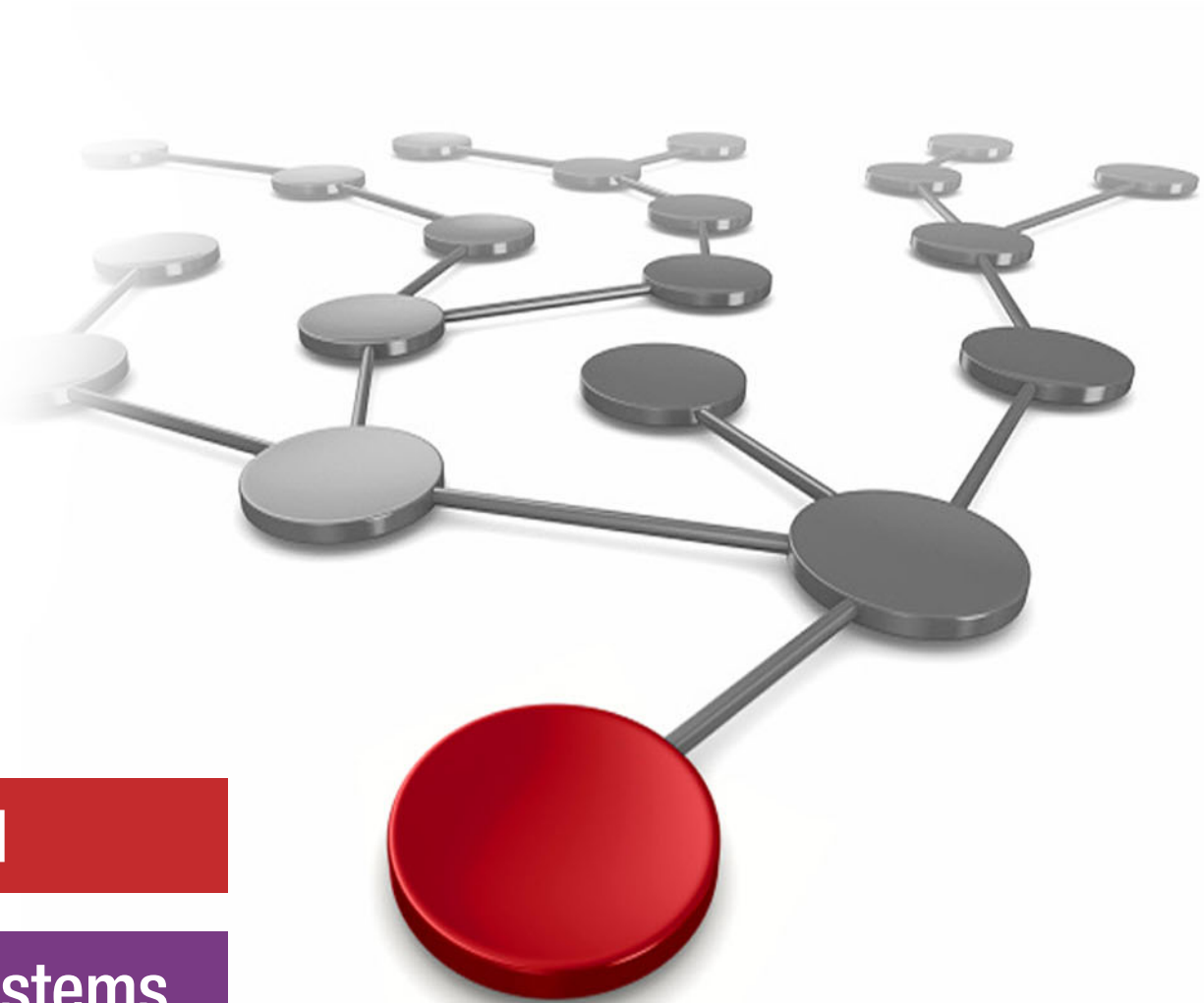# IBM Power E1050
Technical Overview and Introduction

Giuliano Anselmi

Marc Gregorutti

Stephen Lutz

Michael Malicdem

Guido Somers

Tsvetomir Spasov

Cloud

Power Systems

IBM

IBM Redbooks

IBM Power E1050: Technical Overview and Introduction

August 2022

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

First Edition (August 2022)

This edition applies to the IBM Power E1050 (9043-MRX) server and the Hardware Management Console (HMC) Version 10 Release 1.

Contents

Notices

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

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

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

INTERNATIONAL BUSINESS MACHINES CORPORATION PROVIDES THIS PUBLICATION "AS IS" WITHOUT WARRANTY OF ANY KIND, EITHER EXPRESS OR IMPLIED, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF NON-INFRINGEMENT, MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE. Some jurisdictions do not allow disclaimer of express or implied warranties in certain transactions, therefore, this statement may not apply to you.

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

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

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

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

Information concerning non-IBM products was obtained from the suppliers of those products, their published announcements or other publicly available sources. IBM has not tested those products and cannot confirm the accuracy of performance, compatibility or any other claims related to non-IBM products. Questions on the capabilities of non-IBM products should be addressed to the suppliers of those products.

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

This information contains examples of data and reports used in daily business operations. To illustrate them as completely as possible, the examples include the names of individuals, companies, brands, and products. All of these names are fictitious and any similarity to actual people or business enterprises is entirely coincidental.

COPYRIGHT LICENSE:

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

Trademarks

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

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

AIX®	IBM FlashCore®	PowerPC®
DS8000®	IBM FlashSystem®	PowerVM®
Easy Tier®	IBM Security®	QRadar®
HyperSwap®	IBM Spectrum®	Redbooks®
IBM®	Instana®	Redbooks (logo) ®
IBM Cloud®	Micro-Partitioning®	Storwize®
IBM Cloud Pak®	Power Architecture®	Tivoli®
IBM Elastic Storage®	PowerHA®	Turbonomic®

The following terms are trademarks of other companies:

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

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.

LTO, Ultrium, the LTO Logo and the Ultrium logo are trademarks of HP, IBM Corp. and Quantum in the U.S. and other countries.

Red Hat, OpenShift, and Ansible 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.

VMware, and the VMware logo are registered trademarks or trademarks of VMware, Inc. or its subsidiaries in the United States and/or other jurisdictions.

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 E1050 server (9043-MRX) that uses the latest IBM Power10 processor-based technology and supports IBM AIX® and Linux operating systems (OSs). The goal of this paper is to provide a hardware architecture analysis and highlight the changes, new technologies, and major features that are being introduced in this system, such as:

► The latest IBM Power10 processor design, including the dual-chip module (DCM) packaging, which is available in various configurations from 12 - 24 cores per socket.

► Support of up to 16 TB of memory.

► Native Peripheral Component Interconnect Express (PCIe) 5th generation (Gen5) connectivity from the processor socket to deliver higher performance and bandwidth for connected adapters.

► Open Memory Interface (OMI) connected Differential Dual Inline Memory Module (DDIMM) memory cards delivering increased performance, resiliency, and security over industry-standard memory technologies, including transparent memory encryption.

► Enhanced internal storage performance with the use of native PCIe-connected Non-volatile Memory Express (NVMe) devices in up to 10 internal storage slots to deliver up to 64 TB of high-performance, low-latency storage in a single 4-socket system.

► Consumption-based pricing in the Power Private Cloud with Shared Utility Capacity commercial model to allow customers to consume resources more flexibly and efficiently, including AIX, Red Hat Enterprise Linux (RHEL), SUSE Linux Enterprise Server, and Red Hat OpenShift Container Platform workloads.

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

► IBM Power customers
► Sales and marketing professionals
► Technical support professionals
► IBM Business Partners
► Independent software vendors (ISVs)

This paper expands the set of IBM Power documentation by providing a desktop reference that offers a detailed technical description of the Power E1050 Midrange server model.

This paper does not replace the current marketing materials and configuration tools. It is intended as an extra source of information that, together with existing sources, can be used to enhance your knowledge of IBM server solutions.

Authors

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

Giuliano Anselmi is an IBM Power Digital Sales Technical Advisor in IBM Digital Sales Dublin. He joined IBM and focused on Power processor-based technology. For almost 20 years, he covered several technical roles. He is an important resource for the mission of his group, and he serves a reference for IBM Business Partners and customers.

Marc Gregorutti is an Europe, Middle East, and Asia (EMEA) IBM Power Product Field Engineer at IBM France. He started as a IBM service representative in 1998, and then became a remote technical support member for IBM Power for France and then the EMEA. He joined the EMEA IBM Power Product engineering team in 2009 and became one of its leaders. He now focuses on scale-out enterprise midrange systems support to continuously improve the product.

Stephen Lutz is a Certified Leading Technical Sales Professional for IBM Power working at IBM Germany. He holds a degree in Commercial Information Technology from the University of Applied Science Karlsruhe, Germany. He has 23 years of experience in AIX, Linux, virtualization, and IBM Power and its predecessors. He provides pre-sales technical support to clients, IBM Business Partners, and IBM sales representatives in Germany.

Michael Malicdem is a brand technical specialist for IBM Power at IBM Philippines. He has over 17 years of experience in pre-sales solution design, client and partner consultation, presentations, technical enablements relative to IBM Power servers, and IBM storage, which includes 3 years in post-sales services and support. He is a licensed Electronics Engineer and a cum laude graduate from Pamantasan ng Lungsod ng Maynila (PLM). His areas of expertise include server configurations, AIX, RHEL, IBM PowerVM®, PowerSC, PowerVC, and IBM PowerHA®.

Guido Somers a Distinguished Technical Specialist working at IBM Switzerland. He has more than 25 years of experience in the information technology field, 24 years of which were with IBM. He holds several degrees, and he did research in the field of theoretical physics. His areas of expertise include cloud, IT architecture, virtualization, cross-platform infrastructure, IT analytics, ISV Solutions - SAP, Oracle, and Cisco, blockchain, data center migration and optimization, artificial intelligence (AI), and security. He is an IBM Redbooks® Platinum author.

Tsvetomir Spasov is an IBM Power subject matter expert at IBM Bulgaria. His main areas of expertise are Flexible Service Processor (FSP), enterprise Baseboard Management Controller (eBMC), Hardware Management Console (HMC), IBM PowerPC®, and Global Total Microcode Support (GTMS). He has been with IBM since 2016, and provides reactive break-fix, proactive, preventive, and cognitive support. He has conducted different technical trainings and workshops.

The project that produced this publication was managed by:
Scott Vetter, PMP

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

Jesse P Arroyo, Irving Baysah, Nigel Griffiths, Sabine Jordan, Charles Marino, Hariganesh Muralidharan, Hoa Nguyen, Ian Robinson, William Starke, Edward M.H. Tam, Madeline Vega
IBM

Vincencio Michaelis
Bosch

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 LinkedIn:

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

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

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

► Stay current on recent Redbooks publications with RSS Feeds:

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

IBM Power E1050 overview

This chapter provides general description of the new IBM Power E1050 (9043-MRX) server, which is a logical follow-on to the IBM Power E950. The IBM Power E1050 server is an enhanced enterprise class four-socket, 4U server that provides massive scalability and flexibility. These systems are agile and deliver extreme density in an energy-efficient design with best in class reliability and resiliency. They provide enterprise clients with a secure environment that balances mission-critical and modernization applications to deliver a frictionless, hybrid cloud experience.

1.1 System overview

The Power E1050 server is ideal for cloud deployments because of its built-in virtualization, flexible capacity, and high usage. The machine type model number of the Power E1050 server is 9043-MRX. It is a single enclosure server that is four EIA units tall (4U). It can be configured with two, three, or four dual-chip modules (DCMs). Three processor options are available:

► Twelve cores running at a typical 3.35 - 4.00 GHz (max) frequency range

► Eighteen cores running at a typical 3.20 - 4.00 GHz (max) frequency range

► Twenty-four cores running at a typical 2.95 - 3.90 GHz (max) frequency range

A Power E1050 server with four 24-core DCMs offers the maximum of 96 cores. All processor cores can run up to eight simultaneous threads to deliver greater throughput. All sockets must be populated with the same processor modules.

Figure 1-1 shows the Power E1050 server.

Figure 1-1   The Power E1050 server

Figure 1-2 shows a top view of the Power E1050 server with the top lid removed. Under the left metal plate are the fans and Non-volatile Memory Express (NVMe) slots, as shown in Figure 1-3 on page 4. Going to the right, you can see the memory slots that belong the processors to the right of that memory column. Going further to the right, there is another column of memory slots that belong to the processors to the right of them. Under the metal plate at the right edge are the four Titanium class 2300W power supplies and the 11 Peripheral Component Interconnect Express (PCIe) slots, as shown in Figure 1-4 on page 5. The air flow goes from the front to the rear of the server, which in Figure 1-2 is from left to right.
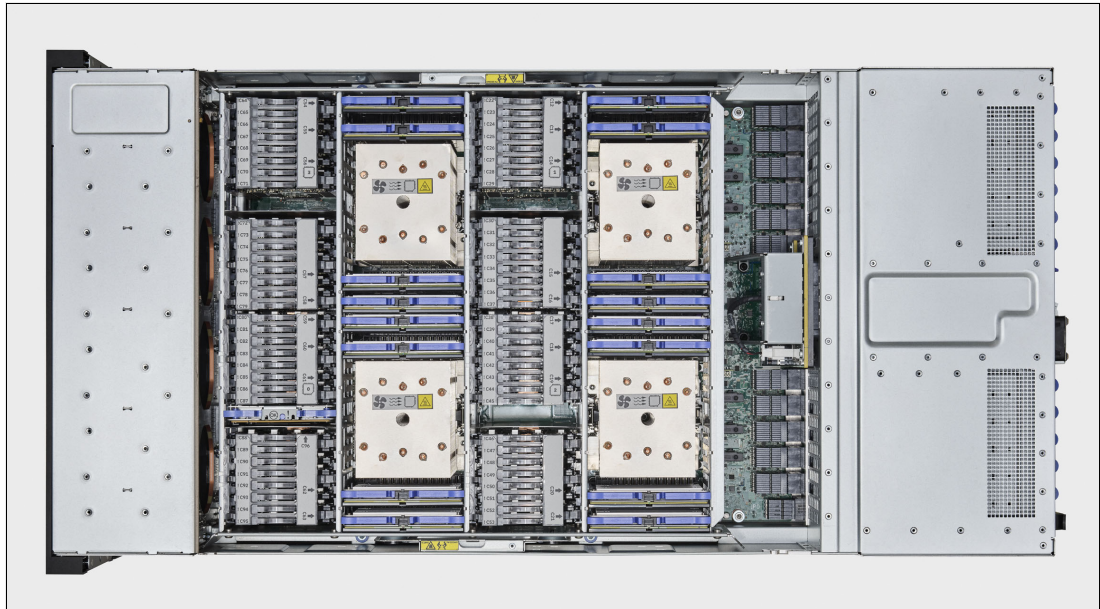
Figure 1-2   Top view of the Power E1050 server with the top lid removed

Each processor module that is added to the system offers 16 Open Memory Interface (OMI) slots that can be populated with 4U Differential Dual Data Rate DIMMs (DDIMMs). At the time of writing, these DDIMMs incorporate Double Data Rate 4 (DDR4) memory chips that deliver an increased memory bandwidth of up to 409 GBps peak transfer rates per socket. With four processor modules, the Power E1050 server provides 64 OMI slots that support up to 16 TB of memory and a maximum peak transfer rate of 1.636 GBps.

The Power E1050 server provides state-of-the-art PCIe Gen5 connectivity. Up to 11 PCIe slots are provided in the system unit with different characteristics:

► Six PCIe Gen4 x16 or PCIe Gen5 x8 slots

► Two PCIe Gen5 x8 slots

► Three PCIe Gen4 x8 slots

The number of available slots depends on the number of available processor modules. For more information about the system diagram, see Figure 2-1 on page 35.

> **Note:** Although some slots are x8 capable only, all connectors in the system have x16 connectors.

If more slots are needed, up to four PCIe Gen3 I/O Drawers with two fanout modules each can be added to the system. Each fanout module provides six slots. With eight fanout modules in four I/O drawers, the maximum number of available slots is 51.

The PCIe slots can be populated with a range of adapters covering local area network (LAN), Fibre Channel (FC), serial-attached SCSI (SAS), Universal Serial Bus (USB), and cryptographic accelerators. At least one network adapter must be included in each system.

The Power E1050 server offers up to 10 internal NVMe U.2 flash bays that can be equipped with 800 GB U.2 Mainstream NVMe drives or U.2 Enterprise class NVMe drives in different sizes up to 6.4 TB. Each NVMe device is connected as a separate PCIe endpoint and can be assigned individually to VMs for best flexibility. The 10 NVMe bays offer a maximum of 64 TB internal storage. For all 10 NVMe bays to b available, the server must be populated with all

four processor cards. With two or three processor cards, the server can be populated with six NVMe devices.

The Power E1050 server does not have internal spinning SAS drives. However, it is possible to attach 19-inch disk expansion drawers that offer SFF Gen2-carrier bays for SAS disks. For more information, see 2.3, "Internal I/O subsystem" on page 65.

In addition to extensive hardware configuration flexibility, the Power E1050 server offers Elastic Capacity on Demand (Elastic CoD) temporarily for both processor cores and memory; IBM Active Memory Expansion; and Active Memory Mirroring (AMM) for hypervisor.

For the best flexibility, the Power E1050 server can be part of an IBM Power Private Cloud with Shared Utility Capacity pool, also known as IBM Power Enterprise Pool 2.0. It consists of Power E1050 servers, Power E950 servers, or a mix of both. In such a pool, Base Capacity can be purchased for processor cores, memory, and operating system (OS) licenses (AIX) or subscriptions (Linux). This Base Capacity is independent of the configuration of the servers. Even if only a small Base Capacity was purchased, all available resources of the servers in the pool can be used. If more resources are used than are available in the Base Capacity of the pool (the sum of all Base Capacities of all servers that are part of the pool), these additional used resources, that is, metered resource consumption, are billed. The metering is done on a per-minute basis. The billing can be pre-paid by purchasing credits upfront, or the billing can be post-pay. In a post-pay pool, IBM generates an invoice monthly.

The Power E1050 server includes IBM PowerVM Enterprise Edition to deliver virtualized environments and support a frictionless hybrid cloud experience. Workloads can run the AIX, and Linux OSs, including the Red Hat OpenShift Container Platform. IBM i is not a supported OS on the Power E1050 server.

The Power E1050 server also provides strong resiliency characteristics, which include Power10 chip capabilities and memory protection. The new 4U DDIMMS that are used in the Power E1050 server offer an enhanced buffer, N+1 voltage regulation, and spare dynamic RAM (DRAM) technology. Also, technologies like Chipkill with advanced error correction code (ECC) protection are included, and transparent Power10 memory encryption with no performance impact. This technology is the same enterprise class technology that is used in the Power E1080 server.

Other resiliency features that are available in the Power E1050 server are hot-plug NVMe bays, hot-plug PCIe slots, redundant and hot-plug power supplies, hot-plug redundant cooling fans, hot-plug Time of Day battery, and even highly resilient architecture for power regulators.

Figure 1-3 shows the front view of a Power E1050 server with the front bezel removed.
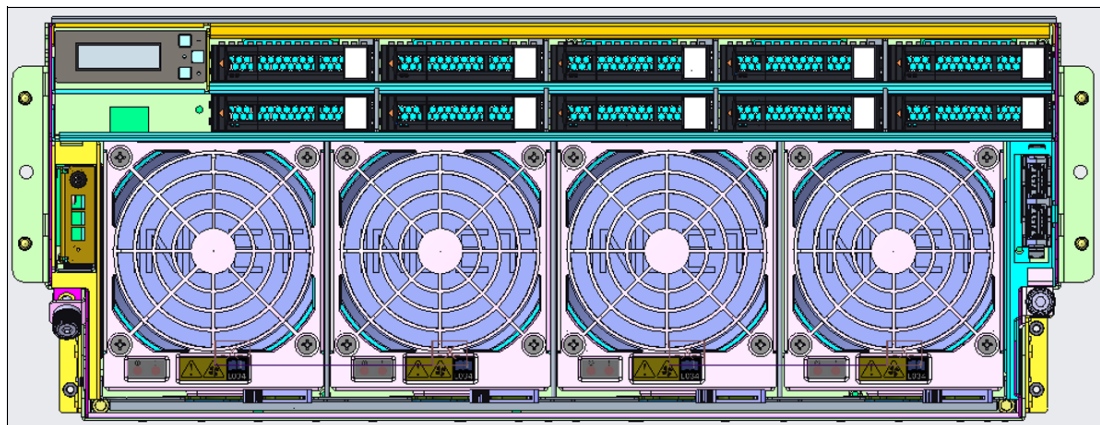
Figure 1-3   Front view of the Power E1050 server

Figure 1-4 shows the rear view of the Power E1050 server. The leftmost slot (P0-C0) is the enterprise Baseboard Management Controller (eBMC) Service Processor Card, and then there are five PCIe slots. On the right side are six more adapter slots.
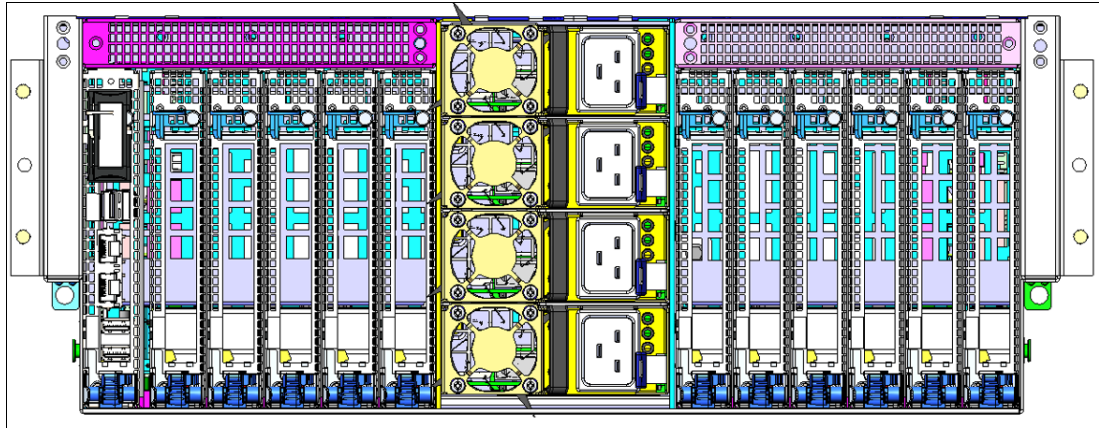
Figure 1-4   Rear view of the Power E1050 server

Table 1-1 shows a summary of features of the Power E1050 server.

Table 1-1   Power E1050 server feature summary

Feature	Comments
Processors	12, 18, or 24 CPU cores per socket.
Sockets	Four sockets are available. 2, 3, or 4 sockets may be populated.
Memory	► Up to 64 OMI slots that can be equipped with 4U DDIMMs. ► DDIMM sizes are 32, 64, 128, and 256 GB. ► 8 TB 16 TB maximum memory.
Integrated PCIe	► Six PCIe Gen4 16-lane or PCIe Gen5 8-lane slots. ► Two PCIe Gen5 8-lane slots. ► Three PCIe Gen4 8-lane slots. ► PCIe slots are full-high and half-length, and use Blind-Swap Cassettes (BSCs).
Internal NVMe Flash bays	Up to 10 U.2 NVMe bays for 15-mm NVMe drives or 7-mm NVMe drives in a 15-mm carrier.
Internal USB ports	USB 3.0. Two front and two rear.
Media bays	DVD through an external USB DVD.
Maximum I/O drawers	Four PCIe Gen3 I/O drawers (#EMX0).
External storage drawers	Up to 64 EXP12SX or ESP24SX drawers.

Table 1-2 shows the major differences between the Power E950 and Power E1050 servers.

