclosure capacity provides great footprint efficiency for your systems. It can be easily placed on standard 24-inch floor tiles.

The IBM 42U Slim Rack has a lockable perforated front steel door that provides ventilation, physical security, and visibility of indicator lights in the installed equipment within. In the rear, either a lockable perforated rear steel door (#EC02) or a lockable Rear Door Heat Exchanger (RDHX)(1164-95X) is used. Lockable optional side panels (#EC03) increase the rack's aesthetics, help control airflow through the rack, and provide physical security. Multiple 42U Slim Racks can be bolted together to create a rack suite (#EC04).

Up to six optional 1U power distribution units (PDUs) can be placed vertically in the sides of the rack. Additional PDUs can be placed horizontally, but they each use 1U of space in this position.

The AC power distribution unit and rack content

For the 42U SlimRack (7965-94Y) rack model, 12-outlet PDUs are available. These PDUs include the AC PDU (#7188) and the AC Intelligent PDU+ (iPDU) (#7109). The iPDU is identical to the PDU, but it is equipped with one Ethernet port, one console serial port, and one RS232 serial port for power monitoring.

The PDUs have 12 client-usable IEC 320-C13 outlets. Six groups of two outlets are fed by six circuit breakers. Each outlet is rated up to 10 A, but each group of two outlets is fed from one 15 A circuit breaker.

Four PDUs can be mounted vertically in the back of the 7965-94Y rack.

Figure A-1 shows the placement of the four vertically mounted PDUs.

