IBM Power System IC922
Technical Overview and Introduction

Stephen Lutz
YoungHoon Cho
IBM Redbooks

IBM Power System IC922 Technical Overview and Introduction

March 2020
Note: Before using this information and the product it supports, read the information in "Notices" on page v.

First Edition (March 2020)

This edition applies to the IBM Power System IC922 (9183-22X).

This document was created or updated on January 30, 2020.

© Copyright International Business Machines Corporation 2020. All rights reserved.
Note to U.S. Government Users Restricted Rights -- Use, duplication or disclosure restricted by GSA ADP Schedule Contract with IBM Corp.
# Contents

- **Notices** ................................................................. v
- **Trademarks** ............................................................. vi
- **Preface** ................................................................. vii
- **Authors** ................................................................. vii
- **Now you can become a published author, too!** .................. viii
- **Comments welcome.** ................................................ viii
- **Stay connected to IBM Redbooks** .................................. viii

## Chapter 1. Architectural and technical description ........................................ 1

1.1 Power IC922 hardware overview .................................................. 2
   1.1.1 Statements of direction .................................................. 3
   1.1.2 Minimum features ........................................................ 3
   1.1.3 Configurable feature matrix ............................................. 4
1.2 Operating system support .......................................................... 4
   1.2.1 Red Hat Enterprise Linux ............................................... 5
1.3 Operating environment ............................................................ 5
1.4 Physical package ........................................................................ 6
1.5 System architecture .................................................................... 7
1.6 The POWER9 processor ............................................................... 9
   1.6.1 POWER9 processor overview ............................................. 9
   1.6.2 Processor features .......................................................... 10
   1.6.3 Supported technologies .................................................... 11
   1.6.4 Simultaneous multithreading ............................................. 12
   1.6.5 Processor feature codes ................................................... 12
   1.6.6 Coherent Accelerator Processor Interface 2.0 ....................... 13
   1.6.7 OpenCAPI 3.0 ............................................................... 14
1.7 Memory subsystem ................................................................. 16
   1.7.1 DIMM placement rules .................................................... 18
   1.7.2 Memory bandwidth ......................................................... 19
1.8 The Trusted Platform Module ...................................................... 19
1.9 Internal I/O subsystem ............................................................... 20
   1.9.1 PCIe Express Controller and PCI Express ......................... 20
   1.9.2 Slot configuration .......................................................... 21
   1.9.3 GPU ........................................................................... 23
   1.9.4 FPGA ......................................................................... 23
   1.9.5 PCI adapters ................................................................. 23
1.10 Internal storage ........................................................................ 26
   1.10.1 Drive backplanes .......................................................... 27
   1.10.2 HDD and SSD .............................................................. 28
   1.10.3 NVMe ................................................................. 28
1.11 System ports ........................................................................... 29
1.12 External I/O subsystems ............................................................. 29
1.13 IBM Systems Storage ................................................................. 29
   1.13.1 IBM Flash Storage ......................................................... 29
   1.13.2 Software-defined storage ................................................ 30
   1.13.3 IBM Hybrid Storage ....................................................... 30
   1.13.4 IBM Storage Area Networks ......................................... 30
Chapter 2. Systems management .......................... 31
  2.1 Service processor ...................................... 32
    2.1.1 Intelligent Platform Management Interface .......... 32
    2.1.2 Petitboot bootloader ................................. 32
    2.1.3 Configuring the network to the BMC ................. 33
    2.1.4 Systems management using the OpenBMC tool ....... 36
    2.1.5 The OpenBMC GUI .................................. 37
    2.1.6 Connect to the BMC using SSH ....................... 39
    2.1.7 Using a HMC for systems management ............... 40
    2.1.8 Working with the BMC over the REST API .......... 40
  2.2 Installation of the operating system .................. 41
    2.2.1 Installation using a virtual media ................... 41
    2.2.2 Installation using an USB stick ...................... 44

Chapter 3. Reliability, availability, and serviceability .... 45
  3.1 Processor and cache reliability ....................... 46
    3.1.1 L3 cache line delete ............................... 46
    3.1.2 Special Uncorrectable Error handling ................ 46
    3.1.3 Thermal management and current/voltage monitoring . 46
    3.1.4 PCI extended error handling ......................... 47
  3.2 Serviceability ......................................... 47
    3.2.1 Concurrent maintenance .............................. 47
    3.2.2 Service processor .................................... 48
    3.2.3 Get SEL data using the OpenBMC tool ............... 48
    3.2.4 The system LEDs .................................... 48
    3.2.5 Factory reset ....................................... 49
    3.2.6 Error handling and reporting ....................... 49

Appendix A. Server racks and energy management ............. 51
  IBM server racks ........................................ 52
  IBM 42U SlimRack 7965-94Y ................................ 52
  IBM 42U SlimRack 7965-S42 ................................ 53
  IBM 7014 Model T42 rack ................................ 53
  The AC power distribution unit and rack content ............ 55
  Rack-mounting rules ...................................... 58
  OEM racks ................................................ 58
  Energy consumption estimation ............................ 58

Related publications ...................................... 59
  IBM Redbooks .............................................. 59
  Online resources .......................................... 59
  Help from IBM ............................................ 61
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®
- IBM®
- IBM Cloud™
- Interconnect®
- OpenCAPI™
- POWER®
- Power Architecture®
- POWER8®
- POWER9™
- PowerVM®
- Redbooks®
- Redbooks (logo)®
- RS/6000™
- System Storage™

The following terms are trademarks of other companies:

The registered trademark Linux® is used pursuant to a sublicense from the Linux Foundation, the exclusive licensee of Linus Torvalds, owner of the mark on a worldwide basis.

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

Red Hat, are trademarks or registered trademarks of Red Hat, Inc. or its subsidiaries in the United States and other countries.

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 System IC922 (9183-22X) server using IBM POWER9™ processor-based technology and supports Linux operating systems (OSes). The objective of this paper is to introduce the system offerings and their capacities and available features.

The IBM Power System IC922 server is built to deliver powerful computing, scaling efficiency, and storage capacity in a cost-optimized design to meet the evolving data challenges of the AI era, and include the following features:

- High throughput and performance for high-value Linux workloads, such as inferencing data or storage rich workloads, or cloud
- Potentially low acquisition cost through system optimization such as utilizing industry standard memory and warranty
- Two IBM POWER9 processor single-chip module (SCM) devices that provide high performance with 24, 32, or 40 fully activated cores and maximum 2 TB of memory
- Up to six NVIDIA T4 GPU accelerators
- Up to 24 2.5 inch SAS/SATA drives
- One dedicated and one shared 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)

This paper expands the current set of IBM Power Systems documentation by providing a desktop reference that offers a detailed technical description of the Power IC922 server.

Authors

This paper was produced by a team of specialists from around the world working at IBM Redbooks, Austin Center.

Stephen Lutz is a Certified Leading Technical Sales Professional for Power Systems working at IBM Germany. He holds a degree in Commercial Information Technology from the University of Applied Science Karlsruhe, Germany. He is an IBM POWER9™ champion and has 20 years of experience in AIX®, Linux, virtualization, and Power Systems and its predecessors. He provides pre-sales technical support to clients, IBM Business Partners, and IBM sales representatives in Germany.

Young Hoon Cho is Power Systems Top Gun with the post-sales Technical Support Team for IBM in Korea. He has over 10 years of experience working on IBM RS/6000®, IBM System p, and Power Systems products. He provides second-line technical support to Field Engineers working on Power Systems and system management.
The project that produced this publication was managed by:
Scott Vetter  
PMP, IBM

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

Ryan Achilles, Tonny Bastiaans, Kimber Carson, Gareth Coates, Martin Fullam, Nigel Griffiths, Jeffrey Jajowka, Ben Mashak, Brandon W. Pederson, Matthew Richardson, Alan Standridge, Justin Thaler, Eka Thirumalai, Kaushik Venkatesh, Alvin Wang  
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 IBM Redbooks® 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, IBM Redbooks
  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:

➤ Stay current on recent Redbooks publications with RSS Feeds:
http://www.redbooks.ibm.com/rss.html
Chapter 1. Architectural and technical description

The IBM Power System IC922 (9183-22X) server is built to deliver powerful computing, scaling efficiency, and storage capacity in a cost-optimized design to meet the evolving data challenges of the AI era.

The IC in IC922 stands for inference and cloud. The I can also stand for I/O.

The new IBM Power System IC922 (9183-22X) server is a powerful 2-socket server that offers 24, 32 or 40 fully activated cores and I/O configuration flexibility to help meet accelerated computing needs. The server features are designed for high-performance data analytics and high-performance computing workloads with up to six NVIDIA Tesla T4 GPU accelerators.
1.1 Power IC922 hardware overview

The IBM Power System IC922 server is engineered to put your Artificial Intelligence (AI) models to work and unlock business insights. It uses cooptimized hardware and software to deliver the necessary components for AI inference. The Power IC922 server fits into your AI workflow and integrates with the IBM Power System AC922 training platform. Thanks to its modular design, the Power IC922 server can scale with business needs.

The Power IC922 server supports two processor sockets, offering 12-core (2.8 to 3.8 GHz), 16-core (3.35 to 4.0 GHz), and 20-core (2.9 to 3.8 GHz) POWER9 technology-based system configurations in a 19-inch rack-mount with a 2U (EIA units) drawer configuration.

The Power IC922 server provides two hot-swap and redundant power supplies and 32 DIMM memory slots. 16 GB, 32 GB and 64 GB supported memory features allow for a maximum system memory of 2 TB.

The Power IC922 server also provides:
- High throughput and performance for high-value Linux workloads, such as inferencing data or storage rich workloads, or cloud
- Potentially low acquisition cost through system optimization such as utilizing industry standard memory and warranty
- Up to six NVIDIA T4 GPU accelerators
- Multiple I/O options in the system with the standard PCIe Riser, including:
  - Two PCIe x16 Gen 3 FHFL slots (supports double-wide accelerators)
  - Two PCIe x16 Gen 4 LP slots
  - Two PCIe x8 Gen 3 FHFL slots (physically x16)
  - Two PCIe x8 Gen 3 FHFL slots
  - Two PCIe x16 Gen 3 LP slots
- Up to 24 2.5 inch SAS/SATA drives
- Optional Mellanox InfiniBand or high-performance Ethernet
- Two rear USB 3.0 ports and one front USB 3.0 port
- 19-inch rack-mount with 2U configuration
- Red Hat Enterprise Linux 7.6 for IBM Power LE (POWER9) or later operating system

The Power IC922 server is shown in Figure 1-1
Chapter 1. Architectural and technical description

1.1.1 Statements of direction

- IBM currently intends to enable inference and training accelerator capabilities such as GPU, FPGAs, and ASICs with the new IBM Power System IC922 server. This flexible server design is currently intended to accommodate up to 8 accelerated devices, such as NVIDIA T4s or Xilinx U50. IBM's aim with the new Power System IC922 server is to help address the growing industry need for AI and enhance IBM's end-to-end, integrated AI platform, which includes IBM Watson® Machine Learning and select IBM POWER9 hardware.