Table 1-2   Comparing the Power E950 and Power E1050 servers

Features	Power E950 server	Power E1050 server
Processor	IBM Power9 (single-chip module (SCM))	Power10 (DCM)
Sockets	2 - 4	2 - 4
Maximum cores	32, 40, 44, or 48	48, 72, or 96
Maximum memory	16 TB	16 TB
DIMM type and DIMM slots count	Up to 128 industry-standard DIMMs	Up to 64 DDIMMs
L4 cache	Yes	Yes
Memory bandwidth	920 GBps	1.636 GBps
Memory DRAM spare	Yes	Yes
I/O expansion slots	Yes	Yes
PCIe slots	11 (eight Gen4 16-lane + two Gen4 8-lane + one Gen3 slots)	11 (six Gen5 x8/Gen4 x16 + two Gen5 x8 + three Gen4 x8 slots)
Acceleration ports	Eight (CAPI 2.0 + IBM OpenCAPI)	Six (IBM OpenCAPI)
PCIe hot-plug support	Yes + blind swap	Yes + blind swap
IO bandwidth	630 GBps	756 GBps
Internal storage bays	12 (eight SAS + four NVMe drives)	10 (10 NVMe drives)
Internal storage controllers	Optional Concurrently maintainable	Optional Concurrently maintainable

1.2 Operating environment

Table 1-3 details the operating environment for the Power E1050 server.

Table 1-3   Operating environment for the Power E1050 server

Power E1050 operating environment		
System	Power E1050 server	
Item	Operating	Non-operating
Temperature	Recommended: 18 - 27 °C (64.4 - 80.6 °F) Allowable: 5 - 40 °C (41.0 - 140.0 °F)	5 - 45 °C (41.0 - 113.0 °F)
Relative humidity	8 - 80%	5 - 85%
Maximum dew point	24 °C (75.2 °F)	27 °C (80.6 °F)
Operating voltage	200 - 240 V AC	N/A

Power E1050 operating environment		
System	Power E1050 server	
Operating frequency	50 - 60 Hz +/- 3 Hz AC	N/A
Maximum power consumption	5,200 W maximum	N/A
Maximum power source loading	5.36 kVA maximum	N/A
Maximum thermal output	17,742 BTU per hour	N/A
Maximum altitude	3,050 m (10,000 ft.)	N/A
Maximum noise level	8.2 bels LwAm (heavy workload on one maximally configured 4-socket enclosure, 2 TB memory, 25°C, 500 m)	N/A

Note: IBM does not recommend operation above 27C, however, one can expect full performance up to 35C for these systems. Above 35C, the system is capable of operating, but possible reductions in performance may occur to preserve the integrity of the system components. Above 40C there may be reliability concerns for components within the system.

Environmental assessment: The IBM Systems Energy Estimator tool can provide more accurate information about the power consumption and thermal output of systems that are based on a specific configuration, including adapters and I/O expansion drawers.

Note: Derate maximum allowable dry-bulb temperature 1°C (1.8°F) per 175 m (574 ft) above 900 m (2,953 ft) up to a maximum allowable elevation of 3050 m (10,000 ft).

Government regulations, such as those prescribed by the Occupational Safety and Health Administration (OSHA) or European Community Directives, may govern noise level exposure in the workplace, which might apply to you and your server installation. The Power E1050 is available with an optional acoustical door feature that can help reduce the noise that is emitted from this system.

The actual sound pressure levels in your installation depend upon various factors, including the number of racks in the installation, the size, materials, and configuration of the room where you designate the racks to be installed, the noise levels from other equipment, the ambient room temperature, and employees' location in relation to the equipment.

Compliance with such government regulations also depends on many more factors, including the duration of employees' exposure and whether employees wear hearing protection. IBM recommends that you consult with qualified experts in this field to determine whether you are in compliance with the applicable regulations.

1.3 Physical package

The system node requires 4U and the PCIe I/O expansion drawer requires 4U. Thus, a single-enclosure system with one PCIe I/O expansion drawer requires 8U.

Table 1-4 lists the physical dimensions of the system node and the PCIe Gen3 I/O Expansion Drawer.

Table 1-4   Physical dimensions of the system node and the PCIe Gen3 I/O Expansion Drawer

Dimension	Power E1050 system node	PCIe I/O expansion drawer
Width	448 mm (17.6 in.)	448 mm (17.6 in.)
Depth	902 mm (35.5 in.)	736.6 mm (29.0 in.)
Height	175 mm (6.9 in.) four EIA units	177.8 mm (7.0 in.) four EIA units
Weight	69 kg (153 lb)	54.4 kg (120 lb)

To ensure installability and serviceability in non-IBM industry-standard racks, review the installation planning information for any product-specific installation requirements.

Note: The EMX0 remote I/O drawer connection in the T42 and S42 racks stops the rear door from closing, so you must have the 8-inch rack extensions.

1.3.1 Rack-mounted model

Table 1-5 shows the physical dimensions of the Power E1050 rack-mounted chassis. The server is available only in a rack-mounted form factor and takes 4U of rack space.

Table 1-5   Physical dimensions of the Power E1050 rack-mounted chassis

Dimension	Power E1050 server 9043-MRX
Width	448 mm (17.6 in.)
Depth	902 mm (35.5 in.)
Height	175 mm (6.9 in.), 4 EIA units
Weight	69 kg (153 lb)

Figure 1-5 shows the front view of the Power E1050 server.

Figure 1-5   Front view of the Power E1050 server

1.4 System features

This section covers the standard system features that are included in the Power E1050 server.

1.4.1 Power E1050 server features

This summary describes the standard features that are available on the Power E1050 (9043-MRX) server model:

► The Power E1050 supports 24 - 96 processor cores with 2 - 4 Power10 processor modules.

► The Power E1050 delivers 256 GB to 16 TB high-performance DDR4 memory with an L4 cache.

► A Power E1050 server supports up to 10 NVMe drives.

► Up to 11 hot-swap PCIe slots may be in the system unit:
  – Six PCIe Gen5 x8 or Gen4 x16 slots.
  – Three PCIe Gen4 x8 slots.
  – Two PCIe Gen5 x 8 slots.
  – With two processor modules, there are seven PCIe slots; with three modules and four modules, there are 11 PCIe slots.

► The PCIe I/O Expansion Drawer (#EMX0) expands the number of full-high, hot-swap slots:
  – Up to two PCIe drawers with two processor modules (maximum of 31 slots on the server).
  – Up to four PCIe drawers with four processor modules (maximum of 51 slots on the server).

► Up to 64 EXP24SX SFF Drawers (#ESLS) can be attached, providing up to 1,536 SAS bays for disks or solid-state drives (SSDs).

- System unit I/O (integrated I/O):
  - USB 3.0 ports: four 3.0 ports, two front and two rear.
  - USB 2.0 ports: two rear 2.0 ports for limited use.
  - HMC ports: two 1 GbE RJ45.
- Four hot-plug and redundant power supplies 2300 W (200 – 240 V AC) (#EB39).
- System unit only 4U in a 19-inch rack-mount hardware.
- Primary OSs:
  - AIX (#2146): small-tier licensing.
  - Linux (#2147): Red Hat Enterprise Linux (RHEL) and SUSE Linux Enterprise Server.

Processor modules
- The Power E1050 supports 24 - 96 processor cores:
  - Twelve-core typical 3.35 – 4.0 GHz (max) #EPEU Power10 processor.
  - Eighteen-core typical 3.20 – 4.0 GHz (max) #EPEV Power10 processor.
- Twenty-four-core typical 2.95 – 3.90 GHz (max) #EPGW Power10 processor.
- A minimum of two and a maximum of four processor modules are required for each system. The modules can be added to a system later through a Miscellaneous Equipment Specification (MES) upgrade, but the system requires scheduled downtime to install. All processor modules in one server must be of the same frequency (same processor module feature number), that is, you cannot mix processor modules of different frequencies.
- Permanent CoD processor core activations are required for the first processor module in the configuration and are optional for the second and fourth modules. Specifically:
  - Two, three, or four 12-core typical 3.35 – 4.0 GHz (max) processor modules (#EPEU) require 12 processor core activations (#EPUR) at a minimum.
  - Two, three, or four 18-core typical 3.20 – 4.0 GHz (max) processor modules (#EPEV) require 18 processor core activations (#EPUS) at a minimum.
- Two, three, or four 24-core typical 2.95 – 3.90 GHz (max) processor modules (#EPGW) require 24 processor core activations (#EPYT) at a minimum.
- Temporary CoD capabilities are optionally used for processor cores that are not permanently activated. An HMC is required for temporary CoD.

System memory
- 256 GB - 16 TB high-performance memory up to 3200 MHz DDR4 OMI:
  - 64 GB DDIMM Memory (#EM75).
  - 128 GB DDIMM Memory (#EM76).
  - 256 GB DDIMM Memory (#EM77).
  - 512 GB DDIMM Memory (#EM7J).
  - Optional Active Memory Expansion (#EMBM).

► As your memory requirements increase, the system capabilities increase as follows:
  – With two processor modules installed, 32 DDIMM slots are available. Minimum memory is 256 GB.
  – With three processor modules installed, 48 DDIMM slots are available. Minimum memory is 384 GB.
  – With four processor modules installed, 64 DDIMM slots are available. Minimum memory is 512 GB. Sixteen DDIMMs are available per socket.
  – The more DDIMM slots that are filled, the larger the bandwidth that is available to the server.

Permanent CoD memory activations are required for at least 50% of the physically installed memory or 256 GB of activations, whichever is larger. Use 1 GB activation (#EMCP) and 100 GB activation (#EMCQ) features to order permanent memory activations.

Temporary CoD for memory is available for memory capacity that is not permanently activated.

► Delivery through Virtual Capacity machine type and model (MTM) (4586-COD) by using the IBM Entitled Systems Support (IBM ESS) process.

► An HMC is required for temporary CoD.

Notes:

► Memory is ordered in a quantity of eight of the same memory feature.

► The minimum memory that is supported per two Power10 processors installed is 256 GB.

► The minimum memory that is supported per three Power10 processors installed is 384 GB.

► The minimum memory that is supported per four Power10 processors installed is 512 GB.

Storage options

The Power E1050 supports up to 10 NVMe 7-mm or 15-mm drives:

► Six NVMe drives within a two or three-socket configuration

► Ten NVMe drives within a four-socket configuration

All NVMe drives are driven directly from the system backplane with no PCIe card or cables required.

The 7-mm NVMe drives from the IBM Power E950 are also supported on the Power E1050 with a carrier conversion feature that is offered to hold these drives.

1.5 Minimum configuration

Minimum configuration refers to the most basic parts of the Power E1050 server in an initial valid order to have a fully working environment. This configuration was tested, validated, and certified by an IBM development team. The Power E1050 server model belongs to the mid-range category, and it can scale up vertically through physical add-on upgrades and resource activations of inactive processors and memory capacity, otherwise known as Capacity Upgrade on Demand (CUoD).

The Power E1050 server is 4-socket based processor where the first two sockets at a minimum must be populated. As for the memory DIMM slots, four out of 16 per socket must be inserted. The smallest set is 256 GB.

Table 1-6 lists the minimum features of a Power E1050 server configuration.

Table 1-6   Selecting the minimum configuration for the Power E1050 server

Feature	Feature Code	Feature Code description	Minimum quantity
Primary OS Feature Code	► #2146 ► #2147	► Primary OS Indicator-AIX ► Primary OS Indicator-Linux	1
Heat sink + thermal interface material (TIM) pad	► #EPLU ► #EPLV	► Front Heat Sink + TIM PAD (For MRX) ► Rear Heat Sink + TIM PAD (For MRX)	► 1 ► 1 Note: Applies to base two sockets populated.
Processor card	► #EPEU ► #EPEV ► #EPGW	► 12-core typical 3.35 - 4.0 GHZ (max) processor ► 18-core typical 3.20 - 4.0 GHZ (max) processor ► 24-core typical 2.95 - 3.90 GHZ (max) processor	Two of any processor Feature Code, and they must be the same.
Processor activation	AIX or Virtual I/O Server (VIOS): ► #EPUR ► #EPUS ► #EPYT Linux only: ► #EPUN ► #EPUP ► #EPUM	AIX or VIOS: ► One core Processor Activation for #EPEU ► One core Processor Activation for #EPEV ► One core Processor Activation for #EPGW Linux only: ► One core Processor Activation for #EPEU Linux only ► One core Processor Activation for #EPEV Linux only ► One core Processor Activation for #EPGW Linux only	AIX or VIOS: ► 12 ► 18 ► 24 Linux only: ► 12 ► 18 ► 24
Memory DIMM	► #EM75 ► #EM76 ► #EM77	► 64 GB (2 x 32 GB) DDR4 Memory DIMM ► 128 GB (2 x 64 GB) DDR4 Memory DIMM ► 256 GB (2 x 128 GB) DDR4 Memory DIMM	Four of the Feature Code for a base 2-socket.
Memory activation	#EMCP	1 GB Memory Activation for MRX	256 GB or 50% memory that is activated, whichever is higher.
NVMe backplane	#EJ0Q	10 NVMe U.2 Flash Drive bays	1

Feature	Feature Code	Feature Code description	Minimum quantity
NVMe device	▸ #EC5X ▸ #EC7T ▸ #ES1E ▸ #ES1G ▸ #ES3E	▸ Mainstream 800 GB SSD PCIe3 NVMe U.2 module for AIX or Linux ▸ 800 GB Mainstream NVMe U.2 SSD 4k for AIX or Linux ▸ Enterprise 1.6 TB SSD PCIe4 NVMe U.2 module for AIX or Linux ▸ Enterprise 3.2 TB SSD PCIe4 NVMe U.2 module for AIX or Linux ▸ Enterprise 6.4 TB SSD PCIe4 NVMe U.2 module for AIX or Linux	One of any of these Feature Codes. Note: Recommend to have two for mirrored copy. Not required if Feature Code # 0837 (SAN Boot Specify) is selected.
Network adapter	▸ #EC2U ▸ #EC66 ▸ #EN0W	▸ PCIe3 2-Port 25/10Gb NIC&ROCE SR/Cu Adapter ▸ PCIe4 2-port 100Gb ROCE EN adapter ▸ PCIe2 2-port 10/1GbE BaseT RJ45 Adapter	Choose one of any of these Feature Codes.
Power supplies	EB39	Power Supply - 2300W for Server (200 - 240 VAC)	4
Power cord	4558	Power Cord To PDU/UPS (100 - 240V/16A)	4
Language group	9300	Language Group Specify - US English	One of any language

1.6 PCIe adapter slots

The Power E1050 server has up to 11 PCIe slots in the systems drawer. A mix of PCIe Gen5 and Gen4 general-purpose hot-plug slots can deliver configuration flexibility and expandability. Two adapter slots are PCIe Gen5 8-lane, three adapter slots are PCIe Gen4 8-lane, and six adapter slots are Gen4 16-lane or Gen5 8-lane. All adapter slots are full-height, half-length in size. BSCs are used to house the adapter in the system unit for easy installation, removal, and service from the rear of the system. All the PCIe slots are single-root I/O virtualization (SR-IOV) capable.

The 16-lane slots can provide up to twice the bandwidth of the 8-lane slots because they offer twice as many PCIe lanes. PCIe Gen5 slots can support up to twice the bandwidth of PCIe Gen4 slots and up to four times the bandwidth of a PCI Gen3 slot, assuming an equivalent number of PCIe lanes. PCIe Gen1, PCIe Gen2, PCIe Gen3, PCIe Gen4, and PCIe Gen5 adapters can be plugged into a PCIe Gen5 slot, if that adapter is supported. The 16-lane slots can be used to attach PCIe Gen3 I/O expansion drawers.

Table 1-7 shows the number of slots that is supported by the number of processor modules.

Table 1-7   Available PCIe slots

Processor modules	2-socket	3-socket or 4-socket
x16 Gen5 or four slots (CAPI-capable)	2	6
x8 Gen4 slots	3	3
x8 Gen5 slots	2	2

Notes:

► The PCIe Genx8 slot, C1, is reserved for an Ethernet adapter to help ensure proper manufacture and test of the server.

► Each NVMe SSD interface is a Gen4 x4 PCIe bus. The NVMe drives can be in an OS-controlled RAID 0, RAID 1 array. Hardware RAID is not supported on the NVMe drives.

► This server has an energy-efficient design for cooling the PCIe adapter environment. The server can sense which IBM PCIe adapters are installed in their PCIe slots. If an adapter is known to require higher levels of cooling, the server automatically speeds up fans to increase airflow across the PCIe adapters.

► The terms "16-lane" and "x16" and "8-lane" and "x8" are interchangeably used in this case with the same meaning.

Figure 1-6 shows the 11 PCIe adapter slots location with labels for the Power E1050 server model.
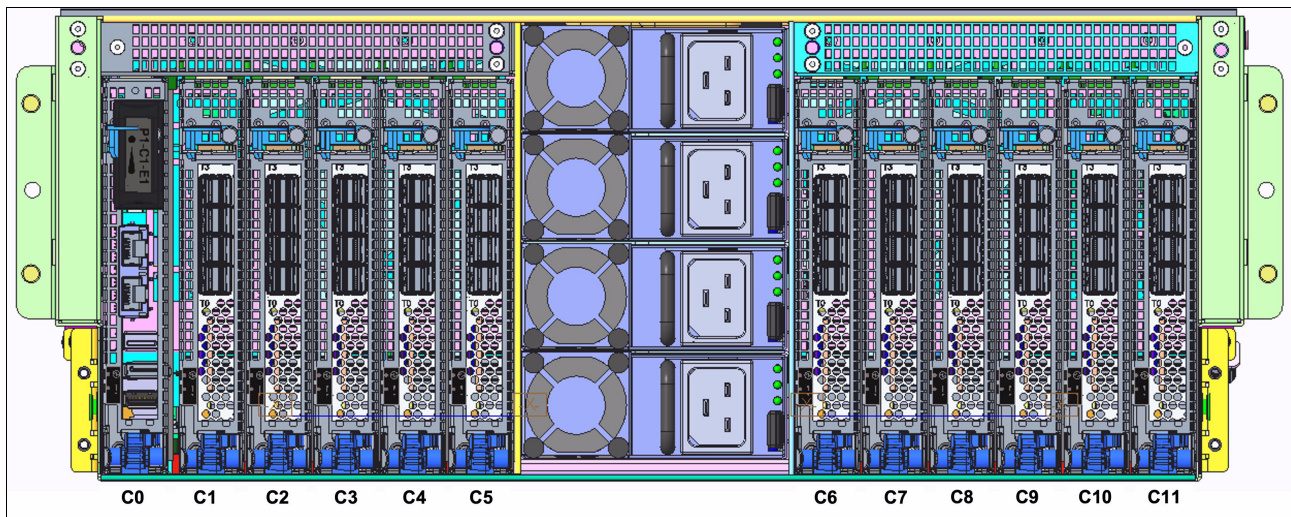
Figure 1-6   PCIe adapter slot locations on the Power E1050 server

Slot C0 is not included in the list. It is meant for only the eBMC service processor card. The total number of PCIe adapter slots that is available can be increased by adding one or more PCIe Gen3 I/O expansion drawers (#EMX0). The maximum number depends on the number of processor modules physically installed. The maximum is independent of the number of processor core activations.

Table 1-8 list the number of maximum number of I/O drawers per populated socket.

Table 1-8   I/O drawers per populated processor socket

Number of processor sockets	Maximum number of PCI I/O drawers	Number of fan-out modules
2	2	Up to 4
3	3	Up to 6
4	4	Up to 8

The connection of each (#EMXG) fanout module in a PCIe Gen3 expansion drawer requires the installation of one (#EJ08)PCIe cable adapter to be placed in one of the PCIe x16 adapter slots of the system unit. For more information about PCIe Gen3 I/O drawer, see 2.3.4, "Attachment of I/O-drawers" on page 71.

1.7 Operating system support

The Power E1050 server supports the following OSs:

► AIX

► Linux

In addition, VIOS can be installed in a special logical partition (LPAR) where its primary function is to host physical I/O adapters like network and storage connectivity, and provide virtualized I/O devices for client LPARs, such as AIX and Linux OSes.

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

The minimum supported levels of IBM AIX and Linux at the time of announcement are described in the following sections. For more information about hardware features, see IBM Power Systems Prerequisites.

IBM Power Systems Prerequisites helps to plan a successful system upgrade by providing the prerequisite information for features in use or that you plan to add to a system. It is possible to choose an MTM (9043-MRX for Power E1050) and discover all the prerequisites, the OS levels that are supported, and other information.

1.7.1 AIX operating system

At the time of announcement, Power E1050 supports the following minimum levels of AIX when installed with direct I/O connectivity:

► AIX 7.3 with the 7300-00 Technology Level and Service Pack 2 or later

► AIX 7.2 with the 7200-05 Technology Level and Service Pack 4 or later

At the time of announcement, Power E1050 supports the following minimum level of AIX when installed with virtual I/O:

► AIX 7.3 with the 7300-00 Technology Level and Service Pack 1 or later

► AIX 7.2 with the 7200-05 Technology Level and Service Pack 1 or later

► AIX 7.2 with the 7200-04 Technology Level and Service Pack 2 or later

► AIX 7.1 with the 7100-05 Technology Level and Service Pack 6 or later

Notes:

► AIX 7.2 with the 7200-04 Technology Level and Service Pack 6 is planned to be available on September 16, 2022.

► AIX 7.1 has been withdrawn from marketing since November 2021. AIX 7.2 is the minimum available version for a new software order.

AIX 7.1 instances must run in an LPAR in IBM Power8 compatibility mode with a VIOS-based virtual storage and network.

AIX 7.2 instances can use both physical and virtual I/O adapters, but must run in an LPAR in Power9 compatibility mode.

AIX 7.3 instances can use both physical and virtual I/O adapters, and can run in an LPAR in native Power10 mode.

IBM periodically releases maintenance packages (service packs (SPs) or technology levels (TLs)) for the AIX OS. For more information about these packages, and downloading and obtaining the installation packages, see Fix Central.

The Service Update Management Assistant (SUMA), which can help you automate the task of checking and downloading OS downloads, is part of the base OS. For more information about the `suma` command, see IBM Documentation.

The AIX OS software is available to order in the following editions:

► AIX Standard Edition: Stand-alone AIX OS only.

► AIX Enterprise Edition: The AIX OS plus the bundled Power software stack.

► IBM Private Cloud Edition with AIX7: An enhanced set from the AIX Enterprise Edition of the bundled software stack intended for private cloud usage. For the list of offering software that is included, see 1.9, "IBM Power solutions" on page 29.

Subscription licensing model

AIX Standard Edition, AIX Enterprise Edition, and IBM Private Cloud Edition with AIX are also available under a subscription licensing model that provides access to an IBM software maintenance for a specified subscription term (1 or 3 years). The subscription term begins on the start date and ends on the expiration date, which is reflected at the IBM ESS website.

Customers are licensed to run the product through the expiration date of the 1- or 3-year subscription term, and then can renew the subscription at the end of it to continue using the product. This model provides flexible and predictable pricing over a specific term, with lower upfront costs of acquisition.

Another benefit of this model is that the licenses are customer number entitled, which means that they are not tied to a specific hardware serial number as with a standard license grant. Therefore, the licenses can be moved between on-premises and cloud if needed, something that is becoming more of a requirement with hybrid workloads.

The product IDs for the subscription licenses are listed in Table 1-9.

Table 1-9   Subscription license product IDs (1- or 3-year terms)

Product ID	Description
5765-2B1	IBM AIX 7 Standard Edition Subscription 7.3.0
5765-2E1	IBM AIX Enterprise Edition Subscription 1.8.0
5765-2C1	IBM Private Cloud Edition with AIX 7 Subscription 1.8.0
5765-6C1	IBM Private Cloud Edition Subscription 8.0 (without AIX)

The subscription licenses are orderable through an IBM configurator. The standard AIX license grant and monthly term licenses for standard edition are still available.