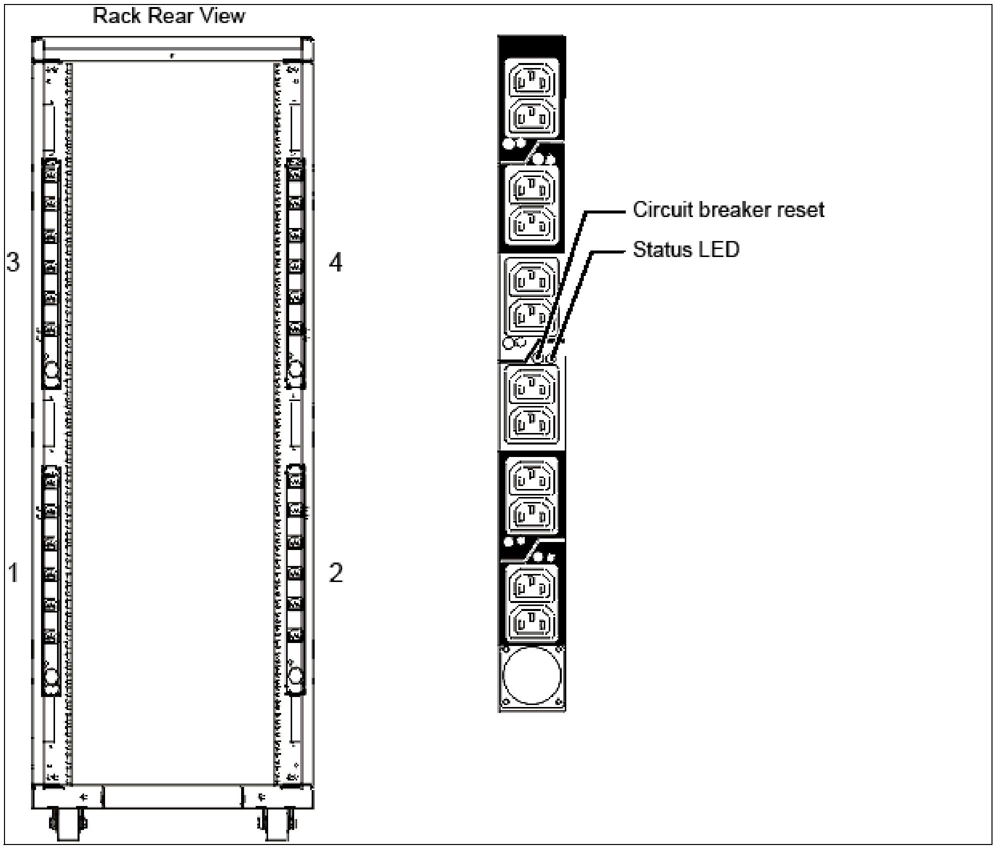


Figure A-1 Power distribution unit placement and view

In the rear of the rack, two additional PDUs can be installed horizontally in the 7965-94Y rack. The four vertical mounting locations are filled first in the 7965-94Y slimline rack. Mounting PDUs horizontally uses 1U per PDU and reduces the space that is available for other racked components. When mounting PDUs horizontally, the preferred approach is to use fillers in the EIA units that are occupied by these PDUs to facilitate the correct airflow and ventilation in the rack.

The PDU receives power through a UTG0247 power-line connector. Each PDU requires one PDU-to-wall power cord. Various power cord features are available for various 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.

Table A-1 shows the available wall power cord options for the PDU and iPDU features, which must be ordered separately.

Table A-1 Wall power cord options for the power distribution unit and intelligent power distribution unit features

Feature code	Wall plug	Rated voltage (V AC)	Phase	Rated amperage	Geography
#6653	IEC 309, 3P+N+G, 16 A	230	3	16 amps/phase	Internationally available
#6489	IEC309 3P+N+G, 32 A	230	3	32 amps/phase	EMEA
#6654	NEMA L6-30	200 - 208, 240	1	24 amps	US, Canada, LA, and Japan
#6655	RS 3750DP (watertight)	200 - 208, 240	1	24 amps	US, Canada, LA, and Japan
#6656	IEC 309, P+N+G, 32 A	230	1	24 amps	EMEA
#6657	PDL	230 - 240	1	32 amps	Australia and New Zealand
#6658	Korean plug	220	1	30 amps	North and South Korea
#6492	IEC 309, 2P+G, 60 A	200 - 208, 240	1	48 amps	US, Canada, LA, and Japan
#6491	IEC 309, P+N+G, 63 A	230	1	63 amps	EMEA

Notes: Ensure that the correct power cord feature is configured to support the power that is being supplied. Based on the power cord that is used, the PDU can supply 4.8 - 19.2 kVA. The power of all of the drawers that are plugged into the PDU must not exceed the power cord limitation.

The Universal PDUs are compatible with previous models.

To better enable electrical redundancy, each server has two power supplies that must be connected to separate PDUs, which are not included in the base order.

For maximum availability, a preferred approach is to connect power cords from the same system to two separate PDUs in the rack, and to connect each PDU to independent power sources.

For detailed power requirements and power cord details about the 7014 racks, see the “Planning for power” section in IBM Knowledge Center.

For detailed power requirements and power cord details about the 7965-94Y rack, see the “Planning for power” section in IBM Knowledge Center.

Rack-mounting rules

Consider the following primary rules when you mount the system into a rack:

- The system can be placed at any location in the rack. For rack stability, start filling a rack from the bottom.
- Any remaining space in the rack can be used to install other systems or peripheral devices if the maximum permissible weight of the rack is not exceeded and the installation rules for these devices are followed.
- Before placing the system into the service position, be sure to follow the rack manufacturer's safety instructions regarding rack stability.

OEM racks

The system can be installed in a suitable square-hole OEM rack if that the rack conforms to the EIA-310-D standard for 19-inch racks. This standard is published by the Electrical Industries Alliance. For more information, see IBM Knowledge Center.

Table A-2 shows the Power LC921 server round hole to square hole conversion feature.

Table A-2 Power LC921 server round hole to square hole conversion

Feature code	Description
#EU0H	Round Hole Rack Conversion Kit for 1U (EIA units) server

Table A-3 shows the Power LC922 server round hole to square hole conversion feature.