- IBM currently intends to deliver the IBM Power System IC922 with significant I/O capabilities to meet the world's demanding data movement requirements. These I/O capabilities are currently intended to enable up to 24 NVMe drives and to deliver OpenCAPI™ technology, such as the Bittware 250-SoC card.

- IBM currently intends to deliver elastic, distributed inferencing capabilities as a new product within the Watson™ Machine Learning Accelerator product family. This product is intended to leverage the accelerator and IO capabilities of the IBM Power System IC922 server to support dynamic scaling for real-time AI inferencing across a cluster of optimized servers.

1.1.2 Minimum features

The minimum Power IC922 server initial order must include two processor modules, 64 GB of memory, two power supplies, two line cords, rack-mounting hardware, a system software indicator, a specified rack integrator, and a specified language group.

Table 1-1 shows the minimum configuration for the Power IC922 server.

<table>
<thead>
<tr>
<th>Feature code</th>
<th>Minimum</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>EK00 or EK01 or EK02</td>
<td>2</td>
<td>12-core 2.8 to 3.8 GHz (max) POWER9 Processor Module</td>
</tr>
<tr>
<td>EM62</td>
<td>4</td>
<td>16 GB DIMMs, 2666 MHz, 4 Gb DDR4 DRAM</td>
</tr>
</tbody>
</table>
### 1.1.3 Configurable feature matrix

Table 1-2 shows configurable feature matrix of the Power IC922 server.

<table>
<thead>
<tr>
<th>Feature code</th>
<th>Minimum</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>EC16</td>
<td>1</td>
<td>Open Power Abstraction Layer (OPAL)</td>
</tr>
<tr>
<td>4650</td>
<td>1</td>
<td>Rack Indicator -- Not Factory Integrated</td>
</tr>
<tr>
<td>EKMP</td>
<td>2</td>
<td>AC Power Supply, 2000 Watt (220 V)</td>
</tr>
<tr>
<td>Power cords</td>
<td>2</td>
<td>Select two power cords from supported list. Feature 4558 is defaulted. (FC 4558 Power Cord To PDU/UPS, (100-240V/16A))</td>
</tr>
<tr>
<td>2147</td>
<td>1</td>
<td>Primary OS Linux</td>
</tr>
<tr>
<td>93xx</td>
<td>1</td>
<td>Language Group Specify (select one from announced features)</td>
</tr>
</tbody>
</table>

**Note:** If you want a rack, it must be ordered as separate machine type and model in the initial system order. It will be shipped at the same time 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 at this time.

---

### 1.2 Operating system support

The Power IC922 server supports 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.

1.2.1 Red Hat Enterprise Linux

Red Hat Enterprise Linux (RHEL) (ppc64le) versions 7.6-alt for POWER9 is supported on the server with support for later distributions as they become available.

Note: There is also a RHEL Version 7.6 for POWER8®. Be sure to use the 7.6-alt version for the Power IC922 server.

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

1.3 Operating environment

Table 1-3 lists the electrical characteristics for the Power IC922 server.

Table 1-3 Electrical characteristics for Power IC922 server

<table>
<thead>
<tr>
<th>Electrical characteristics</th>
<th>Properties</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rated voltage and frequency&lt;sup&gt;a&lt;/sup&gt;</td>
<td>200 - 240 V AC at 50 or 60 Hz plus or minus 3 Hz</td>
</tr>
<tr>
<td>Thermal output (maximum)&lt;sup&gt;b&lt;/sup&gt;</td>
<td>6143 BTU/hr</td>
</tr>
<tr>
<td>Maximum power consumption&lt;sup&gt;b&lt;/sup&gt;</td>
<td>1800 W</td>
</tr>
<tr>
<td>Maximum kVA&lt;sup&gt;c&lt;/sup&gt;</td>
<td>1.855 kVA</td>
</tr>
<tr>
<td>Phase</td>
<td>Single</td>
</tr>
</tbody>
</table>

a. The power supplies automatically accept any voltage with the published, rated-voltage range. If dual power supplies are installed and operating, the power supplies draw approximately equal current from the utility (electrical supply) and provide approximately equal current to the load. Each power supply has an IEC 320-C20 inlet.

b. Power draw and heat load vary greatly by configuration. When you plan for an electrical system, it is important to use the maximum values.

c. To calculate the amperage, multiply the kVA by 1000 and divide that number by the operating voltage.

Power sockets: The Power IC922 server takes IEC 60320 C19/C20 mains power and not C13. Ensure that the correct power cords and PDUs are ordered or available in the rack.

Note: The Power IC922 server should always have both power supplies mounted and running. But in case one power supply failure, the system continues to operate at full performance, for example no CPU throttling takes place.

Table 1-4 lists the environment requirements for the Power IC922 server.
Table 1-4  Environment requirements for Power IC922 server

<table>
<thead>
<tr>
<th>Properties</th>
<th>Recommended operating</th>
<th>Allowable operating(^a\ b\ c)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ASHRAE class</td>
<td></td>
<td>A2 (Fourth edition)</td>
</tr>
<tr>
<td>Airflow direction</td>
<td>Front-to-back</td>
<td></td>
</tr>
<tr>
<td>Temperature</td>
<td>18.0°C – 27.0°C (64.4°F – 80.6°F)</td>
<td>10.0°C – 35.0°C (50.0°F – 95.0°F)</td>
</tr>
<tr>
<td>Low end moisture</td>
<td>-9.0°C (15.8°F) dew point</td>
<td>-12.0°C (10.4°F) dew point and 8% relative humidity</td>
</tr>
<tr>
<td>High end moisture</td>
<td>60% relative humidity and 15°C (59°F) dew point</td>
<td>85% relative humidity and 21.0°C (69.8°F) dew point</td>
</tr>
<tr>
<td>Maximum altitude</td>
<td></td>
<td>5000 m (16,400 ft)</td>
</tr>
</tbody>
</table>

a. Must derate the maximum allowable temperature 1°C (1.8°F) per 175 m (574 ft) up to a maximum allowable elevation of 3050 m (10000 ft).

b. The minimum humidity level is the larger absolute humidity of the -12°C (10.4°F) dew point and the 8% relative humidity. These levels intersect at approximately 25°C (77°F). Below this intersection, the dew point (-12°C) represents the minimum moisture level, while above it, the relative humidity (8%) is the minimum. For the upper moisture limit, the limit is the minimum absolute humidity of the dew point and relative humidity that is stated.

c. The following minimum requirements apply to data centers that are operated at low relative humidity:

Data centers that have do not have ESD floors and where people are allowed to wear non-ESD shoes might want to consider increasing humidity given that the risk of generating 8 kV increases slightly at 8% relative humidity, when compared to 25% relative humidity.

All mobile furnishings and equipment must be made of conductive or static dissipative materials and be bonded to ground.

During maintenance on any hardware, a properly functioning and grounded wrist strap must be used by any personnel who comes in contact with information technology (IT) equipment.

Note: The recommended operating environment is provided as the long-term operating environment that can result in the greatest reliability, energy efficiency, and reliability. The allowable operating environment represents where the equipment is tested to verify functionality. Due to the stresses that operating in the allowable envelope can place on the equipment, these envelopes must be used for short-term operation, not continuous operation.

1.4 Physical package

The Power IC922 server is offered exclusively as a rack-mounted 2U server. The width, depth, height, and weight of the server are shown in Table 1-5.

Table 1-5  Physical dimensions for the Power IC921 server

<table>
<thead>
<tr>
<th>Dimension</th>
<th>Power IC921 server</th>
</tr>
</thead>
<tbody>
<tr>
<td>Width</td>
<td>436 mm (17.2 in.)</td>
</tr>
<tr>
<td>Depth</td>
<td>820 mm (32.3 in.)</td>
</tr>
</tbody>
</table>
**1.5 System architecture**

In this section we show the detailed system architecture of Power IC922 server. Figure 1-3 illustrates the block diagram of the Power IC922 server using the standard PCIe riser cards FC EK30.

The Power IC922 server is a powerful 2-socket server that offers 24, 32, or 40 fully activated cores and I/O configuration flexibility to help meet accelerated computing needs. It is equipped with the latest DD2.3 POWER9 processor modules and supports eight lanes to the memory slots per processor with up to 32 DIMMs in total (up to 16 for each processor).
The system has 10 PCIe slots in different form factors and also bandwidths. Four low profile slots are mounted directly on the system board, with the PCIe cards plugged in vertically. Six additional slots are mounted on two riser cards with the PCIe cards plugged in horizontally. The feature code for the riser cards is FC EK30 and includes both riser cards required for the server.

In a system with the standard FC EK30 riser cards, the most performant slots 1 and 6 are directly connected to the processors and support a Gen 4, x16 bandwidth (32 GBps).

The slots 4 and 9 are also directly connected to the processor, but operate at Gen 3, x16 speed.

The six remaining slots are connected over a PCIe switches (PEX). Although the slots 2, and 7 are Gen3, x16 slots and also are connected, using a Gen 3, x16 connections between the slot and the PEX, the bandwidth is limited to Gen3 x8 (8 GBps). The reason for that is the connection from the PEX to the processor which has a bandwidth of Gen 3, x8.

The remaining slot 3, 5, 8, and 10 have a maximum throughput of a Gen 3, x8 connection up to 8 GBps.

**Note:** Although the slots 3 and 8 have an x8 performance, the slots on the board are x16 slots.

As the bandwidth between PEX and processors is limited to Gen 3, x8 keep this in mind for best performance to consult the PCIe adapter placement rules in the IBM Knowledge Center:


In an Power IC922 server using the FC EK30 riser card, the slots 1,4,6, and 9 are CAPI 2.0 enabled. Only these slots are CAPI 2.0 enabled as these are directly connected to the processor and no PEX is in between.

The system also provides four OpenCAPI 3.0 ports. These ports are on the system planar and could be used with OpenCAPI 3.0 capable cards. Such cards are plugged into an PCIe slot and are additionally cabled to the OpenCAPI 3.0 ports.

For the external ports, such as USB, Ethernet, RS232 and VGA a systems management chip is used to connect the ports to the processors. This chip also provides access to the open Baseboard Management Controller (openBMC). To access the openBMC use the right Ethernet port (rear view). For a redundant openBMC connection use the left Ethernet port (rear view). This port is a shared port with two MAC addresses that is able to provide both, access to the openBMC but could be also used as a standard Ethernet port inside the operating system. For more information about the Ethernet ports for the OpenBMC see Figure 1-16 on page 24.

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.

Figure 1-4 shows an overview of the system board and locations of the various parts.
1.6 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 IC922 server is the most current DD2.3 generation of the IBM POWER® processor family based on the LaGrange chip offered by the OpenPOWER foundation.

For more information about the LaGrange module or the other modules of the IBM POWER processor family, visit the following website:

https://www.ibm.com/systems/power/openpower/

1.6.1 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).

1.6.2 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 eight direct-attached memory channels for scale-out configurations. Each DDR4 channel supports up to 2666 Mbps for one DIMM per channel or 2400 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 42 lanes of PCIe Gen4 adapters at 16 Gbps.