1.7.2 Linux operating system distributions

Linux is an open-source, cross-platform OS that runs on numerous platforms from embedded systems to mainframe computers. It provides an UNIX like implementation across many computer architectures.

The following Linux distributions are supported on the Power E1050 server model.

Red Hat Enterprise Linux

The latest version of the RHEL distribution from Red Hat is supported in native Power10 mode, allowing it to access all the features of the Power10 processor and platform.

At the time of announcement, the Power E1050 server supports the following minimum levels of the RHEL OS:

► Red Hat Enterprise Linux 8.4 for Power Little Endian (LE) or later

► Red Hat Enterprise Linux 9.0 for Power LE or later

► Red Hat Enterprise Linux for SAP with Red Hat Enterprise Linux 8.4 for Power LE or later

Note: RHEL 9.0 for Power LE or later is supported to run once it is announced.

RHEL is sold on a subscription basis, with initial subscriptions and support that are available for 1 year, 3 years, or 5 years. Support is available either directly from Red Hat or from IBM Technical Support Services. An RHEL 8 for Power LE unit subscription covers up to four cores and up to four LPARs, and the subscription can be stacked to cover more cores and LPARs.

When a client orders RHEL from IBM, a subscription activation code is published at the IBM ESS website. After you retrieve this code from IBM ESS, use it to establish proof of entitlement and download the software from Red Hat.

SUSE Linux Enterprise Server

The latest version of the SUSE Linux Enterprise Server distribution of Linux from SUSE is supported in native Power10 mode, allowing it to access all the features of the Power10 processor and platform.

At the time of announcement, the Power E1050 server supports the following minimum levels of SUSE Linux Enterprise Server OS:

► SUSE Linux Enterprise Server 15 Service Pack 3 or later

► SUSE Linux Enterprise Server for SAP with SUSE Linux Enterprise Server 15 Service Pack 3 or later

SUSE Linux Enterprise Server is sold on a subscription basis, with initial subscriptions and support that are available for 1 year, 3 years, or 5 years. Support is available either directly from SUSE or from IBM Technical Support Services. A SUSE Linux Enterprise Server 15 unit subscription covers 1 - 2 sockets or 1 -2 LPARs, and they subscriptions can be stacked to cover more sockets and LPARs.

When a client orders SUSE Linux Enterprise Server from IBM, a subscription activation code is published at the IBM ESS website. After you retrieve this code from IBM ESS, use it to establish proof of entitlement and download the software from SUSE.

Linux and Power10 technology

The Power10 specific toolchain is available in Advance Toolchain 15.0, which allows clients and developers to use all new Power10 processor-based technology instructions when programming. The cross-module function call impact was reduced because of a new PC-relative addressing mode.

One specific benefit of Power10 technology is a 10x - 20x advantage over Power9 processor-based technology for artificial intelligence (AI) inferencing workloads because of increased memory bandwidth and new instructions. One example is the new special purpose-built Matrix Math Accelerator (MMA) that was tailored for the demands of machine learning and deep learning inference. The MMA also supports many AI data types.

Network virtualization is an area with significant evolution and improvements, which benefit virtual and containerized environments. The following recent improvements were made for Linux networking features on Power10 processor-based servers:

► SR-IOV allows virtualization of network cards at the controller level without needing to create virtual Shared Ethernet Adapters (SEAs) in the VIOS partition. It is enhanced with a virtual Network Interface Controller (vNIC), which allows data to be transferred directly from the partitions to or from the SR-IOV physical adapter without transiting through a VIOS partition.

► Hybrid Network Virtualization (HNV) allows Linux partitions to use the efficiency and performance benefits of SR-IOV logical ports and participate in mobility operations, such as active and inactive Live Partition Mobility (LPM) and Simplified Remote Restart (SRR). HNV is enabled by selecting **Migratable** when an SR-IOV logical port is configured.

Security

Security is a top priority for IBM and our distribution partners. Linux security on IBM Power is a vast topic that can be the subject of detailed separate material. However, improvements in the areas of hardening, integrity protection, performance, platform security, and certifications are introduced in this section.

Hardening and integrity protection deal with protecting the Linux kernel from unauthorized tampering while allowing upgrading and servicing of the kernel. These topics become even more important when running in a containerized environment with an immutable OS, such as CoreOS in Red Hat OpenShift.

Performance is a security topic because specific hardening mitigation strategies (for example, against side-channel attacks) can have a significant performance effect. In addition, cryptography can use significant compute cycles.

The Power E1050 features transparent memory encryption at the level of the controller, which prevents an attacker from retrieving data from physical memory or storage-class devices that are attached to the processor bus.

1.7.3 Red Hat OpenShift Container Platform

The Red Hat OpenShift Container Platform is supported to run on IBM Power servers, including the IBM Power E1050 server. Red Hat OpenShift Container Platform is a container orchestration and management platform that provides a resilient and flexible environment to develop and deploy applications. It extends the open source Kubernetes project to provide an enterprise-grade platform to run and manage workloads by using Linux container technology.

A Red Hat OpenShift Container Platform cluster is built from a number of nodes, which can run on either physical or virtual machines (VMs). A minimum of three control plane nodes are required to support the cluster management function, and at least two compute nodes are required to provide the capacity to run workloads. During installation, an extra bootstrap node is required to host the files that are required for the installation and initial setup.

The bootstrap and control plane nodes are all based on RHEL CoreOS, which is a minimal immutable container host version of the RHEL distribution that inherits the associated hardware support statements. The compute nodes can run on either RHEL or RHEL CoreOS.

Red Hat OpenShift Container Platform is available on a subscription basis, with initial subscriptions and support that are available for 1 year, 3 years, or 5 years. Support is available either directly from Red Hat or from IBM Technical Support Services. Red Hat OpenShift Container Platform subscriptions cover two processor cores each, and they can be stacked to cover more cores.

At the time of announcement, the Power E1050 server supports Red Hat OpenShift Container Platform 4.10 or later.

When a client orders Red Hat OpenShift Container Platform for Power from IBM, a subscription activation code is published at the IBM ESS website. After you retrieve this code from IBM ESS, use it to establish proof of entitlement and download the software from Red Hat.

For more information about running Red Hat OpenShift Container Platform on IBM Power, see Red Hat OpenShift documentation.

1.7.4 Virtual I/O Server

VIOS is part of the PowerVM Enterprise edition feature. VIOS is software that is installed in a special LPAR, which facilitates the sharing of physical I/O resources between client LPARs within the server. VIOS provides virtual Small Computer Serial Interface (SCSI) target, virtual FC, SEA, and PowerVM Active Memory Sharing capabilities to client LPARs within the system.

The minimum required level of VIOS for the Power E1050 server model is VIOS 3.1.3.21 or later.

IBM regularly updates the VIOS code. For more information, see IBM Fix Central.

1.7.5 Entitled Systems Support

The IBM ESS website is the IBM go-to place to view and manage IBM Power and IBM Storage software and hardware. In general, most products that are offered by IBM Systems that are purchased through IBM Digital Sales representatives or IBM Business Partners are accessed at this website when IBM e-config on Cloud is used.

The website features the following three main sections:

- ► My Entitled Software: Activities that are related to IBM Power and IBM Storage software, such as downloading licensed, free, and trial software media; placing software update orders; and managing software keys.

- ► My Entitled Hardware: Activities that are related to IBM Power and IBM Storage hardware, such as renewing Update Access Keys (UAKs); buying and using Elastic CoD; assigning or buying credits for new and existing pools in Enterprise Pools 2.0; downloading Storage CoD codes; and managing Hybrid Capacity credits.

- ► My Inventory: Activities that are related to IBM Power and IBM Storage inventory, such as browsing software licenses, software maintenance, and hardware inventory; or managing inventory retrievals by using Base Composer or generating several types of reports.

For initial access and to get more information, see IBM ESS.

> **Note:** A valid registered IBMid is required before a user can sign in to IBM ESS.

1.7.6 Update Access Keys

UAKs are a technology that was introduced with Power8 servers and adapted by the Power9 and Power10 servers. An UAK is an important parameter that is first validated by the system before an update can proceed.

UAKs have two types:

- ► System firmware UAK: This UAK is checked when system firmware updates are applied to the system. The UAK includes an expiration date. System firmware updates contain a release date. When attempting to apply a system firmware update, if the release date for the firmware updates has passed the expiration date for the UAK, the updates are not processed. As UAKs expire, they must be replaced by using either the HMC or the Advanced System Management Interface (ASMI) on the service processor.

- ► AIX UAK: This UAK checks for an active Software Maintenance Agreement (SWMA) when updating the AIX OS. The server uses an AIX update that includes the expiration date for the associated SWMA agreement. The server periodically checks and informs the administrator about AIX UAKs that are about to expire, that have expired, or that are missing. The AIX UAK is a CoD code. There is a single AIX UAK per server (not one per LPAR). As the AIX UAK expires, it must be replaced by using either the HMC or ASMI GUI.

By default, newly delivered systems include an UAK that often expires after 3 years. Thereafter, the UAK can be extended every 6 months, but only if an IBM maintenance contract exists. The contract can be verified at the IBM ESS website (see 1.7.5, "Entitled Systems Support" on page 19).

Figure 1-7 shows an example of viewing the system firmware UAK in the HMC.

Figure 1-7   UAK view from the HMC

Figure 1-8 shows another example of viewing the access key in ASMI.

Figure 1-8   Access key view from the ASMI

1.8 Hardware Management Console overview

The HMC can be a hardware appliance or virtual appliance that can be used to configure and manage your systems. The HMC connects to one or more managed systems and provides capabilities for the following primary functions:

- ► Systems Management functions, such as power off, power on, system settings, CoD, enterprise pools, shared processor pools (SPPs), Performance and Capacity Monitoring, and starting ASMI for managed systems.

- ► Delivers virtualization management through support for creating, managing, and deleting LPARs, LPM, and Remote Restart; configuring SRIOV; and managing VIOSs, dynamic resource allocation, and OS terminals.

- ► Acts as the Service Focal Point (SFP) for systems and supports service functions, including Call Home, dump management, guided repair and verify, concurrent firmware updates for managed systems, and around-the-clock error reporting with Electronic Service Agent (ESA) for faster support.

- ► Provides appliance management capabilities for configuring the network and users on the HMC, and updating and upgrading the HMC.

1.8.1 HMC 7063-CR2

The 7063-CR2 IBM Power HMC (see Figure 1-9) is a second-generation IBM Power processor-based HMC.

Figure 1-9   HMC 7063-CR2

The CR2 model includes the following features:

- ► Six-core Power9 130 W processor chip

- ► 64 GB (4x16 GB) or 128 GB (4x32 GB) of memory RAM

- ► 1.8 TB with RAID 1 protection of internal disk capacity

- ► Four-port 1 Gb Ethernet (RH-45), 2-port 10 Gb Ethernet (RJ-45), two USB 3.0 ports (front side) and two USB 3.0 ports (rear side), and 1 Gb Intelligent Platform Management Interface (IPMI) Ethernet (RJ-45)

- ► Two 900 W power supply units (PSUs)

- Remote Management Service: IPMI port (OpenBMC) and Redfish application programming interface (API)
- Base Warranty is 1 year 9x5 with available optional upgrades

A USB Smart Drive is not included.

> **Note:** The recovery media for V10R1 is the same for 7063-CR2 and 7063-CR1.

The 7063-CR2 is compatible with flat panel console kits 7316-TF3, TF4, and TF5.

> **Note:** The 7316-TF3 and TF4 were withdrawn from marketing.

1.8.2 Virtual HMC

Initially, the HMC was sold only as a hardware appliance, including the HMC firmware. However, IBM extended this offering to allow the purchase of the hardware appliance and a virtual appliance that can be deployed on ppc64le architectures and deployed on x86 platforms.

Any customer with a valid contract can download this offering from the IBM ESS website, or this offering can be included with an initial Power E1050 order.

The virtual HMC supports the following hypervisors:

- On x86 processor-based servers:
  - Kernel-based Virtual Machine
  - Xen
  - VMware
- On Power processor-based servers: PowerVM

The following minimum requirements must be met to install the virtual HMC:

- 16 GB of memory
- Four virtual processors
- Two network interfaces (a maximum of four is allowed)
- One disk drive (500 GB available disk drive)

For an initial Power E1050 order with the IBM configurator (e-config), you can find the HMC virtual appliance by selecting **Add software** → **Other System Offering**s (as product selections) and then select either of the following items:

- 5765-VHP for IBM HMC Virtual Appliance for Power V10
- 5765-VHX for IBM HMC Virtual Appliance x86 V10

For more information about an overview of the Virtual HMC, see this web page.

For more information about how to install the virtual HMC appliance and all requirements, see IBM Documentation.

1.8.3 BMC network connectivity rules for the 7063-CR2

The 7063-CR2 HMC features a baseboard management controller (BMC), which is a specialized service processor that monitors the physical state of the system by using sensors. The OpenBMC that is used on 7063-CR2 provides a GUI that can be accessed from a workstation that has network connectivity to the BMC. This connection requires an Ethernet port to be configured for use by the BMC.

> **Note:** This section describes the BMC of the hardware HMC 7063-CR2. The Power E1050 also uses an eBMC for the systems management, as described in 2.5, "The enterprise Baseboard Management Controller" on page 73.

The 7063-CR2 provides two network interfaces (eth0 and eth1) for configuring network connectivity for BMC on the appliance.

Each interface maps to a different physical port on the system. Different management tools name the interfaces differently. The HMC task **Console Management → Console Settings → Change BMC/IPMI Network Settings** modifies only the Dedicated interface.

The BMC ports are listed in Table 1-10.

Table 1-10   BMC ports

Management tool	Logical port	Shared or dedicated	CR2 physical port
OpenBMC UI	eth0	Shared	eth0
OpenBMC UI	eth1	Dedicated	Management port only
ipmitool	lan1	Shared	eth0
ipmitool	lan2	Dedicated	Management port only
HMC task (change BMC or IMPI Network settings	lan2	Dedicated	Management port only

Figure 1-10 shows the BMC interfaces of the HMC.
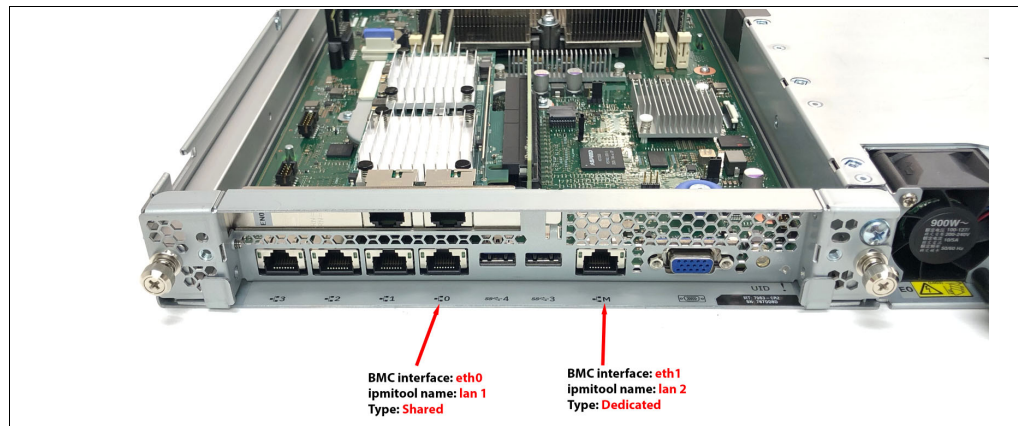
Figure 1-10   BMC interfaces

The main difference is the shared and dedicated interface to the BMC can coexist. Each one has its own LAN number and physical port. Ideally, the customer configures one port, but both can be configured. The rules to connecting IBM Power servers to the HMC remain the same as previous versions.

1.8.4 High availability HMC configuration

For the best manageability and redundancy, a dual HMC configuration is suggested. This configuration can be two hardware appliances, but also one hardware appliance and one virtual appliance or two virtual appliances.

The following requirements must be met:

► Two HMCs are at the same version.

► The HMCs use different subnets to connect to the Flexible Service Processors (FSPs).

► The HMCs can communicate with the servers' partitions over a public network to allow for full synchronization and function.

1.8.5 HMC code level requirements for Power E1050

The minimum required HMC version for Power E1050 is V10R1 M1020. V10R1 is supported only on 7063-CR1, 7063-CR2, and Virtual HMC appliances. It is *not* supported on 7042 machine types. HMC with V10R1 cannot manage Power7 processor-based systems.

An HMC that is running V10R1 M1020 includes the following features:

► HMC OS Secure Boot support for the 7063-CR2 appliance.

► Ability to configure login retries and suspended time and support for inactivity expiration in the password policy.

► Ability to specify the HMC location and data replication for groups.

► VIOS Management Enhancements:

 – Prepare for VIOS maintenance:

 • Validation for redundancy for the storage and network that is provided by VIOS to customer partitions.

 • Switch path of redundant storage and network to start failover.

 • Roll back to original configuration on failure of preparation.

 • Audit various validation and preparation steps that are performed.

 • Report any failure that is seen during preparation.

 – CLI and scheduled operations support for VIOS backup or restore VIOS configurations and SSP configurations

 – Option to back up or restore a Shared Storage Pool configuration in HMC

 – Options to import or export the backup files to external storage

 – Option to fail over all vNICs from one VIOS to another one

► Supports 128 and 256 MB local memory bus (LMB) sizes.

► Automatically choose the fastest network for LPM memory transfer.

- ► HMC user experience enhancements:
  - – Usability and performance improvements
  - – Enhancements to help connect global search
  - – Quick view of serviceable events
  - – More progress information for UI wizards
- ► Allow LPM/Remote Restart when a virtual optical device is assigned to a partition.
- ► UAK support.
- ► Configure the Virtualization Management Interface (VMI) connection to the Power Hypervisor for the Power E1050 and for the Power10 scale-out servers.
- ► Scheduled operation function: In the ESA, there is a new feature that allows customers to receive message alerts only if scheduled operations fail (see Figure 1-11).
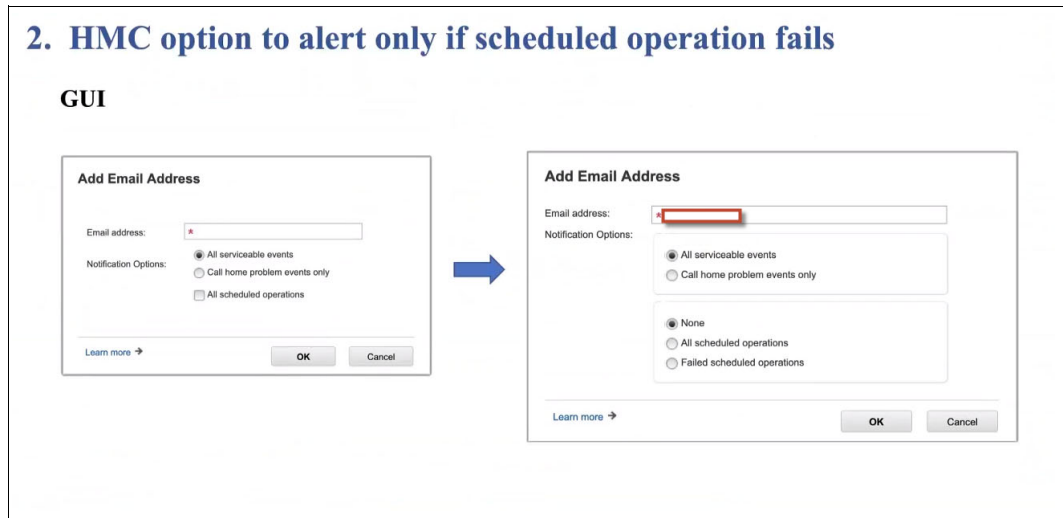
Figure 1-11   HMC alert feature

Log retention of the HMC audit trail also is increased.

1.8.6 Configuring the VMI

Starting with the Power10 scale-out servers and the Power E1050 server, IBM Power servers are migrating to an industry-standard service processor chipset that is known as eBMC. As part of the eBMC transition, the virtualization management communication path was removed from the service processor.

This path is something specific to PowerVM with the HMC. eBMC is based on of the OpenBMC code base, which is platform-neutral. Developers wanted to minimize the number of PowerVM specific functions in eBMC.

The industry-standard service processor communication protocols and the performance of the service processor do not lend themselves well to supporting the virtualization management traffic pass-through like with the FSP in the past. VMI was invented to satisfy the need for a new network connection point for the HMC.

The VMI design is a combination of HMC and Power Hypervisor support for a new VMI network endpoint for handling virtualization management. The service processor now has two IP addresses for each system management network port: one for system management (eBMC), and one for virtualization management (VMI).

Note: Each port also has two MAC addresses, that is, BMC and VMI each have one.

The eBMC IP address is the equivalent of the FSP IP address in previous generations of IBM Power servers. The WebUI, Representational State Transfer (REST) interfaces, and others, all use the eBMC IP address. This IP address is the only one that users interact with directly.

The VMI IP address is used for virtualization management. This IP address is the one that the HMC used to communicate with Power Hypervisor for partition management and consoles. Users do not interact directly with this IP address. From a customer perspective, other than having two IP addresses on the service network instead of one, there is no difference from an HMC user perspective.

All traffic between the HMC and VMI is encrypted with TLS by using a system unique certificate.