Table A-3 Power LC922 server round hole to square hole conversion

Feature code	Description
#EU0J	2U Round Hole Rack Converter Kit

Note: These round hole adapters are automatically included in new systems that fit into IBM branded T42 and 7965-S42 racks, but these features are available in case these interposers are misplaced.

Energy consumption estimation

For Power Systems servers, various energy-related values are important:

- Maximum power consumption and power source loading values

  These values are important for site planning. They are described in IBM Knowledge Center.

  Search for type and model number and “server specifications”. For example, for the Power LC922 servers, search for “9006-22P”.

- An estimation of the energy consumption for a certain configuration

  Calculate the energy consumption for a certain configuration in the IBM Systems Energy Estimator.

  In that tool, select the type and model for the system, and enter details about the configuration and CPU usage that you want. The tool shows the estimated energy consumption and the waste heat at the usage that you want and also at full usage.

Related publications

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

IBM Redbooks

The following IBM Redbooks publications provide more information about the topic in this document. Some publications that are referenced in this list might be available in softcopy only.

- *IBM PowerAI: Deep Learning Unleashed on IBM Power Systems Servers*, SG24-8409
- *IBM Power System L922 Technical Overview and Introduction*, REDP-5496
- *IBM Power System S822LC for High Performance Computing Introduction and Technical Overview*, REDP-5405
- *IBM Power Systems H922 and H924 Technical Overview and Introduction*, REDP-5498
- *IBM Power Systems S922, S914, and S924 Technical Overview and Introduction*, REDP-5497
- *IBM PowerVM Best Practices*, SG24-8062
- *IBM PowerVM Virtualization Introduction and Configuration*, SG24-7940
- *IBM PowerVM Virtualization Managing and Monitoring*, SG24-7590

You can search for, view, download, or order these documents and other Redbooks publications, Redpapers, web docs, drafts, and additional materials, at the following website:

ibm.com/redbooks

Online resources

These websites are also relevant as further information sources:

- IBM Fix Central website

  http://www.ibm.com/support/fixcentral/
- IBM Knowledge Center

  http://www.ibm.com/support/knowledgecenter/
- IBM Knowledge Center: IBM Power Systems Hardware

  http://www-01.ibm.com/support/knowledgecenter/api/redirect/powersys/v3r1m5/index.jsp
- IBM Knowledge Center: Migration combinations of processor compatibility modes for active Partition Mobility

  http://www-01.ibm.com/support/knowledgecenter/api/redirect/powersys/v3r1m5/topic/p7hc3/iphc3pcmcombosact.htm

- IBM Portal for OpenPOWER - POWER9 Monza Module
  https://www.ibm.com/systems/power/openpower/tgcmDocumentRepository.xhtml?aliasId=POWER9_Sforza
- IBM Power Systems
  http://www.ibm.com/systems/power/
- IBM Storage website
  http://www.ibm.com/systems/storage/
- IBM System Planning Tool website
  http://www.ibm.com/systems/support/tools/systemplanningtool/
- IBM Systems Energy Estimator
  http://www-912.ibm.com/see/EnergyEstimator/
- NVIDIA Tesla V100
  https://www.nvidia.com/en-us/data-center/tesla-v100/
- NVIDIA Tesla V100 Performance Guide
  http://images.nvidia.com/content/pdf/volta-marketing-v100-performance-guide-us-r6-web.pdf
- OpenPOWER Foundation
  https://openpowerfoundation.org/
- Power Systems Capacity on Demand website
  http://www.ibm.com/systems/power/hardware/cod/
- Support for IBM Systems website
  http://www.ibm.com/support/entry/portal/Overview?brandind=Hardware~Systems~Power

Help from IBM

IBM Support and downloads

ibm.com/support

IBM Global Services

ibm.com/services

REDP-5495-00

ISBN 0738457140

Printed in U.S.A.