The POWER9 processor consists of the following main components:

- Twenty 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 (SMT4).
- On-chip accelerators.
  - 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)).
  - 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.
- Two memory controllers that support direct-attached DDR4 memory:
  - Supports eight direct-attach memory buses.
  - Supports ×4 and ×8, 4 - 16 GB DRAMs and 3D stacked DRAMs.
- Processor SMP interconnect:
  - Supports two inter-node SMP X-bus links with 4B running at 16 Gbps. This allows a peak inter-node performance of 120 GBps.

Note: The inter-node SMP X-bus connection has a doubled performance compared to a Power LC922 server, as it has two X-busses not only one.

- Maximum two-socket SMP.
- Two 25G Link bricks with support for OpenCAPI 3.0 and NVIDIA NVLink 2.0 interconnect.
  OpenCAPI 3.0 allows a Field Programmable Gate Array (FPGA) or Application-Specific Integrated Circuit (ASIC) on a PCIe card to connect coherently to the POWER9 processor-based SMP using OpenCAPI cables from the PCIe card to a OpenCAPI 3.0 port on the system board. See Figure 1-4 for the location of the ports.
- Three PCIe Controllers (PECs) with 42 lanes of PCIe Gen4 I/O:
  - PEC0: one ×16 lanes.
– PEC1: one ×8 lanes and one ×2 lane (bifurcation).
– PEC2: one ×16 lanes, two ×8 lanes (bifurcation), or one ×8 lane and two ×4 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.

For more information on the LaGrange POWER9 single chip module refer to the POWER9 LaGrange Single-Chip Module data sheet:

https://ibm.ent.box.com/s/1dsn840viar4qb6a7ekzjb13b2xr1v91

Figure 1-5 shows a POWER9 processor with up to 24 cores (maximum available cores in the Power IC922 server is 20).

1.6.3 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
- Little Endian, and strong-ordering support extension
- A 51-bit real address and a 68-bit virtual address

1.6.4 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 1-6 shows a comparison between the different POWER9 processors options for the Power IC922 server and the number of threads that are supported by each SMT mode.

<table>
<thead>
<tr>
<th>Cores per system</th>
<th>SMT 0</th>
<th>SMT 2</th>
<th>SMT 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>24</td>
<td>24 HTpS(^a)</td>
<td>48 HTpS</td>
<td>96 HTpS</td>
</tr>
<tr>
<td>32</td>
<td>32 HTpS</td>
<td>64 HTpS</td>
<td>128 HTpS</td>
</tr>
<tr>
<td>40</td>
<td>40 HTpS</td>
<td>80 HTpS</td>
<td>160 HTpS</td>
</tr>
</tbody>
</table>

\(^a\) Hardware Threads per System.

1.6.5 Processor feature codes

The Power IC922 server supports three processor configurations, as shown in Table 1-7.

<table>
<thead>
<tr>
<th>Feature code</th>
<th>Description</th>
<th>Minimum</th>
<th>Maximum</th>
<th>OS support</th>
</tr>
</thead>
<tbody>
<tr>
<td>EK00</td>
<td>12-core 2.8 to 3.8 GHz (max) POWER9 Processor</td>
<td>2</td>
<td>2</td>
<td>Linux</td>
</tr>
<tr>
<td>EK01</td>
<td>16-core 3.35 to 4.0 GHz (max) POWER9 Processor</td>
<td>2</td>
<td>2</td>
<td>Linux</td>
</tr>
<tr>
<td>EK02</td>
<td>20-core 2.9 to 3.8 GHz (max) POWER9 Processor</td>
<td>2</td>
<td>2</td>
<td>Linux</td>
</tr>
</tbody>
</table>

The Power IC922 server supports only two processor modules, a one processor system is not available. No processor activation features are used or orderable on the Power IC922 server. All processor cores are always fully activated.
1.6.6 Coherent Accelerator Processor Interface 2.0

IBM CAPI 2.0 is the evolution of CAPI that defines a coherent accelerator interface structure for attaching special processing devices to the POWER9 processor bus. As with the original CAPI, CAPI 2.0 can attach accelerators that have coherent shared memory access with the processors in the server and share full virtual address translation with these processors by using standard PCIe Gen 4 buses with twice the bandwidth compared to the previous generation.

Applications can have customized functions in Field Programmable Gate Arrays (FPGAs) and queue work requests directly in shared memory queues to the FPGA. Applications can also have customized functions by using the same effective addresses (pointers) that they use for any threads running on a host processor. From a practical perspective, CAPI enables a specialized hardware accelerator to be seen as an extra processor in the system with access to the main system memory and coherent communication with other processors in the system.

Figure 1-6 shows a comparison of the traditional model, where the accelerator must go through the processor to access memory with CAPI.

The benefits of using CAPI include the ability to access shared memory blocks directly from the accelerator, perform memory transfers directly between the accelerator and processor cache, and reduce the code path length between the adapter and the processors. This reduction in the code path length might occur because the adapter is not operating as a traditional I/O device, and there is no device driver layer to perform processing. CAPI also presents a simpler programming model.

CAPI implements the POWER Service Layer (PSL), which provides address translation and system memory cache for the accelerator functions. The custom processors on the system board, consisting of an FPGA or an ASIC, use this layer to access shared memory regions, and cache areas as though they were a processor in the system. This ability enhances the performance of the data access for the device and simplifies the programming effort to use the device. Instead of treating the hardware accelerator as an I/O device, it is treated as a processor, which eliminates the requirement of a device driver to perform communication. It also eliminates the need for DMA that requires system calls to the OS kernel. By removing these layers, the data transfer operation requires fewer clock cycles in the processor, improving the I/O performance.
The implementation of CAPI on the POWER9 processor enables hardware companies to develop solutions for specific application demands. Companies use the performance of the POWER9 processor for general applications and the custom acceleration of specific functions by using a hardware accelerator with a simplified programming model and efficient communication with the processor and memory resources.

To use CAPI 2.0, the accelerator PCIe card must be mounted into a CAPI 2.0 enabled PCIe slot. In an Power IC922 server using the FC EK30 riser card, the slots 1, 4, 6, and 9 are CAPI 2.0 enabled.

### 1.6.7 OpenCAPI 3.0

Although CAPI is a technology that is present in IBM POWER processors and depends on IBM intellectual property (the PSL), several industry solutions might benefit from having a mechanism of connecting different devices to the processor with low latency, including memory attachment. The PCIe standard is pervasive to every processor technology, but its design characteristics and latency do not enable the attachment of memory for load/store operations.

Therefore, the OpenCAPI Consortium was created, with the goal of defining a device attachment interface to open the CAPI interface to other hardware developers and extending its capabilities. OpenCAPI aims to enable memory, accelerators, network, storage, and other devices to connect to the processor through a high-bandwidth, low-latency interface, becoming the interface of choice for connecting high-performance devices.
By providing a high-bandwidth, low-latency connection to devices, OpenCAPI enables several applications to improve networking, use FPGA accelerators, use expanded memory beyond server internal capacity, and reduce latency to storage devices. Some of these use cases and examples are shown in Figure 1-7.

The design of OpenCAPI enables low latency in accessing attached devices (in the same range of DDR memory access, that is, 10 ns), which enables memory to be connected through OpenCAPI and serve as main memory for load/store operations. In contrast, PCIe latency is 10 times larger (around 100 ns). Therefore, OpenCAPI has a significant enhancement compared to traditional PCIe interconnects.

OpenCAPI is neutral regarding processor architecture, so the electrical interface is not defined by the OpenCAPI consortium or any of its workgroups. On the POWER9 processor, the electrical interface is based on the design from the 25G workgroup within the OpenPower Foundation, which encompasses a 25 GBps signaling and protocol that is built to enable a low-latency interface on CPU and attached devices.
The current design for the adapter is based on a PCIe card that draws power from the PCIe slot while connecting to the OpenCAPI port on the system board through a 25-GBps cable, as shown in Figure 1-8.

![OpenCAPI compatible adapter and 25G link](image)

The OpenCAPI interface uses the same electrical interconnect as NVLink 2.0. Systems can be designed to have an NVLink-attached GPU, an OpenCAPI-attached device, or both. For the Power IC922 server NVLink is not supported, therefore these interconnects could only be used for OpenCAPI 3.0 solutions.

Figure 1-9 shows the four OpenCAPI 3.0 ports named P1T1, P1T2, P1T3, and P1T4.

![OpenCAPI 3.0 ports in the Power IC922 server](image)

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

### 1.7 Memory subsystem

The POWER9 memory subsystem consists of a POWER9 processor that supports eight channels. Each memory port or channel supports single- or dual-drop industry-standard RDIMMs, therefore each channel can accommodate one or two DIMMs. The POWER9 processor connects to the DIMMs by using the standard DDR4 memory interface.

**Note:** Although the DIMMs are industry-standard RDIMMs, only DIMMs sold by IBM are supported in the system.
Figure 1-10 shows the memory slot location of the Power IC922 server.

The corresponding slot location codes of the slots in Figure 1-10 are as shown in Table 1-8 and Table 1-9:

**Table 1-8  DIMM location codes left hand side**

<table>
<thead>
<tr>
<th>Memory DIMM slot</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
<th>15</th>
</tr>
</thead>
<tbody>
<tr>
<td>Location code</td>
<td>P1-C35</td>
<td>P1-C34</td>
<td>P1-C33</td>
<td>P1-C32</td>
<td>P1-C31</td>
<td>P1-C30</td>
<td>P1-C29</td>
<td>P1-C28</td>
<td>P1-C27</td>
<td>P1-C26</td>
<td>P1-C25</td>
<td>P1-C24</td>
<td>P1-C23</td>
<td>P1-C22</td>
<td>P1-C21</td>
<td>P1-C20</td>
</tr>
</tbody>
</table>

**Table 1-9  DIMM location codes right hand side**

<table>
<thead>
<tr>
<th>Memory DIMM slot</th>
<th>16</th>
<th>17</th>
<th>18</th>
<th>19</th>
<th>20</th>
<th>21</th>
<th>22</th>
<th>23</th>
<th>24</th>
<th>25</th>
<th>26</th>
<th>27</th>
<th>28</th>
<th>29</th>
<th>30</th>
<th>31</th>
</tr>
</thead>
<tbody>
<tr>
<td>Location code</td>
<td>P1-C19</td>
<td>P1-C18</td>
<td>P1-C17</td>
<td>P1-C16</td>
<td>P1-C15</td>
<td>P1-C14</td>
<td>P1-C13</td>
<td>P1-C12</td>
<td>P1-C11</td>
<td>P1-C10</td>
<td>P1-C9</td>
<td>P1-C8</td>
<td>P1-C7</td>
<td>P1-C6</td>
<td>P1-C5</td>
<td>P1-C4</td>
</tr>
</tbody>
</table>

Memory DIMMs in the Power IC922 server could be installed in increments of four up to 32 DIMMs per system.

**Note:** A 20 DIMMs and a 28 DIMMs configuration is not supported.

All DIMMs in a system must be of the same size. One feature code shown in Table 1-10 provides one DIMM, therefore the feature codes need also to be ordered in increments of four.
Table 1-10 Memory feature codes for the Power IC922 server

<table>
<thead>
<tr>
<th>Feature code</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>EM62</td>
<td>16 GB 1RX4 2666 MHz DDR4 RDIMM</td>
</tr>
<tr>
<td>EM63</td>
<td>32 GB 2RX4 2666 MHz DDR4 RDIMM</td>
</tr>
<tr>
<td>EM64</td>
<td>64 GB 4RX4 2666 MHz DDR4 2H 3DS RDIMM</td>
</tr>
</tbody>
</table>

Table 1-11 shows the possible memory configurations for the Power IC922 server.