Figure 1-12 shows a dual HMC connection to the eBMC of an Power E1050 server.
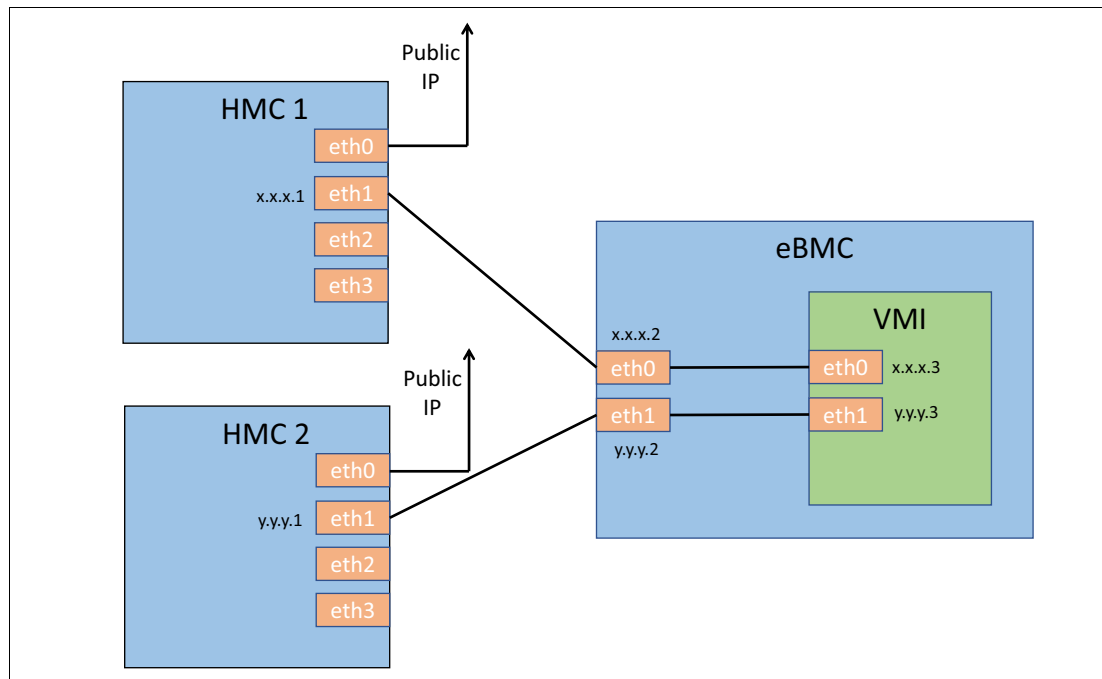
Figure 1-12   Dual HMC connection to a Power E1050 server

VMI supports both DHCP and static IP address configurations. After the server is connected to the HMC and configured by setting the access passwords, you may configure the VMI. To configure the VMI, click **VMI Configuration** in the HMC GUI, as shown in Figure 1-13. In the configuration dialog box, you may switch between DHCP or a static IP address. The default VMI connection setting (not the eBMC connection) for a new Power E1050 server is static, so if you want to use DHCP, you must change the configuration from static to DHCP before you power on the server.
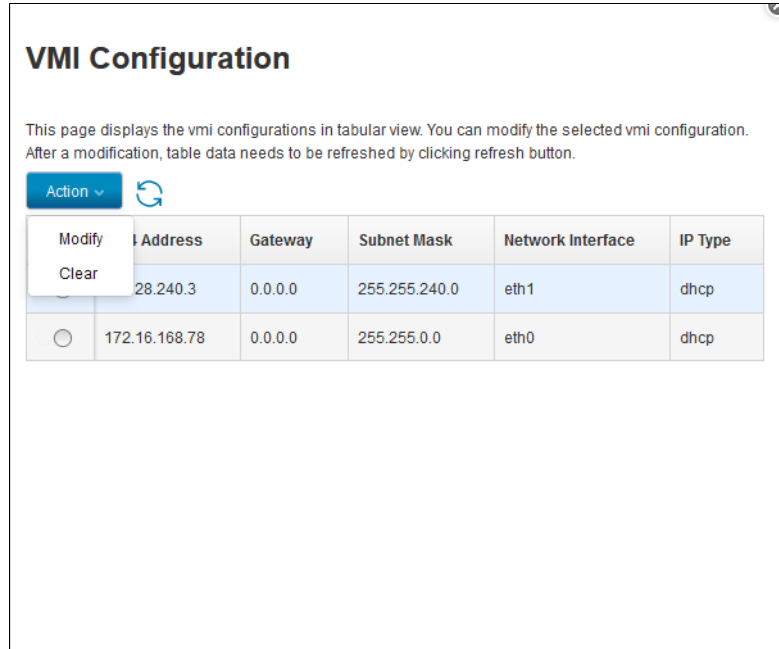
Figure 1-13   VMI configuration of a Power E1050 server

Here is a summary for configuring a new server that has factory settings:

1. Connect the Ethernet cable from the eBMC port to the internal HMC network.

2. Plug in the power cables. The eBMC starts and obtains the IP address configuration from the DHCP server on the HMCs.

3. Enter the access password or, if HMC auto-discovers, the default credentials are used.

4. The server shows as Power Off, but it is now in a manageable state.

5. Configure the VMI to change from static to DHCP, as shown in Figure 1-13.

6. Power on the server. The VMI obtains its IP address, and the HMC to VMI connection is established automatically.

1.8.7 HMC currency

In recent years, cybersecurity emerged as a national security issue and an increasingly critical concern for CIOs and enterprise IT managers.

The IBM Power processor-based architecture always ranked highly in terms of end-to-end security, which is why it remains a platform of choice for mission-critical enterprise workloads.

A key aspect of maintaining a secure IBM Power environment is ensuring that the HMC (or virtual HMC) is current and fully supported (including hardware, software, and IBM Power firmware updates).

Outdated or unsupported HMCs represent a technology risk that can quickly and easily be mitigated by upgrading to a current release.

1.9 IBM Power solutions

The Power E1050 server comes cloud-enabled with integrated PowerVM Enterprise capabilities.

1.9.1 IBM Power Private Cloud Solution with Dynamic Capacity

The IBM Power Private Cloud Solution with Dynamic Capacity is an infrastructure offering that enables you to take advantage of cloud agility and economics while getting the same business continuity and flexibility that you already enjoy from IBM Power servers.

The IBM Power Private Cloud Solution offers:

► Cost optimization with pay-for-use pricing

► Automated and consistent IT management with Red Hat Ansible for IBM Power

► IBM Proactive Support for IBM Power services

► IBM Systems Lab Services Assessment and implementation assistance

Both Elastic and Shared Utility Capacity options are available on all Power E1050 servers through the Virtual Capacity (4586-COD) MTM and the IBM ESS website.

Elastic Capacity on the Power E1050 server enables you to deploy pay-for-use consumption of processor, memory, and supported OSs.

Shared Utility Capacity on Power E1050 servers provides enhanced multisystem resource sharing and by-the-minute tracking and consumption of compute resources across a collection of systems within a Power Enterprise Pools 2.0 (PEP2). Shared Utility Capacity delivers a complete range of flexibility to tailor initial system configurations with the right mix of purchased and pay-for-use consumption of processor, memory, and software across a collection of Power E1050 and Power E950 servers.

Metered Capacity is the extra installed processor and memory resource above each system's Base Capacity. It is activated and made available for immediate use when a pool is started, and then it is monitored by the minute by an IBM Cloud® Management Console (IBM CMC).

For more information, see IBM Power Enterprise Pools 2.0.

Metered resource usage is charged only for minutes that exceed the pool's aggregate base resources, and usage charges are debited in real time against your purchased Capacity Credits (5819-CRD) on account.

IBM offers a Private Cloud Capacity Assessment and Implementation Service that is performed by IBM Systems Lab Services professionals, which can be preselected at time of purchase or requested for qualifying Power E1050 servers.

1.9.2 IBM Private Cloud Edition 1.8

IBM Private Cloud Edition is a complete package that adds flexible licensing models in the cloud. It helps you to rapidly deploy multi-cloud infrastructures with a compelling set of cloud-enabled capabilities. The IBM Power Enterprise Cloud Edition primarily provides value for clients that use both AIX and Linux on Power, with simplified licensing models and advanced features.

The IBM Private Cloud Edition (5765-ECB) includes:

► IBM PowerSC 2.1
► IBM PowerSC Multi-Factor Authentication (MFA)
► IBM Cloud PowerVC for Private Cloud
► IBM VM Recovery Manager DR
► IBM Tivoli® Monitoring

If you use IBM AIX as the primary OS, there is a specific offering for it: IBM Private Cloud Edition with AIX 7 1.8.0 (5765-CBA). The offering includes:

► IBM AIX 7.3 or IBM AIX 7.2
► IBM PowerSC 2.1
► IBM PowerSC MFA
► IBM Cloud PowerVC for Private Cloud
► IBM VM Recovery Manager DR
► IBM Tivoli Monitoring

IBM PowerSC 2.1

IBM PowerSC 2.1 (5765-SC2) provides a security and compliance solution that is optimized for virtualized environments on IBM Power running IBM PowerVM and IBM AIX, or Linux on Power. Security control and compliance are some of the key components that are needed to defend virtualized data centers and a cloud infrastructure against evolving threats.

The PowerSC 2.1 product contains the following enhancements:

► A blacklisting antivirus feature to allow selective, on-demand hash searches across endpoints that are managed through PowerSC
► Linux on Intel support for PowerSC endpoints, including MFA on IBM Power
► Single sign-on (SSO) support

  Users can log in to PowerSC through SSO with their OpenID Connect (OIDC) Enterprise identity provider and MFA, enabling integration with any application user interface (UI).

► MFA support for Rivest-Shamir-Adleman (RSA) web API

  User MFA includes RSA through the web API, and it no longer employs the access control entry (ACE) protocol.

► User-defined alt-disk for TL and SP updates

  Users can specify alt-disk through Trusted Network Connect (TNC) for TL and SP updates on AIX endpoints.

For more information, see the PowerSC 2.1 sales manual.

IBM PowerSC Multi-Factor Authentication

IBM PowerSC MFA provides alternative authentication mechanisms for systems that are used with RSA SecurID-based authentication systems, and certificate authentication options such as Common Access Card (CAC) and Personal Identification Verification (PIV) cards. IBM PowerSC MFA allows the use of alternative authentication mechanisms instead of the standard password.

You can use IBM PowerSC MFA with many applications, such as Remote Shell (rsh), Telnet, and Secure Shell (SSH).

IBM PowerSC MFA raises the level of assurance of your mission-critical systems with a flexible and tightly integrated MFA solution for IBM AIX and Linux on Power virtual workloads running on IBM Power servers.

For more information, see the PowerSC MFA sales manual.

IBM PowerVC for Private Cloud

IBM PowerVC for Private Cloud works with IBM Power Virtualization Center to provide an end-to-end cloud solution. PowerVC for Private Cloud allows you to provision workloads and manage virtual images.

With PowerVC for Private Cloud, you can perform several operations, depending on your role within a project.

Administrators can perform the following tasks:

► Creating projects and assigning images to projects to give team-specific access to images
► Setting policies on projects to specify default virtual machine (VM) expiration, and so on
► Authorizing users to projects
► Creating expiration policies to reduce abandoned VMs
► Specifying which actions require approvals and approving requests
► Creating, editing, and deleting deployment templates
► Deploying an image from a deployment template
► Dispositioning requests
► Performing lifecycle operations on VMs, such as capture, start, stop, delete, resume, and resize
► Monitoring usage (metering) data across the project or per user

Users can perform the following tasks on resources to which they are authorized. Some actions might require administrator approval. When a user tries to perform a task for which approval is required, the task moves to the request queue before it is performed (or rejected).

► Performing lifecycle operations on VMs, such as capture, start, stop, delete, resume, and resize
► Deploying an image from a deployment template
► Viewing and withdrawing outstanding requests
► Requesting VM expiration extension
► Viewing their usage data

PowerVC 2.0 UI

IBM Power Virtualization Center 2.0 introduces a new UI that is based on the Carbon framework. Carbon is the IBM open-source design system for products and digital experiences. With the IBM Design Language as its foundation, the system consists of working code, design tools and resources, human interface guidelines, and a vibrant community of contributors.

IBM Power Virtualization Center 2.0 comes with a new UI, and many new features and enhancements.

Because IBM Power Virtualization Center is built on the OpenStack technology, you might see some terminology in messages or other text that is not the same as what you see elsewhere in PowerVC. There is also some terminology that might be different from what you are used to seeing in other IBM Power products.

Feature support for PowerVC editions

PowerVC offers different functions depending on the edition that you have installed and the hypervisor that are you are using to manage your systems.

IBM Cloud PowerVC Manager includes all the functions of the PowerVC Standard Edition plus the following features:

► A self-service portal that allows the provisioning of new VMs without direct system administrator intervention. An option is for policy approvals for the requests that are received from the self-service portal.

► Deploy templates that simplify cloud deployments.

► Cloud management policies that simplify management of cloud deployments.

► Metering data that can be used for chargeback.

For more information, see the PowerVC 2.0 sales manual.

IBM VM Recovery Manager DR

IBM VM Recovery Manager DR (5765-DRG) is an automated DR solution that enables IBM Power users to achieve low recovery times for both planned and unplanned outages. It introduces support for more storage replication solutions and support for an extra guest OS, which broadens the offering's applicability to a wider range of client requirements.

IBM VM Recovery Manager DR offers support for:

► IBM DS8000® Global Mirror

► IBM SAN Volume Controller (SVC), and IBM Storwize® Metro and Global Mirror

► Extended Memory Controller Symmetrix Remote Data Facility (SRDF) synchronous replication

► Boot device selection for IBM Power8 processor-based systems

For more information, see the VMRM-DR sales manual.

IBM Tivoli Monitoring

IBM Tivoli Monitoring products monitor the performance and availability of distributed OSs and applications. These products are based on a set of common service components that are collectively known as IBM Tivoli Management Services. Tivoli Management Services components provide security, data transfer and storage, notification mechanisms, UI presentation, and communication services in an agent-server-client architecture.

1.10 IBM Power platform modernization

Cloud capabilities are a prerequisite for using enterprise-level IT resources. There is a rich infrastructure around IBM Power to help modernize services with the strategic initiatives of your business.

The most important products are:

► IBM Power Virtual Servers
► Red Hat OpenShift Container Platform for Power

1.10.1 IBM Power Virtual Servers

IBM Power Virtual Servers on IBM Cloud is an infrastructure as a service (IaaS) offering that you can use to deploy a virtual server, also known as an LPAR, in a matter of minutes. IBM Power clients who have typically relied upon an on-premises only infrastructure now can quickly and economically extend their Power IT resources into the cloud. Using IBM Power Virtual Servers on IBM Cloud is an alternative to avoid the large capital expense or added risk when moving your essential workloads to a public cloud.

An IBM Cloud integrates your IBM AIX capabilities into the IBM Cloud experience, which means you get fast, self-service provisioning, flexible management both on-premises and off, and access to a stack of enterprise IBM Cloud services all with pay-as-you-use billing that lets you easily scale up and out.

You can quickly deploy an IBM Power Virtual Servers on IBM Cloud instance to meet your specific business needs. With IBM Power Virtual Servers on IBM Cloud, you can create a hybrid cloud environment that allows you to easily control workload demands.

For more information, see IBM Power Systems Virtual Servers-Getting started.

1.10.2 Red Hat OpenShift Container Platform for Power

Red Hat OpenShift Container Platform for Power (5639-OCP) provides a secure, enterprise-grade platform for on-premises, private platform-as-a-service clouds on IBM Power servers. It brings together industry-leading container orchestration from Kubernetes, advanced application build and delivery automation, and RHEL certified containers for IBM Power.

Red Hat OpenShift Container Platform for Power brings developers and IT operations together on a common platform. It provides applications, platforms, and services for creating and delivering cloud-native applications and management so IT can ensure that the environment is secure and available.

Red Hat OpenShift Container Platform for Power provides enterprises the same functions as the Red Hat OpenShift Container Platform offering on other platforms.

Key features include:

► A self-service environment for application and development teams.
► A pluggable architecture that supports a choice of container run times, networking, storage, Continuous Integration / Continuous Deployment (CI-CD), and more.
► Ability to automate routine tasks for application teams.

Red Hat OpenShift Container Platform subscriptions are offered in two core increments that are designed to run in a virtual guest. The purchase of a separate RHEL subscription (5639-RLE) is *required* as the base OS.

For more information, see Red Hat OpenShift Container Platform for Power.

1.10.3 Hybrid Cloud Management Edition

Many enterprises continue to maintain hybrid cloud environments because they host many core business workloads on-premises while creating or migrating newer workloads to public cloud environments. Managing these environments can be a daunting task. Organizations need the right tools to tackle the challenges that are posed by these heterogeneous environments to accomplish their objectives.

Collectively, the capabilities that are listed in this section work together to create a consistent management platform between client data centers, public cloud providers, and multiple hardware platforms (fully inclusive of IBM Power) to provide all the necessary elements for a comprehensive hybrid cloud platform.

The capabilities are:

► IBM Cloud Pak® for Watson AIOps
► IBM Observability by IBM Instana®
► IBM Turbonomic® (an IBM Company)
► Red Hat Advanced Cluster Management for Kubernetes (RHACM)
► IBM Power servers
► IBM Power Virtual Servers

Architecture and technical overview

This chapter describes the architecture and technical overview of the IBM Power E1050 server.

Figure 2-1 shows an architecture overview diagram of the server.
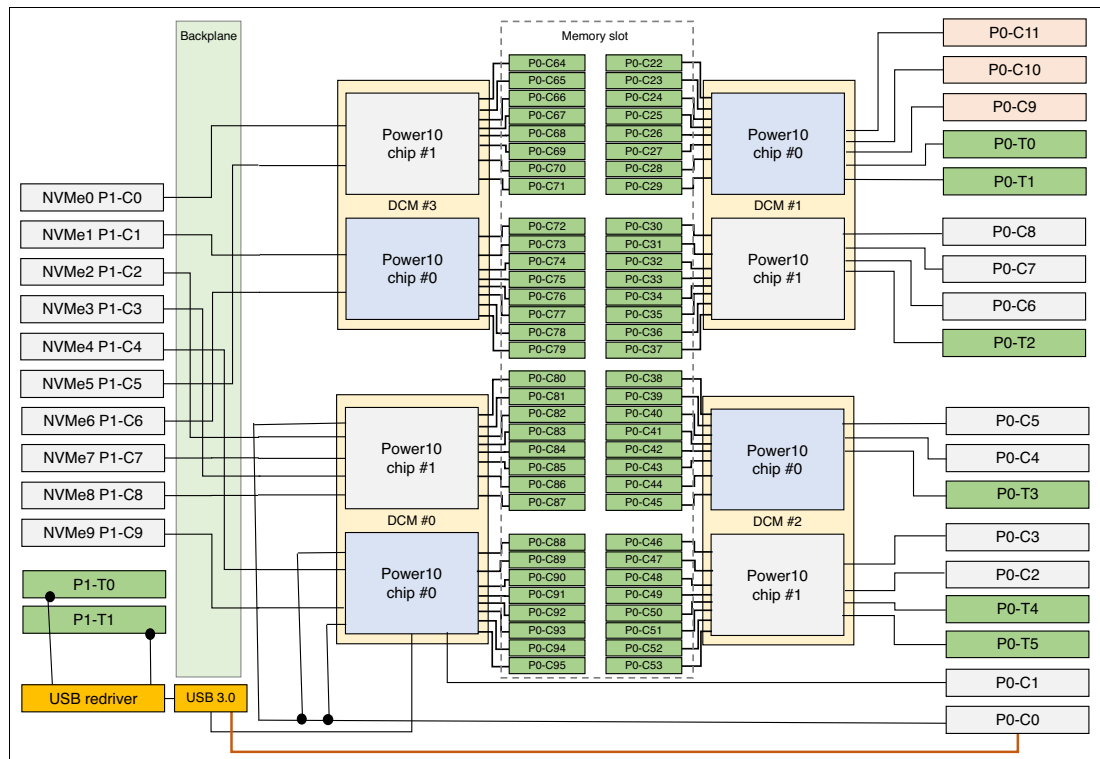
Figure 2-1   Power E1050 architecture overview

2.1 The IBM Power10 processor

The IBM Power10 processor was introduced to the general public on August 17, 2020 at the 32nd HOT CHIPS[1] semiconductor conference. At that meeting, the new capabilities and features of the latest IBM Power processor microarchitecture and the IBM Power instruction set architecture (ISA) v3.1B were revealed and categorized according to the following Power10 processor design priority focus areas:

► Data plane bandwidth focus area

   Terabyte per second signaling bandwidth on processor functional interfaces, petabyte system memory capacities, 16-socket symmetric multiprocessing (SMP) scalability, and memory clustering and memory inception capability

► Powerful enterprise core focus area

   New core micro-architecture, flexibility, larger caches, and reduced latencies

► End-to-end security focus area

   Hardware-enabled security features that are co-optimized with PowerVM hypervisor support

► Energy-efficiency focus area

   Up to threefold energy-efficiency improvement in comparison to Power9 processor technology

► Artificial intelligence (AI)-infused core focus area

   A 10 - 20x matrix-math performance improvement per socket compared to the IBM Power9 processor technology capability

The remainder of this section provides more specific information about the Power10 processor technology as it is used in the Power E1050 server.

The *IBM Power10 Processor* session material that was presented at the 32nd HOT CHIPS conference is available through the HC32 conference proceedings archive at this web page.

2.1.1 Power10 processor overview

The Power10 processor is a superscalar symmetric multiprocessor that is manufactured in complimentary metal-oxide-semiconductor (CMOS) 7-nm lithography with 18 layers of metal. The processor contains up to 15 cores that support eight simultaneous multithreading (SMT8) independent execution contexts.

Each core has private access to 2 MB L2 cache and local access to 8 MB of L3 cache capacity. The local L3 cache region of a specific core also is accessible from all other cores on the processor chip. The cores of one Power10 processor share up to 120 MB of latency optimized non-uniform cache access (NUCA) L3 cache.

---

[1] https://hotchips.org/

The processor supports the following three distinct functional interfaces, which can run with a signaling rate of up to 32 GTps[2]:

► Open Memory Interface (OMI)

The Power10 processor has eight memory controller unit (MCU) channels that support one OMI port with two OMI links each.[3] One OMI link aggregates eight lanes running at 32 GTps speed and connects to one memory buffer-based Differential Dual Inline Memory Module (DDIMM) slot to access main memory. Physically, the OMI interface is implemented in two separate die areas of eight OMI links each. The maximum theoretical full-duplex bandwidth aggregated over all 128 OMI lanes is 1 TBps.

► SMP fabric interconnect (Power A-bus/X-bus/OpenCAPI/Networking (PowerAXON))

A total of 144 lanes are available in the Power10 processor to facilitate the connectivity to other processors in an SMP architecture configuration. Each SMP connection requires 18 lanes, that is, eight data lanes plus one spare lane per direction (2 x(8+1)). Thus, the processor can support a maximum of eight SMP connections with at total of 128 data lanes per processor. This configuration yields a maximum theoretical full-duplex bandwidth aggregated over all SMP connections of 1 TBps.

The generic nature of the interface implementation also allows you to use 128 data lanes to potentially connect accelerator or memory devices through the OpenCAPI protocols. Also, this implementation can support memory cluster and memory interception architectures.

Because of the versatile characteristic of this technology, it is also referred to as *PowerAXON* interface.[4] The OpenCAPI and the memory clustering and memory interception use cases can be pursued in the future, and they are currently not used by th available technology products.

► Peripheral Component Interconnect Express (PCIe) Version 5.0 interface

To support external I/O connectivity and access to internal storage devices, the Power10 processor provides differential PCIe 5.0 interface busses (PCIe Gen 5) with a total of 32 lanes. The lanes are grouped in two sets of 16 lanes that can be used in one of the following configurations:

– One x16 PCIe Gen 4
– Two x8 PCIe Gen 4
– One x8, and two x4 PCIe Gen 4
– One x8 PCIe Gen 5, and one x8 PCIe Gen 4
– One x8 PCIe Gen 5, and two x4 PCIe Gen 4

---

[2] Giga-transfers per second (GTps).
[3] The OMI links are also referred to as OMI subchannels.
[4] A-busses and X-busses provide SMP fabric ports, which are used between central electronic complex (CEC) drawers or within CEC drawers respectively.

Figure 2-2 shows the Power10 processor die with several functional units that are labeled. Sixteen SMT8 processor cores are shown, but the dual-chip module (DCM) with two Power10 processors provides 12-, 18-, or 24-core for Power E1050 server configurations.
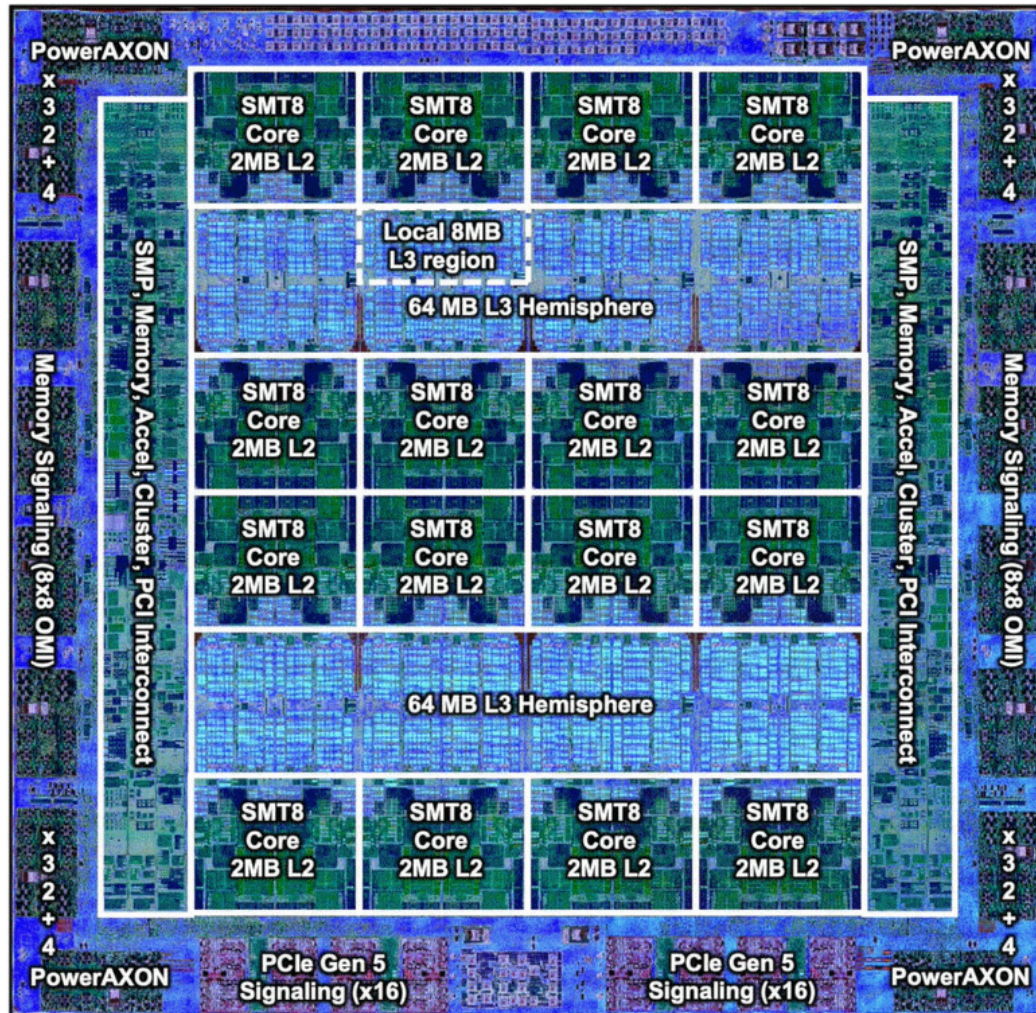
Figure 2-2   The Power10 processor chip (die photo courtesy of Samsung Foundry)

Important Power10 processor characteristics are listed in Table 2-1.

Table 2-1   Summary of the Power10 processor chip and processor core technology

Technology	Power10 processor chip
Processor die size	602 mm^2
Fabrication technology	► CMOS[a] 7-nm lithography ► 18 layers of metal
Maximum processor cores per chip	15
Maximum execution threads per core / chip	8 / 120
Maximum L2 cache core	2 MB
Maximum On-chip L3 cache per core / chip	8 MB / 120 MB

Technology	Power10 processor chip
Number of transistors	18 billion
Processor compatibility modes	Support for Power ISA[b] of Power8 and Power9

a. CMOS
b. Power ISA

The Power10 processor can be packaged as single-chip module (SCM) or DCM. The Power E1050 server implements the DCM version. The DCM contains two Power10 processors plus more logic that is needed to facilitate power supply and external connectivity to the module.

Figure 2-3 shows the logical diagram of the Power10 DCM.
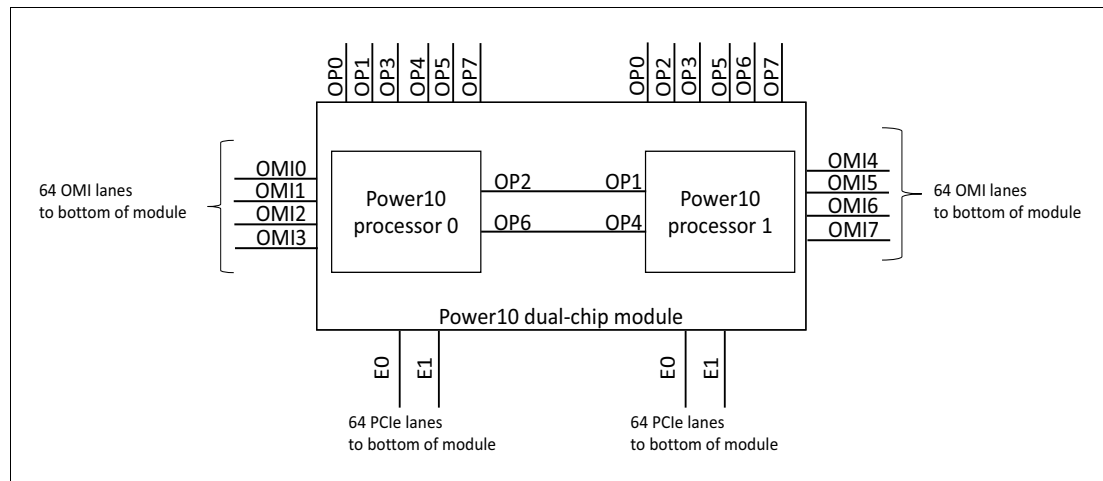
Figure 2-3   Power10 dual-chip module

2.1.2 Dual-chip modules for Power E1050 server

Power E1050 can configure a maximum of four DCMs when the 4-socket (4S) configuration is requested. Power E1050 also offers 2-socket and 3-socket configurations. Eight out of 16 memory OMI busses per Power10 chip are brought out to the module pins for a total of 16 OMI busses per DCM.

Eight OP (SMP) busses from each chip are bought to DCM module pins. Each chip has two x32 PCIe busses brought to DCM module pins.

The details of all busses that are brought out to DCM modules pins are shown in Figure 2-3.

2.1.3 Power10 processor core

The Power10 processor core inherits the modular architecture of the Power9 processor core, but the re-designed and enhanced micro-architecture increases the processor core performance and processing efficiency. The peak computational throughput is markedly improved by new execution capabilities and optimized cache bandwidth characteristics. Extra Matrix Math Accelerator (MMA) engines can deliver performance gains for machine learning, particularly for AI inferencing workloads.

The Power E1050 server uses the Power10 enterprise-class processor variant in which each core can run with up to eight independent hardware threads. If all threads are active, the mode of operation is referred to as SMT8 mode. A Power10 core with SMT8 capability is named a Power10 SMT8 core or SMT8 core for short. The Power10 core also supports modes with four active threads (SMT4), two active threads (SMT2), and one single active thread (single-threaded (ST)).

The SMT8 core includes two execution resource domains. Each domain provides the functional units to service up to four hardware threads.

Figure 2-4 shows the functional units of an SMT8 core where all eight threads are active. The two execution resource domains are highlighted with colored backgrounds in two different shades of blue.
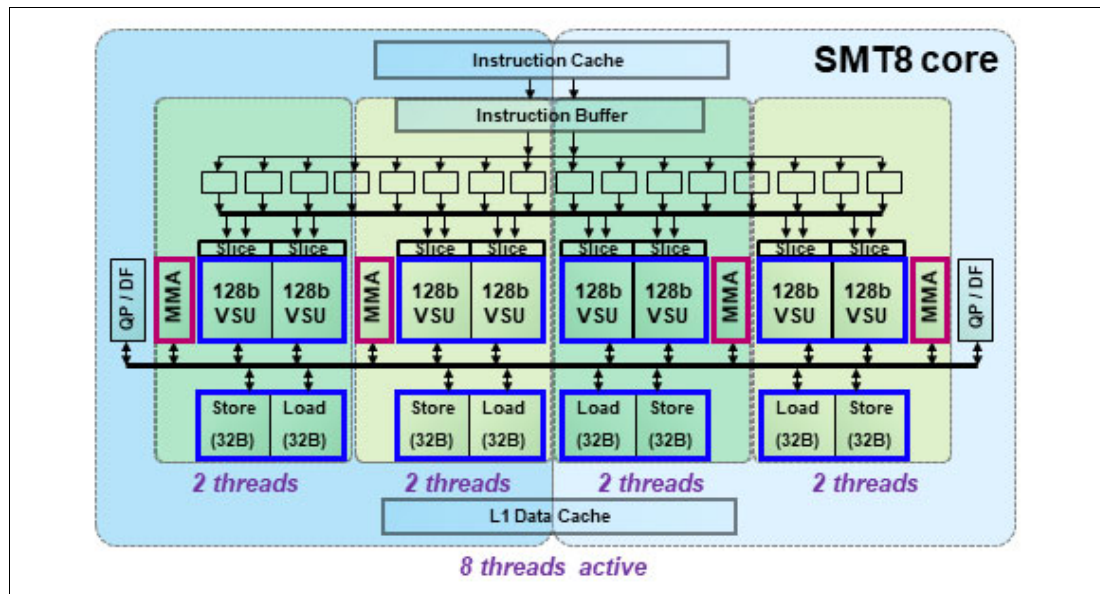
Figure 2-4   Power10 SMT8 core

Each of the two execution resource domains supports 1 - 4 threads and includes four vector scalar units (VSUs) of 128-bit width, two MMAs, and one quad-precision floating-point (QP) and decimal floating-point (DF) unit.

One VSU and the directly associated logic are called an execution *slice*. Two neighboring slices can also be used as a combined execution resource, which is then named *super-slice*. When operating in SMT8 mode, eight simultaneous multithreading (SMT) threads are subdivided in pairs that collectively run on two adjacent slices, as indicated through colored backgrounds in different shades of green.

In SMT4 or lower thread modes, 1 - 2 threads each share a four-slice resource domain. Figure 2-4 on page 40 also indicates other essential resources that are shared among the SMT threads, such as an instruction cache, an instruction buffer, and an L1 data cache.

The SMT8 core supports automatic workload balancing to change the operational SMT thread level. Depending on the workload characteristics, the number of threads that is running on one chiplet can be reduced from four to two and even further to only one active thread. An individual thread can benefit in terms of performance if fewer threads run against the core's executions resources.

Micro-architecture performance and efficiency optimization lead to an improvement of the performance per watt signature compared with the previous Power9 core implementation. The overall energy efficiency is better by a factor of approximately 2.6, which demonstrates the advancement in processor design that is manifested by Power10.

The Power10 processor core includes the following key features and improvements that affect performance:

► Enhanced load and store bandwidth
► Deeper and wider instruction windows
► Enhanced data prefetch
► Branch execution and prediction enhancements
► Instruction fusion

Enhancements in the area of computation resources, working set size, and data access latency are described next. The change in relation to the Power9 processor core implementation is provided in parentheses.

Enhanced computation resources

Here are the major computational resource enhancements:

► Eight VSU execution slices, each supporting 64-bit scalar or 128-bit single instructions multiple data (SIMD) +100% for permute, fixed-point, floating-point, and crypto (Advanced Encryption Standard (AES) and Secure Hash Algorithm (SHA)) +400% operations.

► Four units for MMA, each capable of producing a 512-bit result per cycle (new), and +400% Single and Double precision FLOPS plus support for reduced precision AI acceleration.

► Two units for QP and DF operations for more instruction types.

Larger working sets

The following major changes were implemented in working set sizes:

► L1 instruction cache: Two 48 KB 6-way (96 KB total) (+50%)
► L2 cache: 2 MB 8-way (+400%)
► L2 translation lookaside buffer (TLB): Two 4-K entries (8 K total) (+400%)

Data access with reduced latencies

The following major changes reduce latency for load data:

► L1 data cache access at four cycles nominal with zero penalty for store-forwarding (- 2 cycles) for store forwarding

► L2 data access at 13.5 cycles nominal (-2 cycles)

► L3 data access at 27.5 cycles nominal (-8 cycles)

► TLB access at 8.5 cycles nominal for effective-to-real address translation (ERAT) miss including for nested translation (-7 cycles)

Micro-architectural innovations that complement physical and logic design techniques and specifically address energy efficiency include the following examples:

► Improved clock-gating.

► Reduced flush rates with improved branch prediction accuracy.

► Fusion and gather operating merging.

► Reduced number of ports and reduced access to selected structures.

► Effective address (EA)-tagged L1 data and instruction cache yield ERAT access only on a cache miss.

In addition to improvements in performance and energy efficiency, security represents a major architectural focus area. The Power10 processor core supports the following security features:

► Enhanced hardware support that provides improved performance while mitigating for speculation-based attacks.

► Dynamic Execution Control Register (DEXCR) support.

► Return-oriented programming (ROP) protection.

2.1.4 Simultaneous multithreading

Each core of the Power10 processor supports multiple hardware threads that represent independent execution contexts. If only one hardware thread is used, the processor core runs in ST mode.

If more than one hardware thread is active, the processor runs in SMT mode. In addition to the ST mode, the Power10 processor supports the following different SMT modes:

► SMT2: Two hardware threads active
► SMT4: Four hardware threads active
► SMT8: Eight hardware threads active

SMT enables a single physical processor core to simultaneously dispatch instructions from more than one hardware thread context. Computational workloads can use the processor core's execution units with a higher degree of parallelism. This ability enhances the throughput and scalability of multi-threaded applications and optimizes the compute density for ST workloads.

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 most workloads, especially those workloads with large or frequently changing working sets, such as database servers and web servers.

Table 2-2 lists a historic account of the SMT capabilities that are supported by each implementation of the IBM Power Architecture® since IBM Power4.

Table 2-2   SMT levels that are supported by Power processors

Technology	Maximum cores per system	Supported hardware threading modes	Maximum hardware threads per partition
IBM Power4	32	ST	32
IBM Power5	64	ST and SMT2	128
IBM Power6	64	ST and SMT2	128
IBM Power7	256	ST, SMT2, and SMT4	1024
IBM Power8	192	ST, SMT2, SMT4, and SMT8	1536
IBM Power9	192	ST, SMT2, SMT4, and SMT8	1536
IBM Power10	240	ST, SMT2, SMT4, and SMT8	1920[a]

a. Power Hypervisor supports a maximum 240 SMT8 threads, that is, 1920. AIX support up to 1920 (240 SMT8) total threads in a single partition, starting with AIX 7.3 + Power10.

The Power E1050 server supports the ST, SMT2, SMT4, and SMT8 hardware threading modes. With the maximum number of 96 cores, a maximum of 768 hardware threads per partition can be reached.

2.1.5 Matrix Math Acceleration AI workload acceleration

The MMA facility was introduced by the Power ISA 3.1. The related instructions implement numerical linear algebra operations on small matrices and are meant to accelerate computation-intensive kernels, such as matrix multiplication, convolution, and discrete Fourier transform.

To efficiently accelerate MMA operations, the Power10 processor core implements a *dense math engine* (DME) microarchitecture that effectively provides an accelerator for cognitive computing, machine learning, and AI inferencing workloads.

The DME encapsulates compute efficient pipelines, a physical register file, and an associated data flow that keeps the resulting accumulator data local to the compute units. Each MMA pipeline performs outer-product matrix operations, reading from and writing back to a 512-bit accumulator register.

Power10 implements the MMA accumulator architecture without adding an architected state. Each architected 512-bit accumulator register is backed by four 128-bit Vector Scalar eXtension (VSX) registers.

Code that uses the MMA instructions is included in OpenBLAS and Eigen libraries. This library can be built by using the most recent versions of the GNU Compiler Collection (GCC) compiler. The latest version of OpenBLAS is available at this web page.

OpenBLAS is used by the Python-NumPy library, PyTorch, and other frameworks, which makes it easy to use the performance benefit of the Power10 MMA accelerator for AI workloads.

The Power10 MMA accelerator technology is also used by the IBM Engineering and Scientific Subroutine Library for AIX on Power 7.1 (program number 5765-EAP).

Program code that is written in C/C++ or Fortran can benefit from the potential performance gains by using the MMA facility if the code is compiled by the following IBM compiler products:

► IBM Open XL C/C++ for AIX 17.1 (program numbers 5765-J18, 5765-J16, and 5725-C72)

► IBM Open XL Fortran for AIX 17.1 (program numbers 5765-J19, 5765-J17, and 5725-C74)

For more information about the implementation of the Power10 processor's high throughput math engine, see *A matrix math facility for Power ISA processors*.

For more information about fundamental MMA architecture principles with detailed instruction set usage, register file management concepts, and various supporting facilities, see *Matrix-Multiply Assist Best Practices Guide*, REDP-5612.

2.1.6 Power10 compatibility modes

The Power10 core implements the Processor Compatibility Register (PCR) as described in the Power ISA 3.1, primarily to facilitate Live Partition Mobility (LPM) to and from previous generations of IBM Power hardware.

Depending on the specific settings of the PCR, the Power10 core runs in a compatibility mode that pertains to Power9 (Power ISA 3.0) or Power8 (Power ISA 2.07) processors. The support for processor compatibility modes also enables older operating system (OS) versions of AIX, IBM i, Linux, or Virtual I/O Server (VIOS) environments to run on Power10 processor-based systems.

Note: The Power E 1050 server does not support IBM i.

The Power10 processor-based Power E1050 server supports the Power8, Power9 Base, Power9, and Power10 compatibility modes.

2.1.7 Processor module options

The Power E1050 server uses DCMs. The server can be populated with either two, three, or four DCMs. A one DCM server is not supported because it would provide only a single PCIe slot. Figure 2-1 on page 35 shows how PCIe slots are connected to processors.

Note: All processor modules that are used in a Power E1050 server must be identical (the same Feature Code).

Table 2-3 shows the processor features that are available for the Power E1050 server.

Table 2-3   Power E1050 processor features

Feature Code	Description
#EHC8	Solution Edition for Healthcare typical 2.95 - 3.9 GHz 24-core Processor Module (North America only)
#EPEU	12-core typical 3.35 - 4.0 GHz (maximum) processor

Feature Code	Description
#EPEV	18-core 3.2 - 4.0 GHZ (maximum) processor
#EPGW	24-core 2.95 - 3.9 GHZ (maximum) processor

2.1.8 Processor activations

A physical core in the Power E1050 server must be activated to use the core. There are multiple processor activation types that are available.

Static processor activations

A classical static processor activation allows the permanent usage of the activated cores. For the number of activated cores, IBM generates a key that is integrated into the system for a new, purchased system.

The minimum number of cores that must be activated is one socket. For example, in a server with all four sockets populated with the 12-core option, the minimum number of cores to activate is 12.

There are two kinds of activation features: general-purpose and Linux. Cores with a general-purpose activation can run any supported OS, but cores with a Linux activation can run only Linux OSs. The processor-specific activation features for the Power E1050 server are shown in Table 2-4.

Table 2-4   Processor activation Feature Codes

Processor feature	Static activation feature	Static activation for Linux feature
EHC8 24-core module for Healthcare	#EHCA[a]	N/A
EPEU 12-core module	#EPUR	#ELBW
EPEV 18-core module	#EPUS	#ELBX
EPGW 24-core module	#EPYT	#ELBY

a. #EHCA activates 24 cores for each feature. You must order #EHCA with each #EHC8 module.

Capacity on Demand

Two types of Capacity on Demand (CoD) capability are available for processor and memory on the Power E1050 server:

► Capacity Upgrade on Demand (CUoD) processor activations

If not all cores were activated, it is possible to purchase more core activations through a Miscellaneous Equipment Specification (MES) upgrade order, which results in another key that can be integrated into the system by using the Hardware Management Console (HMC) or the Advanced System Management Interface (ASMI) without requiring a restart of the server or interrupting the business. After entering the code, the additional cores can be used and assigned to LPARs.

► Elastic CoD (Temporary)

With Elastic CoD, you can temporarily activate processors and memory as full-day increments as needed. The processors and memory can be activated and turned off an unlimited number of times whenever you need extra processing resources.

The Elastic CoD capacity can be ordered by using the IBM Entitled Systems Support (IBM ESS) website or through your IBM Business Partner or IBM sales representative. If the resources are ordered through an IBM Business Partner or IBM sales representative, the following Feature Codes apply for the Power E1050 server:

- #ECL1: 1-unit AIX/Linux Processor days Midrange systems
- #ECL2: 100-unit AIX/Linux Processor days Midrange systems
- #ECL3: 10,000-unit AIX/Linux Processor days Midrange systems

After ordering, the capacity is visible at the IBM ESS website. To use the capacity in IBM ESS, select the system where the capacity should be used and the number of days to generate a code that can be entered into the system by using the HMC. For more information, see the IBM Entitled Systems Support (IBM ESS) website or ask your IBM Business Partner or IBM sales representative.

> **Hint:** On the IBM ESS website, you can activate a demonstration mode. In the demonstration mode, you can simulate how to order capacity and how to produce keys without any real execution.

IBM Power Private Cloud with Shared Utility Capacity

Besides the two CoD offerings, the Power E1050 server supports the Power Private Cloud with Shared Utility Capacity solution (Power Enterprise Pools 2.0 (PEP2)), which is an infrastructure offering model that enables cloud agility and cost optimization with pay-for-use pricing. This usage model requires the configuration of the PEP2 Enablement feature (#EP20) for the server, and a minimum of one Base Processor Activation for Pools 2.0 feature is needed. The base processor activations are static and permanent. Additional processor resources needed beyond the capacity that is provided by the base processor activations are metered by the minute and paid through capacity credits.

For more information about PEP2, see *IBM Power Systems Private Cloud with Shared Utility Capacity: Featuring Power Enterprise Pools 2.0*, SG24-8478.

> **Note:** The CUoD technology usage model and the Shared Utility Capacity (PEP2) offering model are all mutually exclusive in respect to each other.

2.1.9 On-chip L3 cache and intelligent caching

The Power10 processor includes a large on-chip L3 cache of up to 120 MB with a NUCA architecture that provides mechanisms to distribute and share cache footprints across a set of L3 cache regions. Each processor core can access an associated local 8 MB of L3 cache. It also can access the data in the other L3 cache regions on the chip and throughout the system.

Each L3 region serves as a victim cache for its associated L2 cache, and it can provide aggregate storage for the on-chip cache footprint.

Intelligent L3 cache management enables the Power10 processor to optimize the access to L3 cache lines and minimize cache latencies. The L3 cache includes a replacement algorithm with data type and reuse awareness. It also supports an array of prefetch requests from the core, including instruction and data, and works cooperatively with the core, memory controller, and SMP interconnection fabric to manage prefetch traffic, which optimizes system throughput and data latency.

The L3 cache supports the following key features:

► Enhanced bandwidth that supports up to 64 bytes per core processor cycle to each SMT8 core.

► Enhanced data prefetch that is enabled by 96 L3 prefetch request machines that service prefetch requests to memory for each SMT8 core.

► Plus-one prefetching at the memory controller for enhanced effective prefetch depth and rate.

► Power10 software prefetch modes that support fetching blocks of data into the L3 cache.

► Data access with reduced latencies.

2.1.10 Nest accelerator

The Power10 processor has an on-chip accelerator that is called the *nest accelerator (NX) unit*. The co-processor features that are available on the Power10 processor are like the features on the Power9 processor. These co-processors provide specialized functions, such as the following examples:

► IBM proprietary data compression and decompression
► Industry-standard Gzip compression and decompression
► AES and SHA cryptography
► Random number generation

Figure 2-5 shows a block diagram of the NX unit.
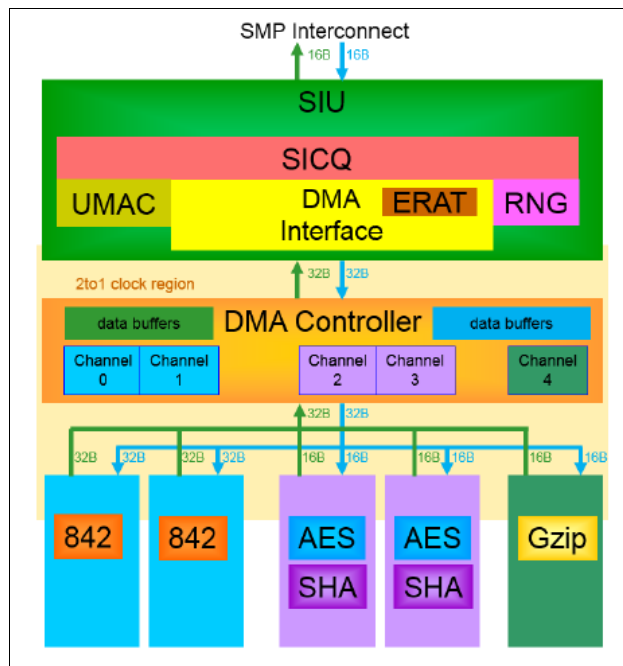
Figure 2-5   Block diagram of the NX unit

Each one of the AES and SHA engines, data compression, and Gzip units consist of a co-processor type, and the NX unit features three co-processor types. The NX unit also includes more support hardware to support co-processor invocation by user code, usage of effective addresses, high-bandwidth storage accesses, and interrupt notification of job completion.