Table 1-11 Possible memory configurations

<table>
<thead>
<tr>
<th>Feature code</th>
<th>Size</th>
<th>4 DIMMs</th>
<th>8 DIMMs</th>
<th>12 DIMMs</th>
<th>16 DIMMs</th>
<th>24 DIMMs</th>
<th>32 DIMMs</th>
</tr>
</thead>
<tbody>
<tr>
<td>EM62</td>
<td>16 GB</td>
<td>64 GB</td>
<td>128 GB</td>
<td>192 GB</td>
<td>256 GB</td>
<td>384 GB</td>
<td>512 GB</td>
</tr>
<tr>
<td>EM63</td>
<td>32 GB</td>
<td>128 GB</td>
<td>256 GB</td>
<td>384 GB</td>
<td>512 GB</td>
<td>768 GB</td>
<td>1024 GB</td>
</tr>
<tr>
<td>EM64</td>
<td>64 GB</td>
<td>256 GB</td>
<td>512 GB</td>
<td>768 GB</td>
<td>1024 GB</td>
<td>1536 GB</td>
<td>2048 GB</td>
</tr>
</tbody>
</table>

Recommendation: Populate 16 slots for a maximum bandwidth or 32 slots for a maximum memory size of 2 TB.

1.7.1 DIMM placement rules

In the system the DIMMs slots populated first are the white plastic colored slots (single-drop). The remaining slots are in black plastic colored. This general rule is also followed in the placement rules for the system. The Power IC922 server has specific placement rules that are shown in Figure 1-11.

Figure 1-11 DIMM placement rules for Power IC922 server

Also see Figure 1-10 to identify the correct slot or have a look into the IBM Knowledge Center:
1.7.2 Memory bandwidth

The system is memory performance optimized with 16 DIMMs populated, because in this configuration all memory DIMMs work in single-drop mode at a speed of 2666 Mbps. If in an memory channel both DIMMs are populated (dual-drop) the speed is reduced to 2400 Mbps.

The maximum bandwidth of the system is dependent of the number of channels used and if the channel operates in single-drop or dual-drop mode. A DDR4 DIMM clocked with 2666 MHz has a maximum transfer rate of 21.3 GB/s. A DDR4 DIMM clocked with 2400 MHz has a maximum transfer rate of 19.2 GB/s. To calculate the maximum bandwidth you may use the following formula:

Maximum Bandwidth = (Number of DIMMs in single-drop mode * 21.3 GBps) + ((Number of DIMMs in dual-drop mode / 2) * 19.2 GBps)

Table 1-12 shows the maximum memory bandwidth for the Power IC922 server dependent on the number of DIMMs populated.

<table>
<thead>
<tr>
<th>DIMM quantity</th>
<th>Maximum Bandwidth in GBps</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>85</td>
</tr>
<tr>
<td>8</td>
<td>170</td>
</tr>
<tr>
<td>12</td>
<td>255</td>
</tr>
<tr>
<td>16</td>
<td>340</td>
</tr>
<tr>
<td>24</td>
<td>323&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td>32</td>
<td>306&lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
</tbody>
</table>

<sup>a</sup> Processor 0 operates at 2400 Mbps  
<sup>b</sup> Both processors operate at 2400 Mbps

1.8 The Trusted Platform Module

The Power IC922 server is equipped with a Trusted Platform Module (TPM). The TPM card is mounted on the system planar as shown in Figure 1-12. The TPM module is able to sign the hardware and firmware by using cryptographic keys. This is the basis for a secure boot of the operating system.

An operating system that is enabled for secure boot makes sure that the hardware and firmware were not modified and may prevent the operating system from booting or at least informs the system administrator if something was modified. The trusted hardware and firmware together with the operating system that securely boots is called a trusted platform.
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 PCIe Controller (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.

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.

The Power IC922 server provides a total of 10 PCIe slots. Slots 1, 2, 6, and 7 are part of the system planar. The rest of PCIe slots provided by PCIe Riser. The server requires one PCIe riser feature (providing two PCIe riser cards).

The Power IC922 can support three different form factors of PCIe adapters:

- PCIe Low Profile (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.

1.9.2 Slot configuration

The Power IC922 server provides PCIe Gen 3 and PCIe Gen 4 slots.

Figure 1-13 shows the rear view of the system with the PCIe adapter slots.

![Figure 1-13 Rear view of the Power IC922 server with PCIe slots indicated](image)

Slots 1, 2, 6, and 7 are part of the system planar and as such, do not depend on the PCIe Riser selected. As part of the system planar, these slots (1, 2, 6 and 7) only support low profile (LP) adapters that are vertically mounted.

The slots provided by PCIe riser are horizontal slots and the PCIe adapters are therefore horizontally mounted. The slots 3, 4, 8, and 9 accept FHFL (full height full length) and FHHL (full height half length) PCIe adapters. The slots 5 and 10 only accept FHHL PCIe adapters.

Table 1-13 shows available PCIe riser for the Power IC922 server.

<table>
<thead>
<tr>
<th>Feature code</th>
<th>CCIN</th>
<th>Description</th>
<th>Min</th>
<th>Max</th>
<th>OS support</th>
</tr>
</thead>
<tbody>
<tr>
<td>EK30</td>
<td>PCI Risers Standard</td>
<td>0</td>
<td>1</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 1-14 lists the PCIe adapter slots and details for the system. The Power IC922 server supports full-height, full-length, and half-length PCIe adapters.

**Table 1-14  PCIe slots and descriptions for the Power IC922 server with FC EK30 Riser**

<table>
<thead>
<tr>
<th>Slot identification</th>
<th>Description</th>
<th>Adapter size</th>
<th>Processor module</th>
<th>Coherent Accelerator Processor Interface (CAPI) 2.0</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>PCIe Gen 4 x16</td>
<td>Low Profile (LP)</td>
<td>CPU 0</td>
<td>Yes</td>
</tr>
<tr>
<td>2</td>
<td>PCIe Gen 3 x16</td>
<td>Low Profile (LP)</td>
<td>CPU 0</td>
<td>No</td>
</tr>
<tr>
<td>3</td>
<td>PCIe Gen 3 x16</td>
<td>Full height, Full length</td>
<td>CPU 0</td>
<td>No</td>
</tr>
<tr>
<td>4</td>
<td>PCIe Gen 3 x16</td>
<td>Full height, Full lengtha</td>
<td>CPU 0</td>
<td>Yes</td>
</tr>
<tr>
<td>5</td>
<td>PCIe Gen 3 x8</td>
<td>Full height, Full length</td>
<td>CPU 0</td>
<td>No</td>
</tr>
<tr>
<td>6</td>
<td>PCIe Gen 4 x16</td>
<td>Low Profile (LP)</td>
<td>CPU 1</td>
<td>Yes</td>
</tr>
<tr>
<td>7</td>
<td>PCIe Gen 3 x16</td>
<td>Low Profile (LP)</td>
<td>CPU 1</td>
<td>No</td>
</tr>
<tr>
<td>8</td>
<td>PCIe Gen 3 x16</td>
<td>Full height, Full length</td>
<td>CPU 1</td>
<td>No</td>
</tr>
<tr>
<td>9</td>
<td>PCIe Gen 3 x16</td>
<td>Full height, Full lengtha</td>
<td>CPU 1</td>
<td>Yes</td>
</tr>
<tr>
<td>10</td>
<td>PCIe Gen 3 x8</td>
<td>Full height, Half length</td>
<td>CPU 1</td>
<td>No</td>
</tr>
</tbody>
</table>

*a. Supports double-wide adapters.*
1.9.3 GPU

Table 1-15 lists the GPUs that are available for the Power IC922 server.

<table>
<thead>
<tr>
<th>Feature code</th>
<th>CCIN</th>
<th>Description</th>
<th>Min</th>
<th>Max</th>
<th>OS support</th>
</tr>
</thead>
<tbody>
<tr>
<td>EK4L</td>
<td>NVIDIA T4 GPU with PCIe Gen 3 x16 LP (16 GB)</td>
<td>0 4 Linux</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>EK4M</td>
<td>NVIDIA T4 GPU with PCIe Gen 3 x16 (16 GB)</td>
<td>0 2 Linux</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The NVIDIA T4 PCIe GPU accelerator adapter is a single slot 6.6-inch PCI Express (PCIe) generation 3 (Gen 3) universal deep learning accelerator based on the TU104 NVIDIA graphics processing unit (GPU). The adapter can be used in either an x8 or x16 PCIe slot in the system. The NVIDIA T4 adapter has 16 GB GDDR6 memory and a maximum power limit of 70 W. The FC EK4L is a short, low-profile adapter. The FC EK4M is a short adapter with a full-height tail stock.

The NVIDIA website provide detailed information about T4 GPUs:

The Power IC922 server supports up to six NVIDIA T4 GPU.

Figure 1-15 shows T4 GPU plug order. The order shown by yellow numbers and PCIe Gen 3 x8 slots outlined in red.

1.9.4 FPGA

There is no available OpenCAPI FPGA at the time of writing.

1.9.5 PCI adapters

This section describes the types and functions of the PCI adapters that are supported by the Power IC922 server.

Local area network adapters
To connect the Power IC922 servers to a LAN, you can use the LAN adapters that are supported in the PCIe slots of the system in addition to the two 1GE RJ45 ports that are present in every system.
Figure 1-16 shows the two 1GE RJ45 ports that are integrated into the system.

Figure 1-16  Integrated LAN port

![Figure 1-16: Integrated LAN port]

Table 1-16 lists the LAN adapters that are available for the Power IC922 server.