The direct memory access (DMA) controller of the NX unit helps to start the co-processors and move data on behalf of co-processors. An SMP interconnect unit (SIU) provides the interface between the Power10 SMP interconnect and the DMA controller.

The NX co-processors can be started transparently through library or OS kernel calls to speed up operations that are related to data compression, LPM migration, IPsec, JFS2 encrypted file systems, PKCS11 encryption, random number generation, and the recently announced logical volume encryption.

In effect, this on-chip NX unit on Power10 systems implements a high-throughput engine that can perform the equivalent work of multiple cores. The system performance can benefit by offloading these expensive operations to on-chip accelerators, which can greatly reduce the CPU usage and improve the performance of applications.

The accelerators are shared among the logical partitions (LPARs) under the control of the PowerVM hypervisor and accessed through a hypervisor call. The OS, along with the PowerVM hypervisor, provides a send address space that is unique per process requesting the co-processor access. This configuration allows the user process to directly post entries to the first in - first out (FIFO) queues that are associated with the NX accelerators. Each NX co-processor type has a unique receive address space corresponding to a unique FIFO for each of the accelerators.

For more information about the usage of the xgzip tool that uses the Gzip accelerator engine, see the following resources:

► *Using the Power9 NX (gzip) accelerator in AIX*

► *Power9 GZIP Data Acceleration with IBM AIX*

► *Performance improvement in OpenSSH with on-chip data compression accelerator in Power9*

► *The nxstat command*

2.1.11 SMP interconnect and accelerator interface

The Power10 processor provides a highly optimized, 32-Gbps differential signaling technology interface that is structured in 16 entities (eight ports that provide two 1x9 xBus). Each entity consists of eight data lanes and one spare lane. This interface can facilitate the following functional purposes:

► First- or second-tier SMP link interface, enabling up to 16 Power10 processors to be combined into a large, robustly scalable single-system image

► Open Coherent Accelerator Processor Interface (OpenCAPI) to attach cache coherent and I/O-coherent computational accelerators, load/store addressable host memory devices, low latency network controllers, and intelligent storage controllers

► Host-to-host integrated memory clustering interconnect, enabling multiple Power10 systems to directly use memory throughout the cluster

Note: The OpenCAPI interface and the memory clustering interconnect are Power10 technology option for future usage.

Because of the versatile nature of signaling technology, the 32-Gbps interface is also referred to as a PowerAXON interface. The IBM proprietary X-bus links connect two processors on a system board with a common reference clock. The IBM proprietary A-bus links connect two processors in different drawers on different reference clocks by using a cable.

OpenCAPI is an open interface architecture that allows any microprocessor to attach to the following items:

► Coherent user-level accelerators and I/O devices
► Advanced memories accessible through read/write or user-level DMA semantics

The OpenCAPI technology is developed, enabled, and standardized by the OpenCAPI Consortium. For more information about the consortium's mission and the OpenCAPI protocol specification, see OpenCAPI Consortium.

The PowerAXON interface is implemented on dedicated areas that are at each corner of the Power10 processor die.

The Power10 processor-based E1050 server uses this interface to implement:

► DCM internal chip-to-chip connections

► Chip-to-chip SMP interconnects between DCMs in a 1-hop topology

► OpenCAPI accelerator interface connections

The chip-to-chip DCM internal interconnects, and connections to the OpenCAPI ports, are shown in Figure 2-6.
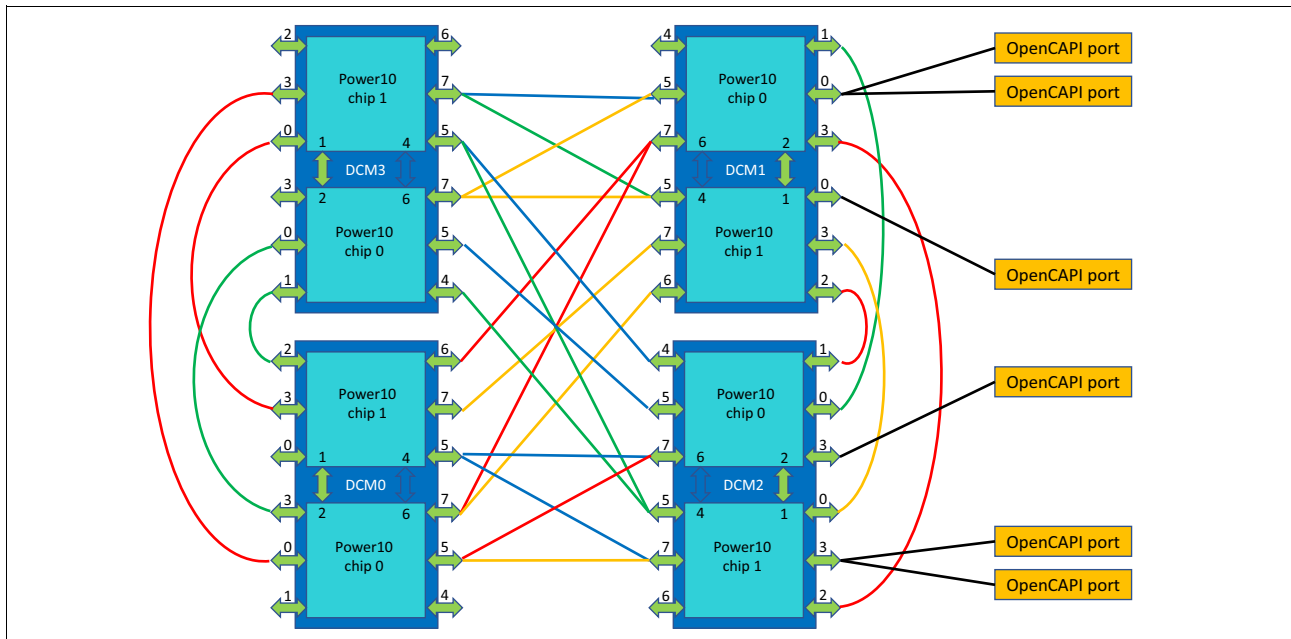
Figure 2-6   SMP xBus 1-hop interconnect and OpenCAPI port connections

Note: The left (front) DCM0 and DCM3 are placed in a 180-degrees rotation compared to the two right (rear) DCM1 and DCM2 to optimize PCIe slots and Non-volatile Memory Express (NVMe) bay wirings.

All ports provide two 1x9 xbusses, so some ports show two connections. All connections going outside the DCM are 1x9, which is also the case for the ports where only one connection is shown.

For the internal connection of the two chips in one DCM, two ports are available, but only one is used. The used port connects the two chips inside the DCM with a 2x9 bus.

> **Note:** The implemented OpenCAPI interfaces can be used in the future, but they are currently not used by the available technology products.

2.1.12 Power and performance management

Power10 processor-based servers implement an enhanced version of the power management EnergyScale technology. As in the previous Power9 EnergyScale implementation, the Power10 EnergyScale technology supports dynamic processor frequency changes that depend on several factors, such as workload characteristics, the number of active cores, and environmental conditions.

Based on the extensive experience that was gained over the past few years, the Power10 EnergyScale technology evolved to use the following effective and simplified set of operational modes:

► Power-saving mode
► Static mode (nominal frequency)
► Maximum performance mode (MPM)

The Power9 dynamic performance mode (DPM) has many features in common with the Power9 MPM. Because of this redundant nature of characteristics, the DPM for Power10 processor-based systems was removed in favor of an enhanced MPM. For example, the maximum frequency is now achievable in the Power10 enhanced MPM (regardless of the number of active cores), which was not always the case with Power9 processor-based servers.

In the Power E1050 server, MPM is enabled by default. This mode dynamically adjusts the processor frequency to maximize performance and enable a much higher processor frequency range. Each of the power saver modes delivers consistent system performance without any variation if the nominal operating environment limits are met.

For Power10 processor-based systems that are under control of the PowerVM hypervisor, the MPM is a system-wide configuration setting, but each processor module frequency is optimized separately.

The following factors determine the maximum frequency that a processor module can run at:

► Processor utilization: Lighter workloads run at higher frequencies.

► Number of active cores: Fewer active cores run at higher frequencies.

► Environmental conditions: At lower ambient temperatures, cores are enabled to run at higher frequencies.

The following Power10 EnergyScale modes are available:

► Power-saving mode

The frequency is set to the minimum frequency to reduce energy consumption. Enabling this feature reduces power consumption by lowering the processor clock frequency and voltage to fixed values. This configuration reduces power consumption of the system while delivering predictable performance.

► Static mode

The frequency is set to a fixed point that can be maintained with all normal workloads and in all normal environmental conditions. This frequency is also referred to as *nominal frequency*.

► MPM

Workloads run at the highest frequency possible, depending on workload, active core count, and environmental conditions. The frequency does not go below the static frequency for all normal workloads and in all normal environmental conditions.

In MPM, the workload is run at the highest frequency possible. The higher power draw enables the processor modules to run in an MPM typical frequency range (MTFR), where the lower limit is well above the nominal frequency and the upper limit is given by the system's maximum frequency.

The MTFR is published as part of the system specifications of a specific Power10 system if it is running by default in MPM. The higher power draw potentially increases the fan speed of the respective system node to meet the higher cooling requirements, which in turn causes a higher noise emission level of up to 15 decibels.

The processor frequency typically stays within the limits that are set by the MTFR, but can be lowered to frequencies between the MTFR lower limit and the nominal frequency at high ambient temperatures above 27 °C (80.6 °F). If the data center ambient environment is less than 27 °C, the frequency in MPM consistently is in the upper range of the MTFR (roughly 10% - 20% better than nominal). At lower ambient temperatures (below 27 °C, or 80.6 °F), MPM mode also provides deterministic performance. As the ambient temperature increases above 27 °C, determinism can no longer be ensured. This mode is the default mode for all Power10 processor-based scale-out servers.

► Idle power saver mode (IPS)

IPS mode lowers the frequency to the minimum if the entire system (all cores of all sockets) is idle. It can be enabled or disabled separately from all other modes.

Figure 2-7 shows the comparative frequency ranges for the Power10 power-saving mode, static or nominal mode, and the MPM. The frequency adjustments for different workload characteristics, ambient conditions, and idle states are also indicated.
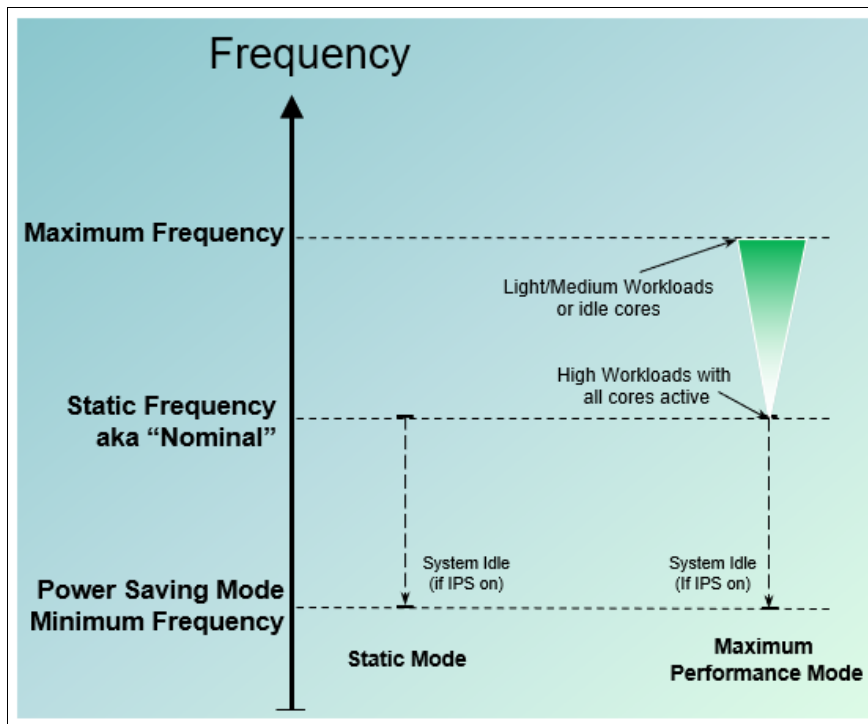
Figure 2-7   Power10 power management modes and related frequency ranges

Table 2-5 shows the power-saving mode and the static mode frequencies and the frequency ranges of the MPM for all processor module types that are available for the Power E1050 server.

Note: For all Power10 processor-based scale-out systems, the MPM is enabled by default.

Table 2-5   Characteristic frequencies and frequency ranges for Power E1050 servers

Feature Code	Cores per single-chip module	Power-saving mode frequency (GHz)	Static mode frequency (GHz)	Maximum performance mode frequency range (GHz)
#EPEU	12	2.0	3.35	3.35 - 4.00 (max)
#EPEV	18	2.0	3.20	3.20 - 4.00 (max)
#EPGW	24	2.0	2.95	2.95 - 3.90 (max)
#EHC8[a]	24	2.0	2.95	2.95 - 3.90 (max)

a. #EHC8 is technically identical to #EPGW, but is part of the Healthcare solution edition (North America (NA) only).

The controls for all power saver modes are available on the ASMI, and can be dynamically modified. A system administrator can also use the HMC to set power saver mode or to enable static mode or MPM.

Figure 2-8 shows the ASMI menu for Power and Performance Mode Setup on a Power E1050 server.
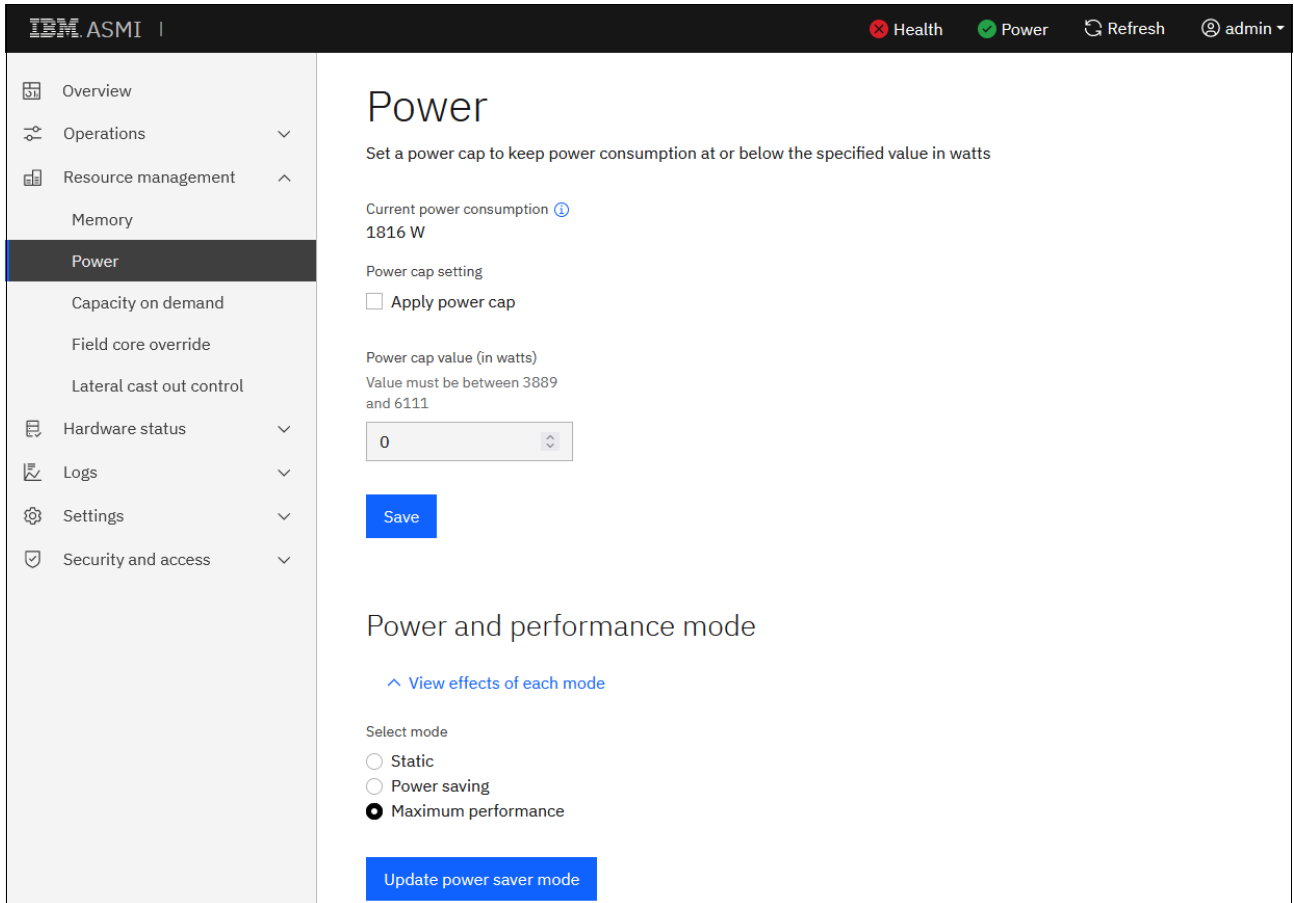
Figure 2-8   ASMI menu for Power and Performance Mode Setup

Figure 2-9 shows the HMC menu for Power and Performance Mode Setup.
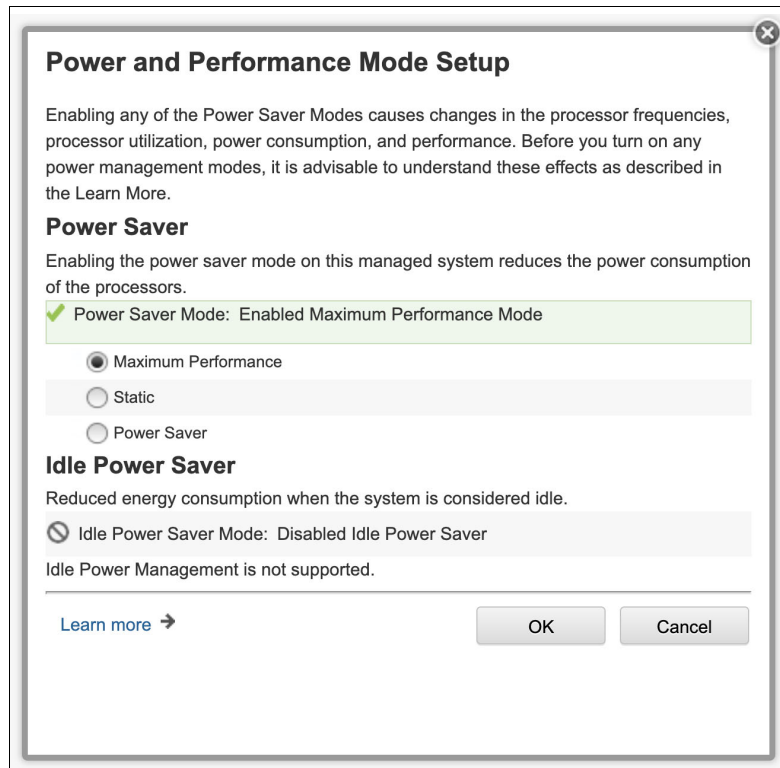
Figure 2-9   HMC menu for Power and Performance Mode Setup

2.1.13 Comparing Power10, Power9, and Power8 processors

The Power10 processor-based systems are using three different processor module packages:

► SCMs, which are based on one Power10 chip

► DCMs, which combine two Power10 chips where both Power10 chips contribute active processor cores

► Entry single-chip modules (eSCMs), which combine two Power10 chips, but all active processor core resources are bundled on one of the two chips

Power E1050 servers use exclusively SCM modules with up to 15 active SMT8-capable cores. These SCM processor modules are structural and performance-optimized for usage in *scale-up multi-socket* systems.

DCM modules with up to 30 active SMT8 capable cores are used in 4-socket Power E1050 servers, and 2-socket Power S1022 and Power S1024 servers. eSCMs with up to eight active SMT8-capable cores are used in 1-socket Power S1014 and 2-socket Power S1022s servers. DCM and eSCM modules are designed to support *scale-out 1- to 4-socket* Power10 processor-based servers.

Table 2-6 compares key features and characteristics of the Power10, Power9, and Power8 processor implementations as used in enterprise class scale-up and the Power10 processor-based scale-out servers.

Table 2-6   Comparison of the Power10 processor technology to prior processor generations

Characteristics	Power10			Power9	Power8
	DCM	eSCM	SCM		
Technology	7 nm			14 nm	22 nm
Die size	2 x 602 mm^2	2 x 602 mm^2	602 mm^2	693 mm^2	649 mm^2
Processor module size	74.5 mm x 85.75 mm	74.5 mm x 85.75 mm	68.5 mm x 77.5 mm	68.5 mm x 68.5 mm	71.5 mm x 71.5 mm
Number of transistors	2 x 18 billion	2 x 18 billion	18 billion	8 billion	4.2 billion
Maximum cores	24	8	15	12	12
Maximum hardware threads per core	8			8	8
Maximum static frequency or high-performance frequency range[a]	3.4 - 4.0 GHz	3.0 - 3.9 GHz	3.6 - 4.15 GHz	3.9 - 4.0 GHz	4.15 GHz
L2 Cache per core	2048 KB			512 KB	512 KB
L3 Cache	8 MB of L3 cache per core, with each core having access to the full 120 MB of L3 cache, on-chip high-efficiency SRAM[b]			10 MB of L3 cache per core, with each core having access to the full 120 MB of L3 cache, on-chip eDRAM[c]	8 MB of L3 cache per core, with each core having access to the full 96 MB of L3 cache, on-chip eDRAM
Supported memory technology	Double Data Rate 4 (DDR4)[d]			DDR4 and DDR3[e]	DDR3 and DDR4
I/O bus	PCIe Gen 5			PCIe Gen 4	PCIe Gen 3

a. Best of class typical frequency range, where the lower limit of the range coincides with the maximum static or nominal frequency.
b. Static random-access memory.
c. Embedded dynamic random-access memory.
d. Power10 processor memory logic and the memory subsystem of Power10 processor-based servers can use DDR5 technology DIMMs.
e. Only DDR3 memory CDIMMs, which are transferred in the context of a model upgrade from Power E870, Power E870C, Power E880, or Power E880C servers to a Power E980 server, are supported.

2.2 Memory subsystem

The Power E1050 server uses the new and innovative OMI. OMI is a technology-neutral memory interface. The Power E1050 server supports DDIMMs that are based on DDR4 technology, which are plugged into the OMI slots. A maximum of 64 OMI slots are available in a server with all four processor sockets populated, with which a maximum of 16 TB of memory can be installed in to the server. In the following sections, the details about OMI and DDIMMs are described in more detail.

2.2.1 Open Memory Interface

The OMI is driven by eight on-chip MCUs, and it is implemented in two separate physical building blocks that lie in opposite areas at the outer edge of the Power10 die. Each area supports 64 OMI lanes that are grouped in four OMI ports. One port in turn consists of two OMI links with eight lanes each, which operate in a latency-optimized manner with unprecedented bandwidth and scale at 32-Gbps speed.

One Power10 processor chip supports the following functional elements to access main memory:

► Eight MCUs

► Eight OMI ports that are controlled one-to-one through a dedicated MCU

► Two OMI links per OMI port, for a total of 16 OMI links

► Eight lanes per OMI link for a total of 128 lanes, all running at 32-Gbps speed

The Power10 processor provides natively an aggregated maximum theoretical full-duplex memory interface bandwidth of 1 TBps per chip.

Memory interface architecture for dual-chip modules

The DCM that is used in the Power E1050 server combines two Power10 processor chips in one processor package. A total of 2 x 8 = 16 OMI ports and 2 x 16 = 32 OMI links are physically present on a Power10 DCM. But because the chips on the DCM are tightly integrated and the aggregated memory bandwidth of 8 OMI ports already culminates at a maximum theoretical full-duplex bandwidth of 1 TBps, only half of the OMI ports are active. Each chip of the DCM contributes four OMI ports and eight OMI links to facilitate main memory access. Figure 2-1 on page 35 and Figure 2-11 on page 61 show details about the OMI port designation and the physical location of the active OMI units of a DCM.

In summary, one DCM supports the following functional elements to access main memory:

► Four active MCUs per chip, for a total of eight MCUs per module.

► Each MCU maps one-to-one to an OMI port.

► Four OMI ports per chip, for at total of eight OMI ports per module.

► Two OMI links per OMI port for a total of eight OMI links per chip and 16 OMI links per module.

► Eight lanes per OMI link for a total of 128 lanes per module, all running at 32-Gbps speed.

The Power10 DCM provides an aggregated maximum theoretical full-duplex memory interface bandwidth of 512 GBps per chip and 1 TBps per module.

2.2.2 Differential Dual Inline Memory Module

The Power10 processor-based E1050 server introduces a new 4U tall DDIMM, which has a new OpenCAPI memory interface that is known as OMI for resilient and fast communication to the processor. This new memory subsystem design delivers solid resiliency features, as described below:

► Memory buffer: The DDIMM contains a memory buffer with key resiliency features, including protection of critical data and address flows by using cyclic redundancy check (CRC), error correction code (ECC), and parity; a maintenance engine for background memory scrubbing and memory diagnostics; and a Fault Isolation Register (FIR) structure, which enables firmware attention-based fault isolation and diagnostics.

► OMI: The OMI interface between the memory buffer and processor memory controller is protected by dynamic lane calibration and a CRC retry/recovery facility to retransmit lost frames to survive intermittent bit flips. A complete lane fail can also be survived by triggering a dynamic lane reduction from eight to four independently for both up and downstream directions. A key advantage of the OMI interface is that it simplifies the number of critical signals that must cross connectors from processor to memory compared to a typical industry-standard DIMM design.

► Memory ECC: The DDIMM includes a robust 64-byte Memory ECC with 8-bit symbols, which can correct up to five symbol errors (one x4 chip and one more symbol) and retry for data and address uncorrectable errors.

► Dynamic row repair: To further extend the life of the DDIMM, the dynamic row repair feature can restore full use of a dynamic RAM (DRAM) for a fault that is contained to a DRAM row while the system continues to operate.

► Spare temperature sensors: Each DDIMM provides spare temperature sensors such that the failure of one does not require a DDIMM replacement.

► Spare DRAMs: 4U DDIMMs include two spare x4 memory modules (DRAMs) per rank, which can be substituted for failed DRAMs during runtime operation. Combined with ECC correction, the two spares allow the 4U DDIMM to continue to function with three bad DRAMs per rank compared to 1 (single device data correct) or 2 (double device data correct) bad DRAMs in a typical industry-standard DIMM design. This setup extends self-healing capabilities beyond what is provided with dynamic row repair capability.

► Spare Power Management Integrated Circuits (PMICs): 4U DDIMMs include PMICs such that the failure of one PMIC does not require a DDIMM replacement.

Note: DDIMMs are also available in a 2U form factor. These 2U DDIMMs are not supported in the Power E1050 server.

Figure 2-10 shows a 4U DDIMM with a plug that connects into an OMI slot.
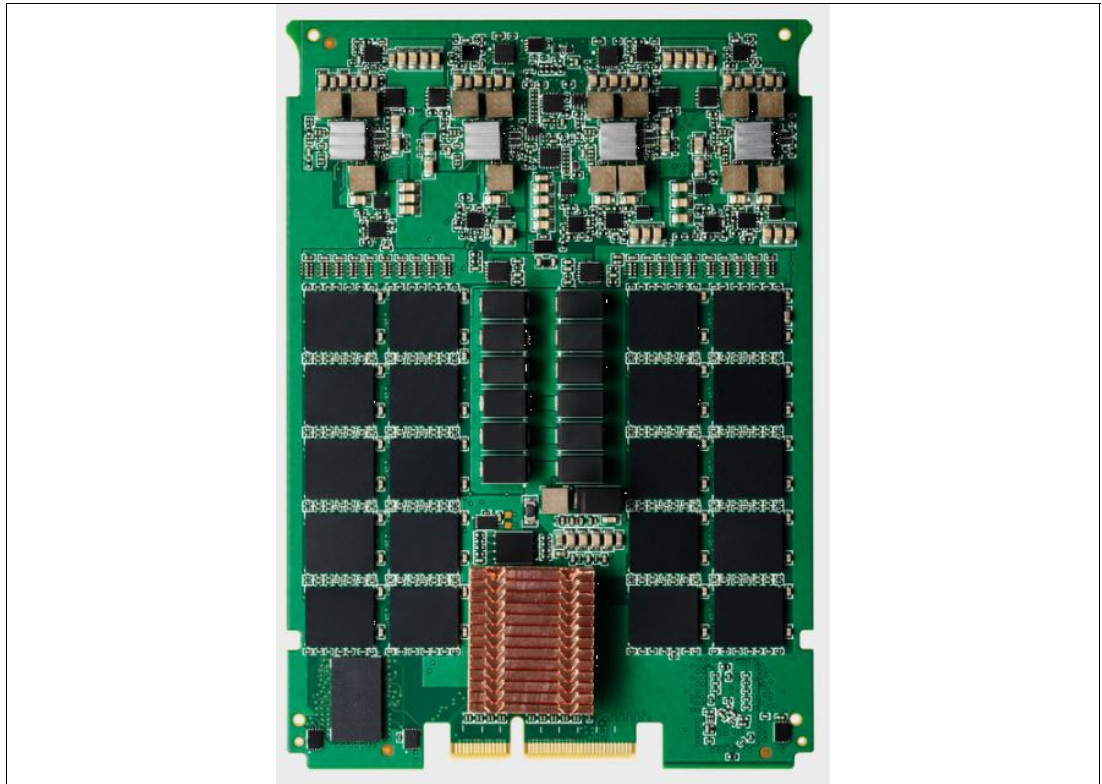
Figure 2-10   4U DDIMM feature

Maximum memory and maximum theoretical memory bandwidth

The OMI physical interface enables low latency, high-bandwidth, and technology-neutral host memory semantics to the processor and allows attaching established and emerging memory elements. With the Power10 processor-based E1050 server, OMI initially supports one main tier, low-latency, and enterprise-grade DDR4 DDIMM per OMI link. This architecture yields a total memory module capacity of 16 DDIMMs per populated processor module (64 with all four modules populated).

The memory bandwidth and the total memory capacity depend on the DDIMM density and the associated DDIMM frequency that are configured for the Power E1050 server. Table 2-7 list the maximum memory and memory bandwidth per populated socket and the maximum values for a fully populated server.

Table 2-7   Maximum theoretical memory and memory bandwidth for the Power E1050 server

Feature Code	DIMM size[a]	DRAM speed	Max memory per socket	Max memory per server	Max memory bandwidth per socket	Max memory bandwidth per server
#EM75	64 GB (2 x32 GB)	3200 MHz	512 GB	2 TB	409 GBps	1.636 GBps
#EM76	128 GB (2 x64 GB)	3200 MHz	1 TB	4 TB	409 GBps	1.636 GBps
#EM77	256 GB (2 x128 GB)	2933 MHz	2 TB	8 TB	375 GBps	1.500 GBps
#EM7J	512 GB (2 x256 GB)	2933 MHz	4 TB	16 TB	375 GBps	1.500 GBps

a. The 128 GB and 256 GB DDIMMs are planned to be available from 9 December 2022.

2.2.3 Memory activations

To use the physical memory, it must be activated. For that action, the Power E1050 server supports several options.

Static memory activations

A classical memory activation allows the permanent usage of the activated memory. For the amount of activated memory, IBM generates a key that is integrated into the system for a new purchased system.

The minimum mount of memory that must be activated is 50% of the installed memory or at least 256 GB. Static memory can be activated in the system configuration by using the Feature Code #EMCP for the activation of 1 GB or #EMCQ for the activation of 100 GB.

Capacity on Demand

Two types of CoD capability are available for processor and memory on the Power E1050 server:

▶ CUoD

If you have not activated memory in your server, you may use CUoD to purchase extra permanent memory (and processor) capacity and dynamically activate it when you need it. This goal can be achieved through a MES upgrade order, which results in another key that can be integrated into the system by using the HMC or the ASMI without restarting the server or interrupting the business. After entering the code, the additional memory can be used and assigned to LPARs.

For activating memory in a configuration, use the Feature Code #EMCP to activate 1 GB or Feature Code #EMCQ to activate 100 GB of memory for any OS.

> **Note:** A minimum of 256 GB or 50% of the installed memory must be activated by default.

▶ Elastic CoD (Temporary)

With Elastic CoD, you can temporarily activate processors or memory as full-day increments as needed.

The Elastic CoD capacity can be ordered by using IBM ESS or through your IBM Business Partner or IBM sales representative. If the resources are ordered through an IBM Business Partner or IBM sales representative, the following Feature Codes apply for the IBM E1050 server:

– #ECM1: 1-unit Memory days for Midrange/High End systems

– #ECM2: 100-unit Memory days for Midrange/High End systems

– #ECM3: 10,000-unit Memory days for Midrange/High End systems

After ordering, the capacity is visible at the IBM ESS website. To use the capacity in IBM ESS, select the system where the capacity should be used and the number of days to generate a code that can be entered into the system by using the HMC. For more information, see have a look into IBM ESS or contact your IBM Business Partner or IBM sales representative.

> **Hint:** On the IBM ESS website, you can activate a demonstration mode. In the demonstration mode, you can simulate how to order capacity and produce keys without any real execution.

IBM Power Private Cloud with Shared Utility Capacity

Besides the two CoD offerings, the Power E1050 server also supports the Power Private Cloud with Shared Utility Capacity solution (PEP2), which is an infrastructure offering model that enables cloud agility and cost optimization with pay-for-use pricing. This use model requires the configuration of the PEP2 Enablement feature (#EP20) for the server, and a minimum of a 256 GB Base Memory Activation for Pools 2.0 feature is needed. The base memory activations are static and permanent. More memory resources that are needed beyond the capacity that is provided by the base memory activations are metered by the minute and paid through capacity credits.

> **Note:** For processors, the metering measures used capacity cycles independent of the entitled capacity that is assigned to LPARs. For memory, this process is different. If memory is assigned to an LPAR, it is used from an Enterprise Pools perspective, even when the OS does not use it.

For more information about PEP2, see *IBM Power Systems Private Cloud with Shared Utility Capacity: Featuring Power Enterprise Pools 2.0*, SG24-8478.

> **Note:** The CUoD technology usage model and the Shared Utility Capacity (PEP2) offering model are all mutually exclusive in respect to each other.

2.2.4 Memory placement rules

Each Power10 chip requires a minimum of two DDIMMs that are installed. Because each processor socket has a dual-chip module, four (2x2 DDIMMs) must be installed per socket. Because each Power E1050 server requires a minimum of two sockets populated, a minimum of eight DDIMMs must be installed. Using the smallest 32 GB DDIMMs, there is a minimum of 256 GB per server in a 2-socket configuration.

The OMI sockets that are numbered as P0-C22 - P0-C95 must be populated in a defined order. Figure 2-11 shows the plugging rules with colors. First, populate the green slots, then the pink slots, then the blue slots, and then the yellow slots.
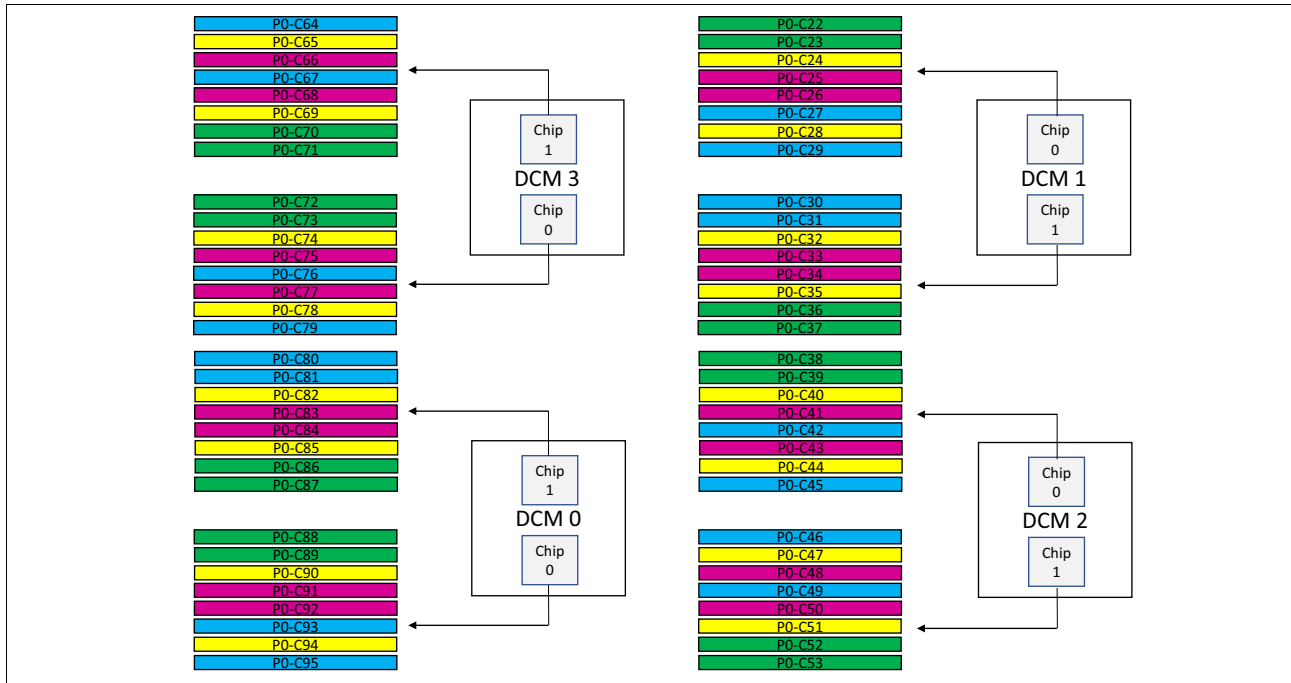
Figure 2-11   OMI ports to DDIMM connector cross-reference

Table 2-8 shows the order in which the DDIMM slots should be populated.

Table 2-8   DDIMM plugging rules

Set	Location codes			
	DCM0	**DCM1**	**DCM2**	**DCM3**
First set of DDIMMs (green in Figure 2-11)	P0-C86	P0-C22	P0-C38	P0-C70
	P0-C87	P0-C23	P0-C39	P0-C71
	P0-C88	P0-C36	P0-C52	P0-C72
	P0-C89	P0-C37	P0-C53	P0-C73
Second set of DDIMMs (pink in Figure 2-11)	P0-C83	P0-C25	P0-C41	P0-C66
	P0-C84	P0-C26	P0-C43	P0-C68
	P0-C91	P0-C33	P0-C48	P0-C75
	P0-C92	P0-C34	P0-C50	P0-C77
Third set of DDIMMs (blue in Figure 2-11)	P0-C80	P0-C27	P0-C42	P0-C64
	P0-C81	P0-C29	P0-C45	P0-C67
	P0-C93	P0-C30	P0-C46	P0-C76
	P0-C95	P0-C31	P0-C49	P0-C79

Set	Location codes			
	DCM0	**DCM1**	**DCM2**	**DCM3**
Fourth set of DDIMMs (yellow in Figure 2-11)	P0-C82	P0-C24	P0-C40	P0-C65
	P0-C85	P0-C28	P0-C44	P0-C69
	P0-C90	P0-C32	P0-C47	P0-C74
	P0-C94	P0-C35	P0-C51	P0-C78

Note: The left (front) DCM0 and DCM3 are placed in a 180-degrees rotation compared to the two right (rear) DCM1 and DCM2 to optimize PCIe slots and NVMe bay wirings.

2.2.5 Pervasive memory encryption

The Power10 MCU provides the system memory interface between the on-chip SMP interconnect fabric and the OMI links. This design qualifies the MCU as an ideal functional unit to implement memory encryption logic. The Power10 on-chip MCU encrypts and decrypts all traffic to and from system memory that is based on the AES technology.

The Power10 processor supports the following modes of operation:

► AES XTS mode

XTS is the xor–encrypt–xor based tweaked-codebook mode with ciphertext stealing. AES XTS provides a block cipher with strong encryption, which is useful to encrypt persistent memory.

Persistent DIMM technology retains the data that is stored inside the memory DIMMs, even if the power is turned off. A malicious attacker who gains physical access to the DIMMs can steal memory cards. The data that is stored in the DIMMs can leave the data center in the clear if not encrypted.

Also, memory cards that leave the data center for repair or replacement can be a potential security breach. Because the attacker might have arbitrary access to the persistent DIMM data, the stronger encryption of the AES XTS mode is required for persistent memory. The AES XTS mode of the Power10 processor is supported for future use if persistent memory solutions become available for IBM Power servers.

► AES CTR mode

CTR is the *Counter* mode of operation, and it designates a low-latency AES bock cipher. Although the level of encryption is not as strong as with the XTS mode, the low-latency characteristics make it the preferred mode for memory encryption of volatile memory. AES CTR makes it more difficult to physically gain access to data through the memory card interfaces. The goal is to protect against physical attacks, which becomes increasingly important in the context of cloud deployments.

The Power10 processor-based scale-out servers support the AES CTR mode for pervasive memory encryption. Each Power10 processor holds a 128-bit encryption key that is used by the processor's MCU to encrypt the data of the DDIMMs that are attached to the OMI links.

The MCU cryptoengine is transparently integrated into the data path, which ensures that the data fetch and store bandwidth are not compromised by the AES CTR encryption mode. Because the encryption has no noticeable performance effect and because of the obvious security benefit, the pervasive memory encryption is enabled by default, and it cannot be turned off through any administrative interface.

Note: The pervasive memory encryption of the Power10 processor does not affect the encryption status of a system dump content. All data that is coming from the DDIMMs is decrypted by the MCU before it is passed onto the dump devices under the control of the dump program code. This statement applies to the traditional system dump under the OS control and the firmware assist dump utility.

Note: The PowerVM LPM data encryption does not interfere with the pervasive memory encryption. Data transfer during an LPM operation uses the following general flow:

1. On the source server, the Mover Server Partition (MSP) provides the hypervisor with a buffer.

2. The hypervisor of the source system copies the partition memory into the buffer.

3. The MSP transmits the data over the network.

4. The data is received by the MSP on the target server and copied in to the related buffer.

5. The hypervisor of the target system copies the data from the buffer into the memory space of the target partition.

To facilitate LPM data compression and encryption, the hypervisor on the source system presents the LPM buffer to the on-chip NX unit as part of process in step 2. The reverse decryption and decompress operation is applied on the target server as part of the process in step 4.

The pervasive memory encryption logic of the MCU decrypts the memory data before it is compressed and encrypted by the NX unit on the source server. The logic also encrypts the data before it is written to memory but after it is decrypted and decompressed by the NX unit of the target server.

2.2.6 Active Memory Mirroring

The Power E1050 server can mirror the Power Hypervisor code across multiple memory DDIMMs. If a DDIMM that contains the hypervisor code develops an uncorrectable error, its mirrored partner enables the system to continue to operate uninterrupted.

Active Memory Mirroring (AMM) is an optional feature (#EM81).

The hypervisor code logical memory blocks are mirrored on distinct DDIMMs to enable more usable memory. There is no specific DDIMM that hosts the hypervisor memory blocks, so the mirroring is done at the logical memory block level, not at the DDIMM level. To enable the AMM feature, the server must have enough free memory to accommodate the mirrored memory blocks.

Besides the hypervisor code itself, other components that are vital to the server operation are also mirrored:

► Hardware page tables (HPTs), which are responsible for tracking the state of the memory pages that are assigned to partitions

► Translation control entities (TCEs), which are responsible for providing I/O buffers for the partition's communications

► Memory that is used by the hypervisor to maintain partition configuration, I/O states, virtual I/O information, and partition state

It is possible to check whether the AMM option is enabled and changes its status by using the HMC. The relevant information and controls are in the Memory Mirroring section of the General Settings window of the selected Power E1050 server (Figure 2-12).
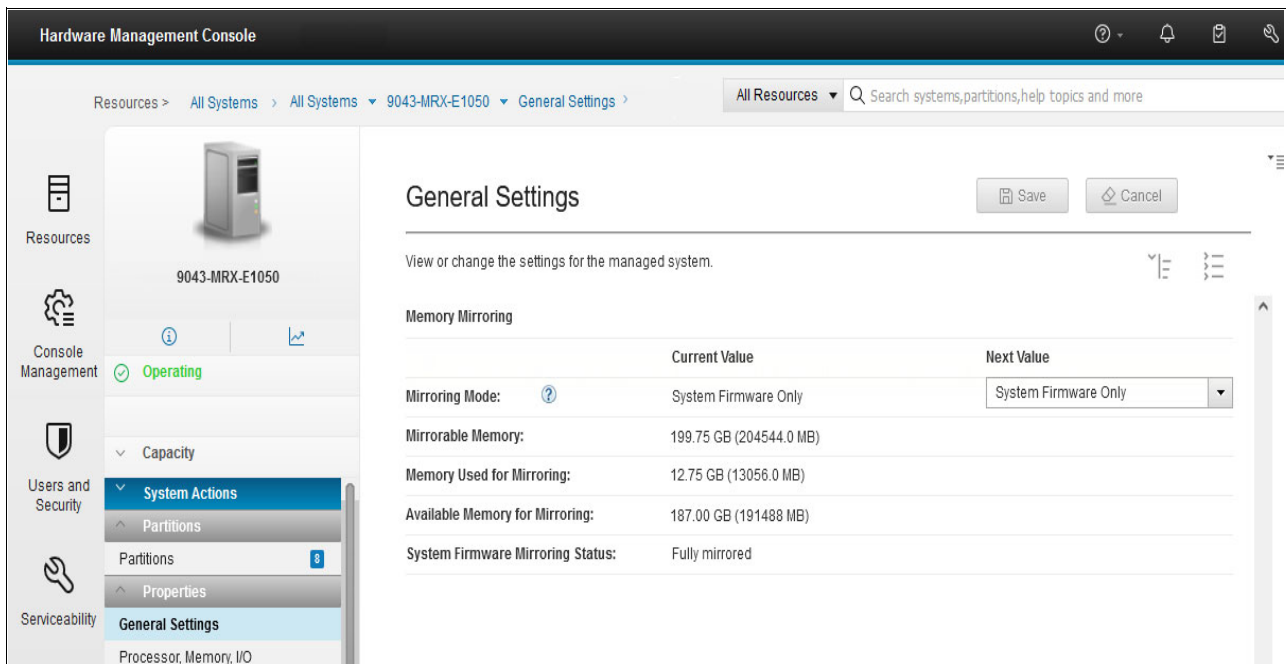
Figure 2-12   Memory Mirroring section in the General Settings window on the HMC enhanced GUI

After a failure occurs on one of the DDIMMs that contains hypervisor data, all the server operations remain active and the enterprise Baseboard Management Controller (eBMC) service processor isolates the failing DDIMMs. The system stays in the partially mirrored state until the failing DDIMM is replaced.

Memory that is used to hold the contents of platform dumps is not mirrored, and AMM does *not* mirror partition data. AMM mirrors only the hypervisor code and its components to protect this data against a DDIMM failure. With AMM, uncorrectable errors in data that are owned by a partition or application are handled by the existing Special Uncorrectable Error (SUE) handling methods in the hardware, firmware, and OS.