Table 1-16  Available LAN adapters for the Power IC922 server

<table>
<thead>
<tr>
<th>Feature code</th>
<th>CCIN</th>
<th>Description</th>
<th>Min</th>
<th>Max</th>
<th>OS support</th>
</tr>
</thead>
<tbody>
<tr>
<td>EC2R</td>
<td>58FA</td>
<td>PCIe3 LP 2-Port 10 Gb NIC&amp;ROCE SR/Cu Adapter</td>
<td>0</td>
<td>4</td>
<td>Linux</td>
</tr>
<tr>
<td>EC2S</td>
<td>58FA</td>
<td>PCIe3 2-Port 10 Gb NIC&amp;ROCE SR/Cu Adapter</td>
<td>0</td>
<td>6</td>
<td>Linux</td>
</tr>
<tr>
<td>EC2T</td>
<td>58FB</td>
<td>PCIe3 LP 2-Port 25/10 Gb NIC&amp;ROCE SR/Cu Adapter</td>
<td>0</td>
<td>4</td>
<td>Linux</td>
</tr>
<tr>
<td>EC2U</td>
<td>58FB</td>
<td>PCIe3 2-Port 25/10 Gb NIC&amp;ROCE SR/Cu Adapter</td>
<td>0</td>
<td>6</td>
<td>Linux</td>
</tr>
<tr>
<td>EC62</td>
<td>2CF1</td>
<td>PCIe4 LP 1-port 100 Gb EDR IB CAPI adapter</td>
<td>0</td>
<td>2</td>
<td>Linux</td>
</tr>
<tr>
<td>EC64</td>
<td>2CF2</td>
<td>PCIe4 LP 2-port 100 Gb EDR IB CAPI adapter</td>
<td>0</td>
<td>2</td>
<td>Linux</td>
</tr>
<tr>
<td>EC67</td>
<td>2CF3</td>
<td>PCIe4 LP 2-port 100 Gb ROCE EN LP adapter</td>
<td>0</td>
<td>2</td>
<td>Linux</td>
</tr>
<tr>
<td>EL3Z</td>
<td>2CC4</td>
<td>PCIe2 LP 2-port 10/1 GbE BaseT RJ45 Adapter</td>
<td>0</td>
<td>4</td>
<td>Linux</td>
</tr>
<tr>
<td>EL4L</td>
<td>576F</td>
<td>PCIe2 4-port 1 GbE Adapter</td>
<td>0</td>
<td>6</td>
<td>Linux</td>
</tr>
<tr>
<td>EL4M</td>
<td>576F</td>
<td>PCIe2 LP 4-port 1 GbE Adapter</td>
<td>0</td>
<td>4</td>
<td>Linux</td>
</tr>
<tr>
<td>EL55</td>
<td>2CC4</td>
<td>PCIe2 2-port 10/1 GbE BaseT RJ45 Adapter</td>
<td>0</td>
<td>6</td>
<td>Linux</td>
</tr>
<tr>
<td>EN0S</td>
<td>2CC3</td>
<td>PCIe2 4-Port (10 Gb+1 GbE) SR+RJ45 Adapter</td>
<td>0</td>
<td>6</td>
<td>Linux</td>
</tr>
</tbody>
</table>
Fibre Channel adapters
The servers support a direct or SAN connection to devices that use Fibre Channel connectivity.

Table 1-17 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>
<thead>
<tr>
<th>Feature code</th>
<th>CCIN</th>
<th>Description</th>
<th>Min</th>
<th>Max</th>
<th>OS support</th>
</tr>
</thead>
<tbody>
<tr>
<td>EN0T</td>
<td>2CC3</td>
<td>PCIe2 LP 4-Port (10 Gb+1 GbE) SR+RJ45 Adapter</td>
<td>0</td>
<td>4</td>
<td>Linux</td>
</tr>
<tr>
<td>EN0U</td>
<td>2CC3</td>
<td>PCIe2 4-port (10 Gb+1 GbE) Copper SFP+RJ45 Adapter</td>
<td>0</td>
<td>6</td>
<td>Linux</td>
</tr>
<tr>
<td>EN0V</td>
<td>2CC3</td>
<td>PCIe2 LP 4-port (10 Gb+1 GbE) Copper SFP+RJ45 Adapter</td>
<td>0</td>
<td>4</td>
<td>Linux</td>
</tr>
</tbody>
</table>

SAS/SATA host bus adapters
Table 1-18 lists the SAS/SATA Host Bus Adapters (HBAs) that are available for the Power IC922 server. These adapters are used to connect internally from the rear of the system to the three drive backplanes in front of the system as discussed in Chapter 1.10, “Internal storage” on page 26.

<table>
<thead>
<tr>
<th>Feature code</th>
<th>CCIN</th>
<th>Description</th>
<th>Min</th>
<th>Max</th>
<th>OS support</th>
</tr>
</thead>
<tbody>
<tr>
<td>EK41</td>
<td></td>
<td>Broadcom 9300-8i SAS/SATA HBA PCIe3 x8 LP</td>
<td>0</td>
<td>2</td>
<td>Linux</td>
</tr>
<tr>
<td>EK43</td>
<td></td>
<td>Broadcom 9305-16i SAS/SATA HBA PCIe3 x8 LP</td>
<td>0</td>
<td>1</td>
<td>Linux</td>
</tr>
<tr>
<td>EK47</td>
<td></td>
<td>Broadcom (LSI) MegaRAID 9361-8i SAS3 Controller with 8 internal ports (2 GB Cache) PCIe 3.0 x8 with cables</td>
<td>0</td>
<td>2</td>
<td>Linux</td>
</tr>
</tbody>
</table>
1.10 Internal storage

The Power IC922 server supports up to 24 2.5-inch SAS/SATA HDD/SSD. It does this through three front-mounted backplanes. Each backplane has 8 drive slots.

The Power IC922 server does not have an integrated disk controller. You need to install SAS/SATA HBAs depending on how many disk backplanes are installed. The Broadcom MegaRAID 9361-8i HBA (FC EK47) supports RAID levels 0, 1, 5, 6, 10, 50, and 60, and Just a Bunch of Disks (JBOD). The other HBAs only support JBOD.

There are various internal storage configurations for the server:

- One FC EK41 or FC EK47 HBA in PCIe slot 5 and one drive backplane 2.5-inch SAS/SATA (FC EK36)
  
  You can install a maximum up to eight 2.5" SAS/SATA disks.

- One FC EK41 or FC EK47 HBA in PCIe slot 5, one FC EK42 in PCIe slot 10, and three drive backplanes 2.5-inch SAS/SATA (FC EK36)
  
  You can install a maximum up to twenty-four 2.5" SAS/SATA disks.
1.10.1 Drive backplanes

Up to three drive backplanes are supported in the Power IC922 server. Table 1-19 shows the available drive backplanes.

<table>
<thead>
<tr>
<th>Feature code</th>
<th>CCIN</th>
<th>Description</th>
<th>Min.</th>
<th>Max</th>
<th>OS support</th>
</tr>
</thead>
<tbody>
<tr>
<td>EK36</td>
<td></td>
<td>Drive backplane 2.5-inch SAS/SATA, 8x bays, 8x carriers, 12V Power cable</td>
<td>0</td>
<td>3</td>
<td>Linux</td>
</tr>
</tbody>
</table>

Figure 1-19 shows the location of the three front backplanes.

A backplane filler is required for each drive backplane that is not installed. The backplane fillers will be installed in the factory automatically, there is no need to order them separately.
1.10.2 HDD and SSD

The Power IC922 server supports hard disk drives (HDDs) or Solid-State-Drives (SSDs) or a mixture of HDDs and SSDs in 2.5-inch SAS/SATA drive backplanes. Mixing HDDs and SSDs applies even within a single disk backplane.

Table 1-20 lists the HDDs and SSDs that are available for the Power IC922 server.

Table 1-20 Available HDDs and SDDs for the Power IC922 server

<table>
<thead>
<tr>
<th>Feature code</th>
<th>CCIN</th>
<th>Description</th>
<th>Min.</th>
<th>Max</th>
<th>OS support</th>
</tr>
</thead>
<tbody>
<tr>
<td>EK27</td>
<td></td>
<td>3.8 TB 2.5-inch SAS3 1-DWPD NONSED(^a)</td>
<td>0</td>
<td>24</td>
<td>Linux</td>
</tr>
<tr>
<td>EK20</td>
<td></td>
<td>240 GB 2.5-inch SATA SSD 1.4-DWPD(^b) NONSED(^a)</td>
<td>0</td>
<td>24</td>
<td>Linux</td>
</tr>
<tr>
<td>EK21</td>
<td></td>
<td>960 GB 2.5-inch SATA SSD 2.5-DWPD(^b) NONSED(^a)</td>
<td>0</td>
<td>24</td>
<td>Linux</td>
</tr>
<tr>
<td>EK22</td>
<td></td>
<td>1.9 TB 2.5-inch SATA SSD 2.5-DWPD(^b) NONSED(^a)</td>
<td>0</td>
<td>24</td>
<td>Linux</td>
</tr>
<tr>
<td>EK23</td>
<td></td>
<td>3.8 TB 2.5-inch SATA SSD 2.5-DWPD(^b) NONSED(^a)</td>
<td>0</td>
<td>24</td>
<td>Linux</td>
</tr>
<tr>
<td>EK1A</td>
<td></td>
<td>600 GB 2.5-inch SAS3 HHD 4KN/512e SED</td>
<td>0</td>
<td>24</td>
<td>Linux</td>
</tr>
<tr>
<td>EK1B</td>
<td></td>
<td>1.2 TB 2.5-inch SAS3 HDD 4KN/512e SED</td>
<td>0</td>
<td>24</td>
<td>Linux</td>
</tr>
<tr>
<td>EK1C</td>
<td></td>
<td>2.4 TB 2.5-inch SAS3 HDD 4KN/512e NONSED(^a)</td>
<td>0</td>
<td>24</td>
<td>Linux</td>
</tr>
</tbody>
</table>

\(^a\) NONSED = Non Self Encrypting Drive
\(^b\) DWPD = Whole Drive Writes Per Day

Note: The Power IC922 server is allowed to be ordered with zero disk features. Each FC EK36 can contain up to eight HDDs or SSDs.

1.10.3 NVMe

NVM Express (NVMe) is a specification defining how host software communicates with non-volatile memory across a PCIe bus. It is the industry standard for PCIe SSDs in all form factors (U.2, M.2, AIC, EDSFF).

There are many form factors of NVMe Solid State Drive for IBM Power systems as follows:

- Add-in Card (ASIC)
  - x4 / x8 PCIe Gen3/4
  - High Performance and Power (25-50 Watts)
  - Focus on High capacity
  - Best performance and lowest latency

- 2.5" SFF (U.2)
  - 2x2 / 1x4 PCIe Gen3/4, Dual port support, 8639 Connector
  - Split into two power offerings: 15-25W (15mm), 9-12W (7mm)
– Expect to be direct attached to CPU vs through HBA

- M.2
- x2 / x4 PCIe Gen3, Size: 22x30, 22x42, 22x60, 22x80, 22x110mm
- Power less than 9 watts, un-encapsulated, No hot swap

Any NVMe disk isn't supported for the Power IC922 server at the time of writing.

1.11 System ports

The rear of the Power IC922 server has a VGA port, a serial port (RS232), an BMC port, two USB 3.0 ports, and two 1 Gbps RJ45 ports. The front of the server has a USB 3.0 port. They are all integrated into the system board. The server supports redundant power supplies.

Figure 1-20 shows the system ports in the rear view of the Power IC922 server.

1.12 External I/O subsystems

The Power IC922 server does 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.
Chapter 2. Systems management

The IBM Power System IC922 server is 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 Service processor

The Power System IC922 server among others use as a service processor a baseboard management controller (BMC) for system service management, monitoring, maintenance, and control. The BMC is a specialized service processor that monitors the physical state of the system by using sensors. The BMC also provides access to the system event log files (SEL). A system administrator or service representative can communicate with the BMC through an independent connection.

The BMC is provided by the OpenBMC project as a part of the Linux Foundation.

There are several ways to work with the BMC as discussed in the following sections.

**Note:** Do not work with the openBMC when the system is under heavy load.

### 2.1.1 Intelligent Platform Management Interface

The Intelligent Platform Management Interface (IPMI) is an open standard for monitoring, logging, recovery, inventory, and control of hardware that is implemented independently of the main processor, I/O, and the operating system. It is the default console to use when you configure the server.

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 RedHat, run the following command:

```bash
# yum install ipmitool
```

**Note:** There is also an ipmitool available for Microsoft Windows.

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:
   ```bash
   # ipmitool -I lanplus -H fsp_ip_address -U <user> -P <password> power on
   ```
3. Activate your IPMI console by running the following command:
   ```bash
   # ipmitool -I lanplus -H fsp_ip_address -U <user> -P <password> sol activate
   ```

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

### 2.1.2 Petitboot bootloader

Petitboot is a kexec-based bootloader that is able to start an operating system. 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.

### 2.1.3 Configuring the network to the BMC

In this section you can learn how to find the right network port for the BMC, how to set the initial password and how to do an initial network configuration.

#### Initial Password change

When the system is in factory default, the default password `openBmc` for root is invalidated for security reasons. You first need to change/set the password. This can be done by login into the OpenBMC GUI or SSH, by using the ipmitool in a local session with local keyboard and screen, and by using the REST API. You can not use a remote ipmitool to change the password, as the necessary password to enter the BMC is invalidated and a change password dialog is not available using the remote ipmitool.

When using the GUI or SSH a change password dialog will appear.

In a local session you can use the ipmitool to change the password. To change it, you need the ID for the user. For user root this is usually 1. To make sure that user root is ID 1, the following command can be used:

```
# ipmitool user list 1
```

**Note:** The 1 stands for the channel number. There are two channel numbers (1 and 2). Channel number 1 is used for the dedicated BMC port of the system (the right Ethernet port). Channel number 2 is used for the shared BMC port (the left Ethernet port). By differentiating 1 and 2 it is possible to define different access rights on the two ports. But the password is the same for both channels, therefore you do not need to specify it when changing the password.

To change the password, use the following command:

```
# ipmitool user set password <ID of the user> <password>
```

The requirement for a password are 8 to 20 characters and the password will be checked against pam_cracklib. Passwords must not confirm to dictionary words, palindrome etc. For example “abcd1234” is not an allowed password.

Refer to pam_cracklib for more details on allowed password:

[https://linux.die.net/man/8/pam_cracklib](https://linux.die.net/man/8/pam_cracklib)

#### Network ports for accessing the BMC interface

The Power IC922 server has three 1 GB network ports with RJ-45 connector. Figure 1-16 on page 24 shows these ports on the backside of the system. For accessing the BMC two ports
could be used. The right-hand port is the dedicated BMC port. The left-hand port is a shared port that could be used to create a redundant connection to the BMC, but also as a network port for the operating system. To make that possible, this port has two MAC addresses, one is available in the BMC environment and one in the operating system. The port in the middle is for usage in the operating system. Table 2-1 gives an overview about the ports and their usage.

\[
\text{Table 2-1} \quad \text{Network ports and their usage}
\]

<table>
<thead>
<tr>
<th>Port location</th>
<th>MAC example</th>
<th>Port name</th>
<th>Accessible in</th>
<th>Usage</th>
</tr>
</thead>
<tbody>
<tr>
<td>left</td>
<td>08:..........:92</td>
<td>eth1</td>
<td>BMC</td>
<td>Shared BMC port</td>
</tr>
<tr>
<td>left</td>
<td>08:..........:93</td>
<td>eth0</td>
<td>OS</td>
<td>NIC port 0</td>
</tr>
<tr>
<td>middle</td>
<td>08:..........:94</td>
<td>eth1</td>
<td>OS</td>
<td>NIC port 1</td>
</tr>
<tr>
<td>right</td>
<td>08:..........:91</td>
<td>eth0</td>
<td>BMC</td>
<td>Dedicated BMC port</td>
</tr>
</tbody>
</table>

**Initial configuration of the BMC network ports**

There are two ways to initially configure the network ports to access the BMC:

1. **DHCP**

   A system in factory default, both BMC network ports are configured with DHCP. When you plug-in the network cable and your network provides a DHCP server an IP address will be assigned to the network port. When you can find out the IP address that was assigned, you can use that one for an initial connect to the BMC. For example use the OpenBMC GUI to set the network parameters of the management ports. In section 2.1.5, “The OpenBMC GUI” you can read how to connect to the GUI.

   \[\text{Note: If the system is in factory default, the password for root is invalidated. You first need to set the password. This can not be done using a remote ipmitool. Use the OpenBMC GUI, a local attached keyboard and screen, SSH or the REST API to set the password.}\]

2. **Attach VGA monitor and keyboard**

   If you do not have a DHCP server, the easiest way to configure the BMC network is to attach a local VGA monitor and a USB keyboard. When the system starts up it will enter the Petitboot loader. There you can exit to a shell as shown in Figure 2-1.
In the Petitboot loader you can use the `ipmitool` command to configure the network as follows.

- **To configure the *dedicated* management port with a static IP address do:**
  ```
  # ipmitool lan set 2 ipsrc static
  # ipmitool lan set 2 ipaddr <Your IP address for the BMC>
  # ipmitool lan set 2 netmask <Your Subnet Mask>
  # ipmitool lan set 2 defgw ipaddr <Your gateway server>
  ```

- **To set the port back to DHCP do:**
  ```
  # ipmitool lan set 2 ipsrc dhcp
  ```

- **To view the current configuration do:**
  ```
  # ipmitool lan print 2
  ```

- **To configure the *shared* management port with a static IP address do:**
  ```
  # ipmitool lan set 1 ipsrc static
  # ipmitool lan set 1 ipaddr <Your IP address for the BMC>
  # ipmitool lan set 1 netmask <Your Subnet Mask>
  # ipmitool lan set 1 defgw ipaddr <Your gateway server>
  ```

- **To set the port back to DHCP do:**
  ```
  # ipmitool lan set 1 ipsrc dhcp
  ```

- **To view the current configuration do:**
  ```
  # ipmitool lan print 1
  ```

**Note:** It is also possible to use a console or a terminal program attached to the serial port in order to access the system monitor. If you want to do this, try at 115200 baud.
2.1.4 Systems management using the OpenBMC tool

The OpenBMC tool provides a communication method to the BMC, by using a command-line interface. The OpenBMC tool can be used either from a remote system, or from the host operating system console window.

**Note:** The minimum required version for the OpenBMC tool is 1.17 (recommended 1.18). The new version is also compatible with other IBM systems that support the OpenBMC tool.

You may download the OpenBMC tool from the following website:


To run the OpenBMC tool, you need the following pre-requisites installed on the system:
- Python 3
- Python 3 requests

**Note:** When you installed python 3 and you are looking for python 3 request you may enter the following command to get it installed:

```
# python -m pip install requests
```

To use the OpenBMC tool enter the connection data and the credentials for the system and a command. The following examples shows how to see the power status and to power on/off the system:

- To check the power status of the system, run the following command:
  
  ```
  # openbmctool -U <username> -P <password> -H <BMC IP address or BMC host name> chassis power status
  ```

- To power on the system, run the following command:
  
  ```
  # openbmctool -U <username> -P <password> -H <BMC IP address or BMC host name> chassis power on
  ```

- To power off the system normally, run the following command:
  
  ```
  # openbmctool -U <username> -P <password> -H <BMC IP address or BMC host name> chassis power softoff
  ```

- To power off the system immediately, run the following command:
  
  ```
  # openbmctool -U <username> -P <password> -H <BMC IP address or BMC host name> chassis power hardoff
  ```

Figure 2-2 shows an example how to query the power status of the system.

```
# openbmctool.py -U root -P <secret> -H <system> chassis power status
Attempting login...
Chassis Power State: On
Host Power State: Running
BMC Power State: Ready
User root has been logged out
```

*Figure 2-2 OpenBMC tool example*
A complete list of commands could be found in the Managing BMC-based systems with OpenBMC document in the IBM Knowledge Center:


### 2.1.5 The OpenBMC GUI

The OpenBMC GUI is a feature rich web access to the OpenBMC service processor. To access the OpenBMC GUI open a supported web browser. In the address bar, enter the IP address of the BMC that you want to connect to. For example, you can use the format https://<BMC IP> in the address bar of the web browser.

**Notes:**

Only HTTPS connections are possible (HTTP does not work).

The default user for the OpenBMC is `root`, and the default password is `openBmc` (0=zero, not O).

When you log in to the BMC of a system that is in the factory default state, the password is invalidated and you have to change the password immediately.

After the login, you will see a window similar to Figure 2-3.

![Figure 2-3 The OpenBMC GUI](image)

The BMC GUI 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.

In the BMC GUI, the following tasks could be accomplished:

- Get server health state and see the event logs.
- Get a detailed hardware inventory, including serial numbers of many components.
- Show sensor data, such as the temperature of many components, the RPMs of the fans, or the total power consumption in watts. Figure 2-4 shows a screenshot with some sensor data examples.
Figure 2-4  Sensor data examples in the OpenBMC

- **Server control:**
  - Server power operations
    Shutdown or reboot the server.
  - Manage power usage
    The minimum allowed power capping for the Power IC922 server is 500W and cannot be maintained in all system configurations and/or environments. The System firmware will attempt to maintain the power capping set by reducing processor core frequency down to the minimum (~2.3GHz), but the firmware is not guaranteed to maintain 500W over all system configurations and/or environments. If the power capping cannot be maintained or processor core frequency and system performance is unacceptable, then the power capping should be raised.
  - Server LED
    Turn on and turn off the server LED to identify the system.
  - Reboot BMC
    A reboot of the BMC does not affect the operating system. You can reboot the BMC even when the OS is up and running.
  - Serial over LAN console
    With a Serial over LAN (SOL) console, when the operating system is running, you can access the Linux even if no network is configured yet.
    When no operating system is running, you can access the Petitboot menu and there you may use the ipmitool to configure the system, for example the network access to the BMC.
  - Virtual Media
    Under Virtual Media you can upload an ISO image that will be presented to the system as a virtual DVD. This allows you to remote install the operating system.
Server configuration
- Configuration of the network settings of the BMC.
  See “Network ports for accessing the BMC interface” to identify the right port.
- SNMP settings
  Here you can add an SNMP server to send SNMP traps to.
- Manage the BMC and server firmware.
- Date and time settings
  Here it is also possible to define a Network Time Protocol (NTP) server.

Configure access control
- LDAP configuration
  It is possible to manage users and roles within OpenLDAP or Active Directory.
- Local user management
  Here you can add additional local users and their privileges or change the password of a user.
- Manage SSL certificates.

2.1.6 Connect to the BMC using SSH

The OpenBMC also supports SSH access. Simply use the IP or host name of the BMC and login with a user and password.

Notes:
The default user for the OpenBMC is root and the default password is openBmc (0=zero, not O).
When you log in to the BMC of a system that is in the factory default state, the password is invalidated, and you have to change the password immediately.

The SSH access provides the following command (among others):

- **obmcutil**
  - obmcutil state
    Shows you the state of the system, the BMC, and the operating system.
  - obmcutil poweron
    Powers on the system.
  - obmcutil poweroff
    Powers off the system.

- **obmc-console-client**
  Opens a shell into the operating system. When the system has no operating system running, it opens a shell into the IPL dialog or Petitboot loader.
2.1.7 Using a HMC for systems management

The Power IC922 server can also be attached to a Hardware Management Console (HMC). If the system is attached to an HMC, the following tasks can be accomplished:

- Power on / off the system.
- Launch BMC System Management.
- Call home

When the HMC is configured to receive SNMP traps and the Power IC922 server is configured to send SNMP traps, then the HMC is able to send events to IBM using the call home feature of the HMC.

- Firmware updates
- Activate and deactivate the identification LED.

For details about the attachment to an HMC, see Managing BMC-based systems by using the Hardware Management Console in the IBM Knowledge Center:


2.1.8 Working with the BMC over the REST API

OpenBMC-based systems such as the Power IC922 server can be managed by using the DMTF Redfish APIs. Redfish is a Representational State Transfer Application Programming Interface (REST API) used for platform management and is standardized by the Distributed Management Task Force, Inc. Redfish enables platform management tasks to be controlled by client scripts that are developed by using secure and modern programming paradigms.

The following example shows how to get the status of the identification LED, how to activate it and how to deactivate it using the REST API of the system:

- Log in to OpenBMC:
  
  ```shell
  # curl -c cjar -k -X POST -H "Content-Type: application/json" -d '{"data":
  ["root", "OpenBmc"]}' https://<BMC_IP>/login
  ```

- Turn on:
  
  ```shell
  # curl -b cjar -k -H "Content-Type: application/json" -X PUT -d '{"data": 1}'
  https://<BMC_IP>/xyz/openbmc_project/led/groups/enclosure_identify/attr/Asserted
  ```

- Turn off:
  
  ```shell
  # curl -b cjar -k -H "Content-Type: application/json" -X PUT -d '{"data": 0}'
  https://<BMC_IP>/xyz/openbmc_project/led/groups/enclosure_identify/attr/Asserted
  ```

Notes:

- Do not write to the persistent storage of the BMC.
- To close the shell enter ~.
2.2 Installation of the operating system

There are several options to install the operating system. All options start from the Petitboot loader. The installation ISO, media or network boot loader must be recognized by the Petitboot loader. Possible installation methods are:

- USB stick with the installation content
- Locally attached USB DVD drive
- Network boot using a PXE server
- Virtual DVD using the OpenBMC

2.2.1 Installation using a virtual media

To use the virtual DVD of the OpenBMC first download the Linux ISO image to your computer that you use to access the OpenBMC using a web browser. Then enter the OpenBMC GUI, select Server control, Virtual Media to network attach the ISO to the Power IC922 server as shown in Figure 2-5. After pressing the start button, the ISO is attached to the system.

![OpenBMC Virtual media](image)
After starting the system, the installation media shows up in Petitboot and you can select the boot entry, but do not start the installation yet. If you do so, you will not be able to finish the installation and the system will show errors similar to:

```
  dracut-initqueue[3537]: Warning: dracut-initqueue timeout - starting timeout scripts
```

Before you continue, remember the UUID of the virtual media boot device. In the example shown in Figure 2-6 the UUID is “2018-10-10-22-13-55-00”.

![Figure 2-6  Red Hat installation from Petitboot](image)

Before you boot the installation media, press `e` for edit when the Install Red Hat line is highlighted. Then add the boot arguments as shown in Figure 2-7. Replace the UUID with the one from your virtual media device from above.
Then press **OK**, which will bring you back into the boot menu, as shown in Figure 2-6. Now you can start the installation by pressing **Enter** on the highlighted Install Red Hat line in Petitboot.

After booting, the Anaconda installer will appear, as shown in Figure 2-8, and you can now continue with defining the details of your installation.
2.2.2 Installation using an USB stick

To use an USB stick for a Linux installation, first download the installation ISO image and copy it to an USB stick as in the following example:

```
# dd if=RHEL-ALT-7.6-20181010.0-Server-ppc64le-dvd1.iso of=/dev/usbms0 bs=32k
```

After that plug the USB stick into any of the USB ports and start the system into the Petitboot loader. There the USB stick will show up for installation similar to an installation from a virtual media device as shown in Figure 2-6.

The steps for an installation from USB sticks are the same as from a virtual media device. Follow the steps as described in 2.2.1, “Installation using a virtual media” on page 41.

More information about the installation of Linux for Power Systems can also be found in the Knowledge Center:

Reliability, availability, and serviceability

The Power System IC922 server brings 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). 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 runtime, 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

The Special Uncorrectable Error (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 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.

### 3.2 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:

- Installing and configuring
  

- Removing and replacing parts
  

- Troubleshooting, service, and support
  

- Installing the Linux operating system
  

The documentation could also be downloaded as a PDF, such as:

- IBM Power System IC922 (9183-22X) Quick Install Guide
- Installing the System and Ordered Parts for the IBM Power System IC922 (9183-22X)
- Servicing the IBM Power System IC922 (9183-22X)

The PDFs are available on the following website:


### 3.2.1 Concurrent maintenance

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

- Hard disk drives (HDDs) and solid state drives (SSDs)
- Redundant hot-plug power supplies
When replacing a power supply, the system continues to run at full performance. There is no performance degradation with only one running power supply.

- Redundant hot-plug cooling fans

All replacement parts in the system are customer replaceable parts.

### 3.2.2 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 using BMC GUI, the IPMI tool, the OpenBMC tool, the REST API, 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. A Firmware code update is supported through the BMC GUI, the IPMI interface and the OpenBMC tool. The firmware image can be updated or flashed regardless of its current state.

For details on how to work with the OpenBMC GUI, see chapter 2.1.5, “The OpenBMC GUI” on page 37.

### 3.2.3 Get SEL data using the OpenBMC tool

The OpenBMC tool is a good way to list service event log data or collect service data.

To list the SELs do:

```bash
# openbmctool -U <user> -P <password> -H <IP> sel print
```

To collect service data do:

```bash
# openbmctool -U <user> -P <password> -H <IP> collect_service_data
```

On a Linux system the data will be stored under: `/tmp/<IP>--<timestamp>`

A complete list of commands for the OpenBMC tool could be found in the Managing BMC-based systems with OpenBMC document in the Knowledge Center:


### 3.2.4 The system LEDs

The server contains a system identification LED button on the front and an additional LED on the rear side. The system identification LED could be turned on and off by pressing the system identification LED button on the front for at least 0.5 seconds. Another way to turn the LED on or off is the usage of the OpenBMC tool:

- **Activate the blue system identify LED:**
  ```bash
  # openbmctool -U <username> -P <password> -H <BMC IP address or BMC host name> chassis identify on
  ```
3.2.5 Factory reset

In case the password to log in to the openBMC was lost, a factory reset of the system might be needed. To do that boot into the local Petitboot shell (using a monitor and a keyboard) and enter the following command:

```bash
# ipmitool raw 0x3a 0x11
```

The factory reset function can take up to 15 minutes to complete.

3.2.6 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.
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 IC922 server.
IBM server racks

The Power IC922 server mounts in 42U Slim Racks 7965-S42, 7014-T42 and 7965-94Y along with square hole industry-standard racks. These racks are built to the 19-inch EIA 310D standard.

**Order information:** The Power IC922 server cannot be integrated into these racks during the manufacturing process, and are not orderable together with servers. If the Power IC922 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 58.

An IBM Business Partner can offer installation of these servers into a rack before delivery to a client’s 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 (FC EC02) or a lockable Rear Door Heat Exchanger (RDHX)(1164-95X) is used. Lockable optional side panels (FC 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 (FC 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.

**Note:** The 7965-94Y was withdrawn from marketing, but the Power IC922 server is also supported to be mounted in this type of rack.
IBM 42U SlimRack 7965-S42

The 2.0-meter (79-inch) Model 7965-S42 is compatible with past and present Power Systems servers, and 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.

Compared to the 7965-94Y Slim Rack, the Enterprise Slim Rack provides more strength and shipping/installation flexibility.

The 7965-S42 rack has space for up to four PDUs in side pockets. More PDUs beyond four are mounted horizontally and use 1U of rack space.

The Enterprise Slim Rack front door, which can be Basic Black/Flat (FC ECRM), High-End appearance (FC ECRF), or original equipment manufacturer (OEM) black (FC ECRE), has perforated steel, providing ventilation, physical security, and visibility of indicator lights in the installed equipment within. It comes standard with a lock that is identical to the locks in the rear doors. The door (FC ECRG and FC ECRE only) can be hinged on either the left or right side.

IBM 7014 Model T42 rack

The 2.0-meter (79.3-inch) Model T42 is compatible with past and present Power Systems servers. The features of the T42 rack are as follows:

► Has 42U (EIA units) of usable space.
► Has optional removable side panels.
► Has optional side-to-side mounting hardware for joining multiple racks.
► Has increased power distribution and weight capacity.
► Supports both AC power only.
► Up to four PDUs can be mounted in the PDU bays (see Figure A-2 on page 56), but others can fit inside the rack. For more information, see “The AC power distribution unit and rack content” on page 55.

► Ruggedized Rack Feature

For enhanced rigidity and stability of the rack, the optional Ruggedized Rack Feature (FC 6080) provides more hardware that reinforces the rack and anchors it to the floor. This hardware is designed primarily for use in locations where earthquakes are a concern. The feature includes a large steel brace or truss that bolts into the rear of the rack.

It is hinged on the left side so it can swing out of the way for easy access to the rack drawers when necessary. The Ruggedized Rack Feature also includes hardware for bolting the rack to a concrete floor or a similar surface, and bolt-in steel filler panels for any unoccupied spaces in the rack.

► Weights are as follows:
  – T42 base empty rack: 261 kg (575 lb)
  – T42 full rack: 930 kg (2045 lb)
Some of the available door options for the T42 rack are shown in Figure A-1.

<table>
<thead>
<tr>
<th>Trim kit (no front door)</th>
<th>Plain front door</th>
<th>Acoustic doors (front and rear)</th>
<th>Optional front door</th>
</tr>
</thead>
<tbody>
<tr>
<td>FC 6272</td>
<td>FC 6069</td>
<td>FC 6249</td>
<td>FC ERG7</td>
</tr>
</tbody>
</table>

The door options are explained in the following list:

- The 2.0 m Rack Trim Kit (FC 6272) is used if no front door is used in the rack.
- The front door for a 2.0 m Rack (FC 6069) is made of steel, with a perforated flat front surface. The perforation pattern extends from the bottom to the top of the door to enhance ventilation and provide some visibility into the rack. This door is non-acoustic and has a depth of about 25 mm (1 in.).
- The 2.0 m Rack Acoustic Door (FC 6249) consists of a front door and a rear door to reduce noise by approximately 6 dB(A). It has a depth of approximately 191 mm (7.5 in.).
- The FC ERG7 provides an attractive black full height rack door. The door is steel, with a perforated flat front surface. The perforation pattern extends from the bottom to the top of the door to enhance ventilation and provide some visibility into the rack. The non-acoustic door has a depth of about 134 mm (5.3 in.).

**Rear Door Heat Exchanger**

To lead away more heat, a special door that is named the Rear Door Heat Exchanger (FC 6858) is available. This door replaces the standard rear door on the rack. Copper tubes that are attached to the rear door circulate chilled water that is provided by the customer. The chilled water removes heat from the exhaust air being blown through the servers and attachments that are mounted in the rack. With industry-standard quick couplings, the water lines in the door attach to the customer-supplied secondary water loop.

For more information about planning for the installation of the IBM Rear Door Heat Exchanger, see the IBM Knowledge Center at:

The AC power distribution unit and rack content

Using previously provided IBM PDU features FC 7188, FC 7109, and FC 7196 reduces the number of Power IC922 servers and other equipment that can be held most efficiently in a rack. The high-function PDUs provide more electrical power per PDU and offer better “PDU footprint” efficiency. In addition, they are intelligent PDUs that provide insight to actual power usage by receptacle and also provide remote power on/off capability for easier support by individual receptacle. The new PDUs can be ordered as FC EPTJ, FC EPTL, FC EPTN, and FC EPTQ.

IBM Manufacturing integrates only the newer PDUs with the Power IC922 server. IBM Manufacturing does not support integrating earlier PDUs, such as FC 7188, FC 7109, or FC 7196, because these PDUs do not offer C19/C20 connectors. Clients can choose to use older IBM PDUs in their racks, but must install those earlier PDUs at their site.

The high-function PDU feature codes are shown in Table A-1.

<table>
<thead>
<tr>
<th>PDUs</th>
<th>Feature code for 1-phase or 3-phase depending on country wiring standards</th>
<th>Feature code for 3-phase 208 V depending on country wiring standards</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nine C19 receptacles(^a)</td>
<td>EPTJ</td>
<td>EPTL</td>
</tr>
<tr>
<td>Twelve C13 receptacles</td>
<td>EPTN</td>
<td>EPTQ</td>
</tr>
</tbody>
</table>

\(^a\) The Power IC922 server has an AC power supply with a C19/C20 connector.

Important: The Power IC922 server takes IEC 60320 C19/C20 mains power and not C13. Ensure that the correct power cords and PDUs are ordered or available in the rack.

In addition, the following high-function PDUs were announced in October 2019:

- **High Function 9xC19 PDU plus (FC ECJJ):**
  This is an intelligent, switched 200-240 volt AC Power Distribution Unit (PDU) plus with nine C19 receptacles on the front of the PDU. The PDU is mounted on the rear of the rack making the nine C19 receptacles easily accessible. For comparison, this is most similar to the earlier generation FC EPTJ PDU.

- **High Function 9xC19 PDU plus 3-Phase (FC ECJL):**
  This is an intelligent, switched 208 volt 3-phase AC Power Distribution Unit (PDU) plus with nine C19 receptacles on the front of the PDU. The PDU is mounted on the rear of the rack making the nine C19 receptacles easily accessible. For comparison, this is most similar to the earlier generation FC EPTL PDU.

- **High Function 12xC13 PDU plus (FC ECJN):**
  This is an intelligent, switched 200-240 volt AC Power Distribution Unit (PDU) plus with twelve C13 receptacles on the front of the PDU. The PDU is mounted on the rear of the rack making the twelve C13 receptacles easily accessible. For comparison, this is most similar to the earlier generation FC EPTN PDU.

- **High Function 12xC13 PDU plus 3-Phase (FC ECJQ):**
  This is an intelligent, switched 208 volt 3-phase AC Power Distribution Unit (PDU) plus with twelve C13 receptacles on the front of the PDU. The PDU is mounted on the rear of the rack making the twelve C13 receptacles easily accessible. For comparison, this is most similar to the earlier generation FC EPTQ PDU.
the rack making the twelve C13 receptacles easily accessible. For comparison, this is most similar to the earlier generation FC EPTQ PDU.

Table A-2 lists the feature codes for the high-function PDUs announced in October 2019.

Table A-2 High-function PDUs available after October 2019

<table>
<thead>
<tr>
<th>PDUs</th>
<th>Feature code for 1-phase or 3-phase depending on country wiring standards</th>
<th>Feature code for 3-phase 208 V depending on country wiring standards</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nine C19 receptacles&lt;sup&gt;a&lt;/sup&gt;</td>
<td>ECJJ</td>
<td>ECJL</td>
</tr>
<tr>
<td>Twelve C13 receptacles</td>
<td>ECJN</td>
<td>ECJQ</td>
</tr>
</tbody>
</table>

<sup>a</sup> The Power IC922 server has an AC power supply with a C19/C20 connector.

Four PDUs can be mounted vertically in the back of the S42 and T42 racks. Figure A-2 shows the placement of the four vertically mounted PDUs. In the rear of the rack, two more PDUs can be installed horizontally in the T00 rack and three in the S42 and T42 rack. The four vertical mounting locations are filled first in the S42 and T42 racks.

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 proper air-flow 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-3 shows the available wall power cord options for the PDU and intelligent power distribution unit (iPDU) features, which must be ordered separately.

### Table A-3  Wall power cord options for the power distribution unit and iPDU features

<table>
<thead>
<tr>
<th>Feature code</th>
<th>Wall plug</th>
<th>Rated voltage (V AC)</th>
<th>Phase</th>
<th>Rated amperage</th>
<th>Geography</th>
</tr>
</thead>
<tbody>
<tr>
<td>6653</td>
<td>IEC 309, 3P+N+G, 16 A</td>
<td>230</td>
<td>3</td>
<td>16 amps/phase</td>
<td>Internationally available</td>
</tr>
<tr>
<td>6489</td>
<td>IEC 309, 3P+N+G, 32 A</td>
<td>230</td>
<td>3</td>
<td>32 amps/phase</td>
<td>EMEA</td>
</tr>
<tr>
<td>6654</td>
<td>NEMA L6-30</td>
<td>200 - 208, 240</td>
<td>1</td>
<td>24 amps</td>
<td>US, Canada, Latin America, and Japan</td>
</tr>
<tr>
<td>6655</td>
<td>RS 3750DP (watertight)</td>
<td>200 - 208, 240</td>
<td>1</td>
<td>24 amps</td>
<td>US, Canada, Latin America, and Japan</td>
</tr>
<tr>
<td>6656</td>
<td>IEC 309, P+N+G, 32 A</td>
<td>230</td>
<td>1</td>
<td>24 amps</td>
<td>EMEA</td>
</tr>
<tr>
<td>6657</td>
<td>PDL</td>
<td>230 - 240</td>
<td>1</td>
<td>32 amps</td>
<td>Australia, New Zealand</td>
</tr>
<tr>
<td>6658</td>
<td>Korean plug</td>
<td>220</td>
<td>1</td>
<td>30 amps</td>
<td>North and South Korea</td>
</tr>
<tr>
<td>6492</td>
<td>IEC 309, 2P+G, 60 A</td>
<td>200 - 208, 240</td>
<td>1</td>
<td>48 amps</td>
<td>US, Canada, Latin America, and Japan</td>
</tr>
<tr>
<td>6491</td>
<td>IEC 309, P+N+G, 63 A</td>
<td>230</td>
<td>1</td>
<td>63 amps</td>
<td>EMEA</td>
</tr>
<tr>
<td>6667</td>
<td>PDL, 3P+N+G</td>
<td>230 - 240</td>
<td>1</td>
<td>32 amps</td>
<td>Australia, New Zealand</td>
</tr>
<tr>
<td>ECJ5</td>
<td>IEC 309, 3P+G</td>
<td>200 - 240</td>
<td>3</td>
<td>24 amps</td>
<td>United States, Canada, Latin America, Japan, and Taiwan</td>
</tr>
<tr>
<td>ECJ7</td>
<td>IEC 309, 3P+G</td>
<td>200 - 240</td>
<td>3</td>
<td>48 amps</td>
<td>United States, Canada, Latin America, Japan, and Taiwan</td>
</tr>
</tbody>
</table>

**Notes:** Ensure that the appropriate 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 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 and 7965-94Y racks, see the IBM Knowledge Center at:

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. The pins of the rack mounting rails fit into both round and square holes. This standard is published by the Electrical Industries Alliance.

For more information, see the IBM Knowledge Center at:

https://www.ibm.com/support/knowledgecenter/9183-22X/p9ia4/p9ia4_oemrack.htm#p9ia4_oemrack

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. For more information, see the IBM Knowledge Center at:
- 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 listed in this section are considered particularly suitable for a more detailed discussion of the topics 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
- Networking Design for HPC and AI on IBM Power Systems, REDP-5478
- IBM Power System AC922 Technical Overview and Introduction REDP-5494
- IBM Power Systems LC921 and LC922 Technical Overview and Introduction, REDP-5495
- IBM Power System L922 Technical Overview and Introduction, REDP-5496
- IBM Power Systems S922, S914, and S924 Technical Overview and Introduction, REDP-5497
- IBM Power Systems H922 and H924 Technical Overview and Introduction, REDP-5498
- IBM Power Systems E950 Technical Overview and Introduction, REDP-5509
- IBM Power Systems E980 Technical Overview and Introduction, REDP-5510

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:

- PCIe adapter placement rules in the IBM Knowledge Center:

- IBM POWER processor family, visit the following website:
  https://www.ibm.com/systems/power/openpower/

- POWER9 LaGrange Single-Chip Module data sheet:
  https://ibm.ent.box.com/s/1dsn840viar4qb6a7ekzjb13b2xr1v91

- For more information about the LaGrange module or the other modules of the IBM POWER processor family, visit the following website:
  https://www.ibm.com/systems/power/openpower/

- IBM Knowledge Center: memory placement rules

- The NVIDIA website provide detailed information about T4 GPUs:

- Details on allowed password for the OpenBMC can be found at:
  https://linux.die.net/man/8/pam_cracklib

- You may download the openbmc_tool from the following website:

- A complete list of commands could be found in the Managing BMC-based systems with OpenBMC document in the IBM Knowledge Center:

- For details about the attachment to an HMC, see Managing BMC-based systems by using the Hardware Management Console in the IBM Knowledge Center:

- For more information about the REST API of the Power IC922 server see the document Managing BMC-based systems with OpenBMC in the Knowledge Center:

- For more details about the DMTF Redfish APIs in general, see:
  http://www.dmtf.org/standards/redfish

IBM Knowledge Center provides current documentation to service effectively the system:
- Installing and configuring

- Removing and replacing parts

- Troubleshooting, service, and support

- Installing the Linux operating system

- A complete list of commands for the OpenBMC tool could be found in the Managing BMC-based systems with OpenBMC document in the Knowledge Center:

- For more information about planning for the installation of the IBM Rear Door Heat Exchanger, see the IBM Knowledge Center at:

- For detailed power requirements and power cord details about the 7014 and 7965-94Y racks, see the IBM Knowledge Center at:
For more information about OEM Racks, see the IBM Knowledge Center at:
https://www.ibm.com/support/knowledgecenter/9183-22X/p9ia4/p9ia4_oemrack.htm#p9ia4_oemrack

IBM Fix Central website
http://www.ibm.com/support/fixcentral/

IBM Knowledge Center: IBM Power Systems Hardware
http://www-01.ibm.com/support/knowledgecenter/api/redirect/powersys/v3r1m5/index.jsp

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/

OpenPOWER Foundation
https://openpowerfoundation.org/

Support for IBM Systems website

Help from IBM

IBM Support and downloads
ibm.com/support

IBM Global Services
ibm.com/services