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

2.3 Internal I/O subsystem

The internal I/O subsystem of the Power E1050 server is connected to the PCIe Express controllers on a Power10 chip in the system. A Power10 chip has two PCI Express controllers (PECs) of 16 lanes each for a total of 32 Gen5/Gen4 lanes per chip and 64 Gen5/Gen4 lanes per DCM.

Each PEC supports up to three PCI host bridges (PHBs) that directly connect to PCIe slots or devices. Both PEC0 and PEC1 can be configured as follows:

► One x16 Gen4 PHB or one x8 Gen5 PHB

► One x8 Gen5 and one x8 Gen4 PHB

► One x8 Gen5 PHB and two x4 Gen4 PHBs

The usage or configurations of the PECs are shown in the notation of the ports. There are two notations, the E-Bus notation and the PHB notation, which describe the split of the PEC. Table 2-9 gives an overview.

Table 2-9   PHB cross-reference to E-Bus notation

PEC configuration	E-Bus ports	PHB
One x16 at DCM0	E0	PHB0
Two x8 at DCM0	E0A and E0B	PHB0 and PHB1
One x8 and two x4 at DCM0	E0A, E0B, and E0C	PHB0, PHB1, and PHB2
One x16 at DCM1	E1	PHB3
Two x8 at DCM1	E1A and E1B	PHB3 and PHB4
One x8 and two x4 at DCM1	E1A, E1B, and E1C	PHB3, PHB4, and PHB5

Figure 2-13 shows a diagram of the I/O subsystem of the Power E1050 server. The left (front) DCM0 and DCM3 are placed in a 180-degrees rotation compared to the two right (rear) DCM1 and DCM2 to optimize PCIe slots and NVMe bay wirings.
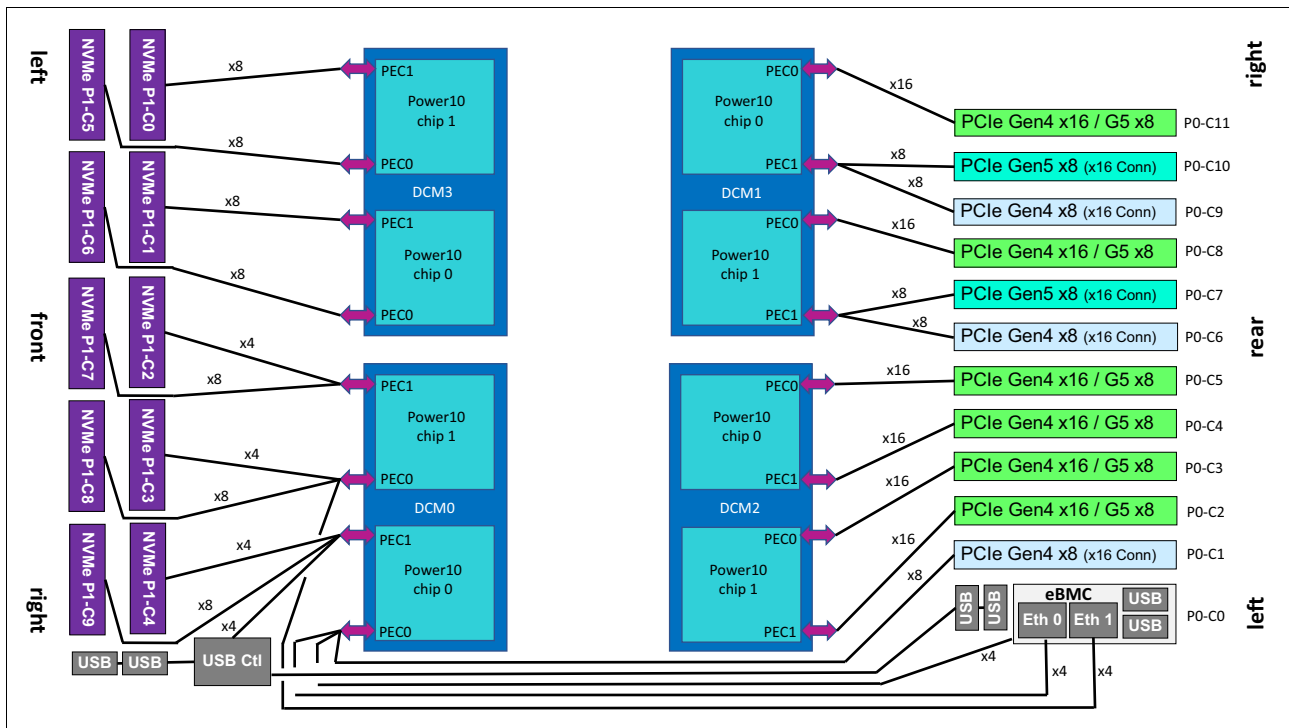
Figure 2-13   Power E1050 I/O subsystem diagram

On the left (front) side, you find 10 NVMe bays. Six NVMe bays are connected to DCM0 and four NVMe bays are connected to DCM3. To make all 10 NVMe bays available, all four processor sockets must be populated. The NVMe bays are connected mainly by using a x8 PHB, some with a x4 PHB. But because the NVMe devices can use four lanes (x4), this fact is not relevant from a performance point of view.

On the right (rear) side, you find 11 PCIe slots in a different manner. Six slots use up to 16 lanes (x16) and can operate either in PCI Gen4 x16 mode or in Gen5 x8 mode. The remaining five slots are connected with eight lanes. Two of them are Gen5 x8, and three of them are Gen4 x8. In a 2-socket processor configuration, seven slots are available for use (P0-C1 and P0-C6 to P0-C11). To make all the slots available, at least three processor sockets must be populated.

The x16 slots can provide up to twice the bandwidth of x8 slots because they offer twice as many PCIe lanes. PCIe Gen5 slots can support up to twice the bandwidth of a PCIe Gen4 slot, and PCIe Gen4 slots can support up to twice the bandwidth of a PCIe Gen3 slot, assuming an equivalent number of PCIe lanes.

Note: Although some slots provide a x8 connection only, all slots have an x16 connector.

All PCIe slots support hot-plug adapter installation and maintenance and enhanced error handling (EEH). PCIe EEH-enabled adapters respond to a special data packet that is generated from the affected PCIe 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 the most devices, although some third-party PCI devices might not provide native EEH support.

All PCIe adapter slots support hardware-backed network virtualization through single-root IO virtualization (SR-IOV) technology. Configuring an SR-IOV adapter into SR-IOV shared mode might require more hypervisor memory. If sufficient hypervisor memory is not available, the request to move to SR-IOV shared mode fails. The user is instructed to free extra memory and try the operation again.

The server PCIe slots are allocated DMA space by using the following algorithm:

► All slots are allocated a 2 GB default DMA window.

► All I/O adapter slots (except the embedded Universal Serial Bus (USB)) are allocated Dynamic DMA Window (DDW) capability based on installed platform memory. DDW capability is calculated assuming 4 K I/O mappings:

 – The slots are allocated 64 GB of DDW capability.

 – Slots can be enabled with Huge Dynamic DMA Window (HDDW) capability by using the I/O Adapter Enlarged Capacity setting in the ASMI.

 – HDDW-enabled slots are allocated enough DDW capability to map all installed platform memory by using 64 K I/O mappings.

 – Slots that are HDDW-enabled are allocated the larger of the calculated DDW capability or HDDW capability.

The Power E1050 server is smarter about energy efficiency when cooling the PCIe adapter environment. It senses which IBM PCIe adapters are installed in their PCIe slots, and if an adapter requires higher levels of cooling, they automatically speed up fans to increase airflow across the PCIe adapters. Faster fans increase the sound level of the server. Higher wattage PCIe adapters include the PCIe3 serial-attached SCSI (SAS) adapters and solid-state drive (SSD)/flash PCIe adapters (#EJ10, #EJ14, and #EJ0J).

USB ports

The first DCM (DCM0) also hosts the USB controller that is connected by using four PHBs, although the USB controller uses only one lane. DCM0 provides four USB 3.0 ports, with two in the front and two in the back. The two front ports provide up to 1.5 A USB current, mainly to support the external USB DVD (Feature Code EUA5). The two rear USB ports current capacity is 0.9 A.

> **Note:** The USB controller is placed on the trusted platform module (TPM) card because of space reasons.

Some customers require that USB ports must be deactivated for security reasons. You can achieve this task by using the ASMI menu. For more information, see 2.5.1, "Managing the system by using the ASMI GUI" on page 73.

2.3.1 PCIe adapter slot details

The PCIe adapter slots have different characteristics. Table 2-10 lists the PCIe adapter slot locations, the slot type, the connection to the processor, the OpenCAPI capability, the support to host a PCIe Gen4 cable card (#EJ2A) to attach the EMX0 PCIe Gen3 I/O expansion drawer, and the I/O adapter enlarged capacity enablement order.

Table 2-10   PCIe slot locations and capabilities for the Power E1080 servers

Location code	Description	Processor module	OpenCAPI	Cable card for I/O drawer	I/O adapter enlarged capacity enablement order[a]
P0-C0[b]	eBMC	DCM0-P0-E0-PHB1	N/A	N/A	N/A
P0-C1[c]	PCIe4 x8	DCM0-P0-E0-PHB0	Yes	No	9
P0-C2	PCIe5 x8 or PCIe4 x16	DCM2-P1-E1-PHB3	Yes	Yes	1
P0-C3	PCIe5 x8 or PCIe4 x16	DCM2-P1-E0-PHB0	Yes	Yes	5
P0-C4	PCIe5 x8 or PCIe4 x16	DCM2-P0-E1-PHB3	No	Yes	2
P0-C5	PCIe5 x8 or PCIe4 x16	DCM2-P0-E0-PHB0	No	Yes	6
P0-C6	PCIe4 x8	DCM1-P1-E1-PHB4	Yes	No	10
P0-C7	PCIe5 x8	DCM1-P1-E1-PHB3	Yes	Yes	7
P0-C8	PCIe5 x8 or PCIe4 x16	DCM1-P1-E0-PHB0	Yes	Yes	3
P0-C9	PCIe4 x8	DCM1-P0-E1-PHB4	Yes	No	11
P0-C10	PCIe5 x8	DCM1-P0-E1-PHB3	Yes	Yes	8
P0-C11	PCIe5 x8 or PCIe4 x16	DCM1-P0-E0-PHB0	Yes	Yes	4

a. Enabling the I/O adapter enlarged capacity option affects only Linux partitions.
b. Only used for an eBMC card.
c. P0-C1 is generally used for the base Ethernet adapter.

The following characteristics are not part of the table because they apply for all slots:

► All slots are SR-IOV capable.

► All slots have an x16 connector, although some provide only a x8 connection.

► All PCIe adapters are installed in an I/O Blind-Swap Cassette (BSC) as the basis for concurrent maintenance. For more information, see 2.3.2, "I/O Blind-Swap Cassettes" on page 69.

► All I/O cassettes can hold Half Length Full Height (HLFH) and Half Length Half Height (HLHH) PCIe adapters.

Figure 2-14 shows the rear view of the Power E1050 server with the location codes for the PCIe adapter slots.
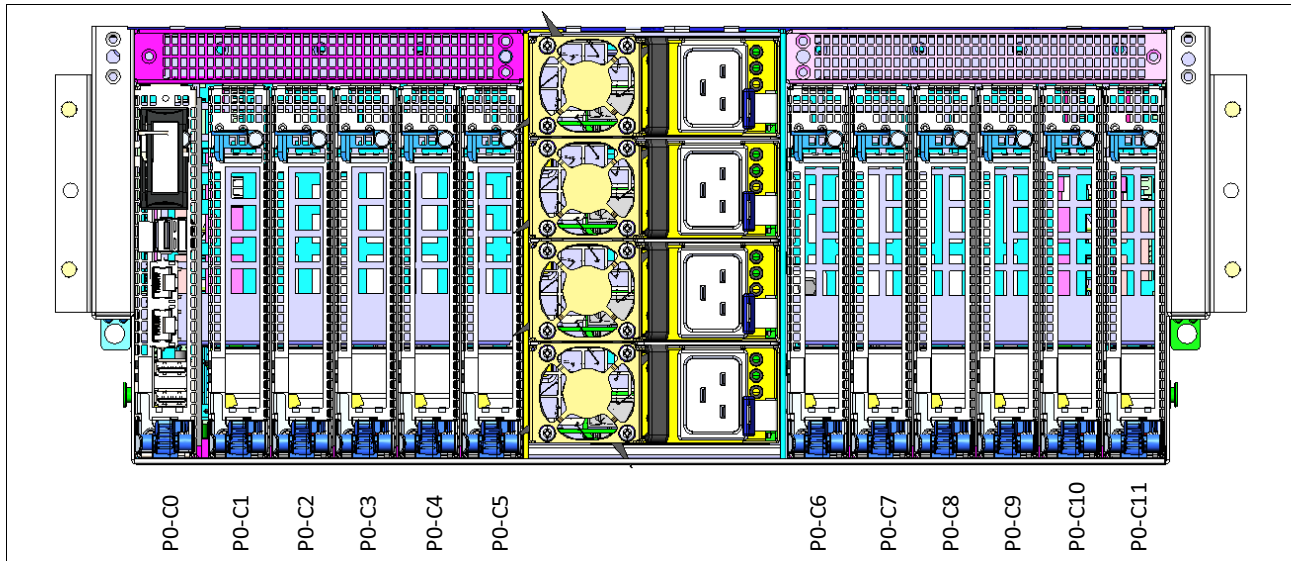
Figure 2-14   Rear view of a Power E1050 server with PCIe slots location codes

Note: Slot P0-C0 is not a PCIe slot; instead, it holds a special I/O Cassette for the eBMC Service Processor Card.

2.3.2 I/O Blind-Swap Cassettes

The Power E1050 server supports I/O BSCs for easy hot-plugging of PCIe adapters. PCIe adapters are installed inside a BSC before inserting it into the system from the rear. Each BSC has one x16 PCIe connector. There is always a BCS in each slot (C1 - C11), either with an adapter or a blank. All PCIe slots C1 - C11 are concurrently maintainable by using a BCS.

Caution: To hot plug a BCS, first go to the HMC or AIX diag to start a hot-plug action to remove the power from a slot. Do not pull a cassette when the slot is still under power.

Slot C0 is a special I/O cassette for the eBMC Service Processor card. The eBMC card is not concurrently maintainable. For more information, see 2.3.5, "System ports" on page 71.

Figure 2-15 is an illustration of the BCS. The Power E1050 I/O BSC is like the one for the Power E950 server, but it has an impel backplane connector to support Gen4 PCIe. Its lock latch is redesigned to include initial turns before engaging the BCS to make the latch more robust.
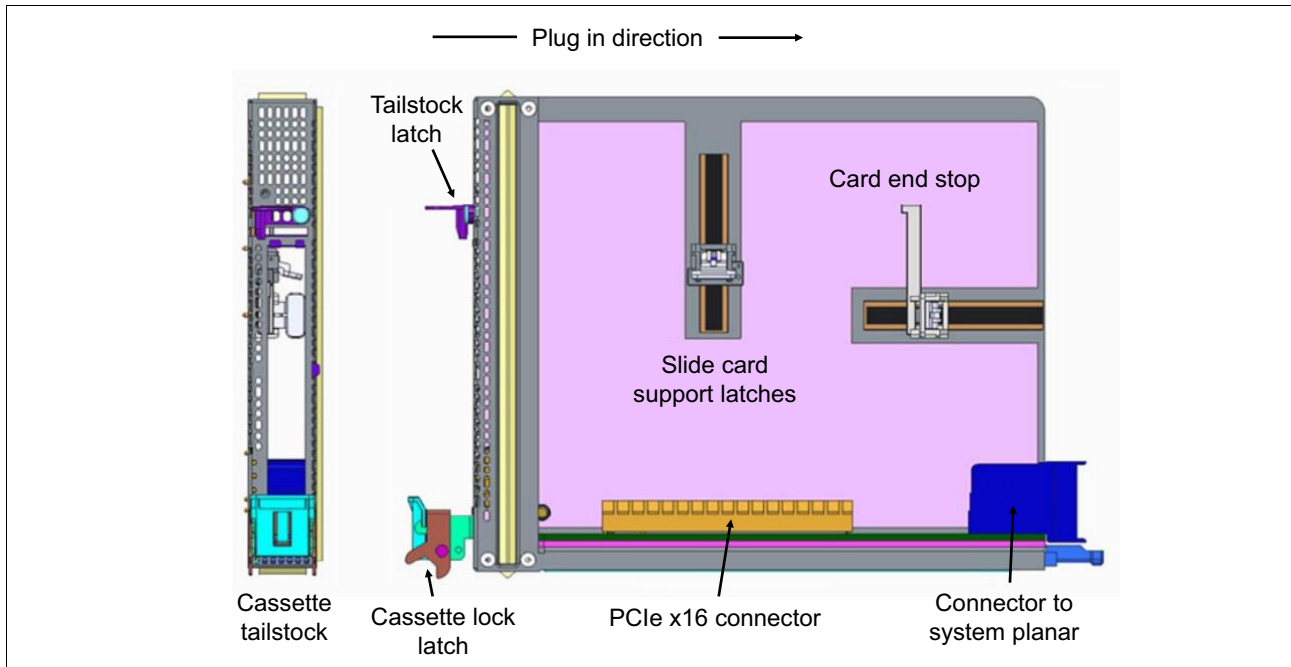
Figure 2-15   Power E1050 Blind-Swap Cassette

2.3.3 Non-volatile Memory Express bays

The IBM Power E1050 server has a 10-NVMe backplane that is always present in the server. It offers up to 10 NVMe bays with all four processor sockets populated and six NVMe bays if only two or three processor sockets are populated. It connects to the system board through three Molex Impact connectors and a power connector. There is no cable between the NVMe backplane and the system board.

The wiring strategy and backplane materials are chosen to ensure Gen4 signaling to all NVMe drives. All NVMe connectors are PCIe Gen4 connectors. For more information about the internal connection of the NVMe bays to the processor chips, see Figure 2-13 on page 66.

Each NVMe interface is a Gen4 x4 PCIe bus. The NVMe drives can be in an OS-controlled RAID array. A hardware RAID is not supported on the NVMe drives. The NVMe thermal design supports 18 W for 15-mm NVMe drives and 12 W for 7-mm NVMe drives.

For more information about the available NVMe drives and how to plug the drives for best availability, see 3.5, "Internal storage" on page 92.

2.3.4 Attachment of I/O-drawers

The Power E1050 server can expand the number of its I/O slots by using I/O Expansion Drawers (#EMX0). The number of I/O drawers that can be attached to an Power E1050 server depends on the number of populated processor slots, which changes the number of available internal PCIe slots of the server. Only some slots can be used to attach an I/O Expansion Drawer by using the #EJ2A CXP Converter adapter, also referred to as a cable card.

Feature Code #EJ2A is an IBM designed PCIe Gen4 x16 cable card. It is the only supported cable card to attach fanout modules of an I/O Expansion Drawer in the Power E1050 server. Previous cards from a Power E950 server cannot be used. Feature Code #EJ2A supports copper and optical cables for the attachment of a fanout module.

> **Note:** The IBM e-config configurator adds 3-meter copper cables (Feature Code #ECCS) to the configuration if no cables are manually specified. If you want to have optical cables make sure to configure them.

Table 2-11 lists the PCIe slot order for the attachment of an I/O Expansion Drawer, the maximum number of I/O Expansion Drawers and Fanout modules, and the maximum number of available slots (dependent on the populated processor sockets).

Table 2-11   I/O Expansion Drawer capabilities depend on the number of populated processor slots

Processor sockets populated	Expansion adapter slots order	Maximum number of I/O Expansion Drawers	Maximum number of fanout modules	Total PCIe slots
DCM0 and DCM1	P0-C8, P0-C11, P0-C7, and P0-C10	2	4	27
DCM0, DCM1, and DCM2	P0-C8, P0-C11, P0-C2, P0-C3, P0-C4, and P0-C5	3	6	41
DCM0, DCM1, DCM2, and DCM3	P0-C8, P0-C11, P0-C2, P0-C3, P0-C4, P0-C5, P0-C7, and P0-C10	4	8	51

For more information about the #EMX0 I/O Expansion Drawer, see 3.9.1, "PCIe Gen3 I/O expansion drawer" on page 99.

2.3.5 System ports

The Power E1050 server has two 1-Gbit Ethernet ports and two USB 2.0 ports to connect to the eBMC service processor. The two eBMC Ethernet ports are used to connect one or two HMCs. There are no other HMC ports, as in servers that have a Flexible Service Processor (FSP). The eBMC USB ports can be used for a firmware update from a USB stick.

The two eBMC Ethernet ports are connected by using four PCIe lanes each, although the eBMC Ethernet controllers need only one lane. The connection is provided by the DCM0, one from each Power10 chip. For more information, see Figure 2-13 on page 66.

The eBMC module with its two eBMC USB ports also is connected to the DCM0 at chip 0 by using a x4 PHB, although the eBMC module uses only one lane.

For more information about attaching an Power E1050 server to an HMC, see Accessing the eBMC so that you can manage the system.

For more information about how to do a firmware update by using the eBMC USB ports, see Installing the server firmware on the service processor or eBMC through a USB port.

2.4 Component summary per processor socket configuration

The Power E1050 server is modular and every populated processor module activates more features and ports. This section provides an overview about the features that are available in a 2-, 3-, and 4-processor configurations.

The following features are available in a *2-socket configuration*:

► Two processor sockets for two Power10 processors with 24 - 48 cores (installed)
► Thirty-two OMI memory ports for 256 GB, with up to 8 TB of memory
► Seven PCIe ports:
    – Two PCIe Gen4 x16 / PCIe Gen5 x8
    – Two PCIe Gen5 x8
    – Three PCIe Gen4 x8
► Up to two I/O Expansion drawers supporting:
    – Four fanout modules
    – Twenty-seven PCIe slots (server total)
► Six individual NVMe ports supporting up to 38.4 TB
► Four USB 3.0 ports
► Two eBMC 1-Gbit Ethernet ports
► Two eBMC 2.0 USB ports

The following features are available in a *3-socket configuration*:

► Three processor sockets for three Power10 processors with 36 - 72 cores (installed)
► Forty-eight OMI memory ports for 384 GB, with up to 12 TB of memory
► Eleven PCIe ports:
    – Six PCIe Gen4 x16 / PCIe Gen5 x8
    – Two PCIe Gen5 x8
    – Three PCIe Gen4 x8
► Up to three I/O Expansion drawers supporting:
    – Six fanout modules
    – Forty-one PCIe slots (server total)
► Six individual NVMe ports supporting up to 38.4 TB
► Four USB 3.0 ports
► Two eBMC 1-Gbit Ethernet ports
► Two eBMC 2.0 USB ports

The following features are available in a *4-socket configuration*:

► Four processor sockets for four Power10 processors with 48 - 96 cores (installed)
► Sixty-four OMI memory ports for 512 GB, with up to 16 TB of memory
► Eleven PCIe ports:
    – Six PCIe Gen4 x16 / PCIe Gen5 x8
    – Two PCIe Gen5 x8
    – Three PCIe Gen4 x8
► Up to four I/O Expansion drawers supporting:
    – Eight fanout modules
    – Fifty-one PCIe slots (server total)
► Ten individual NVMe ports supporting up to 64 TB
► Four USB 3.0 ports
► Two eBMC 1-Gbit Ethernet ports
► Two eBMC 2.0 USB ports

2.5 The enterprise Baseboard Management Controller

The Power E1050 server and the Power10 scale-out systems use an eBMC for system service management, monitoring, maintenance, and control. The eBMC also provides access to the system event log (SEL) files. The eBMC is a specialized service processor that monitors the physical state of the system by using sensors. A system administrator or service representative can communicate with the eBMC through an independent connection.
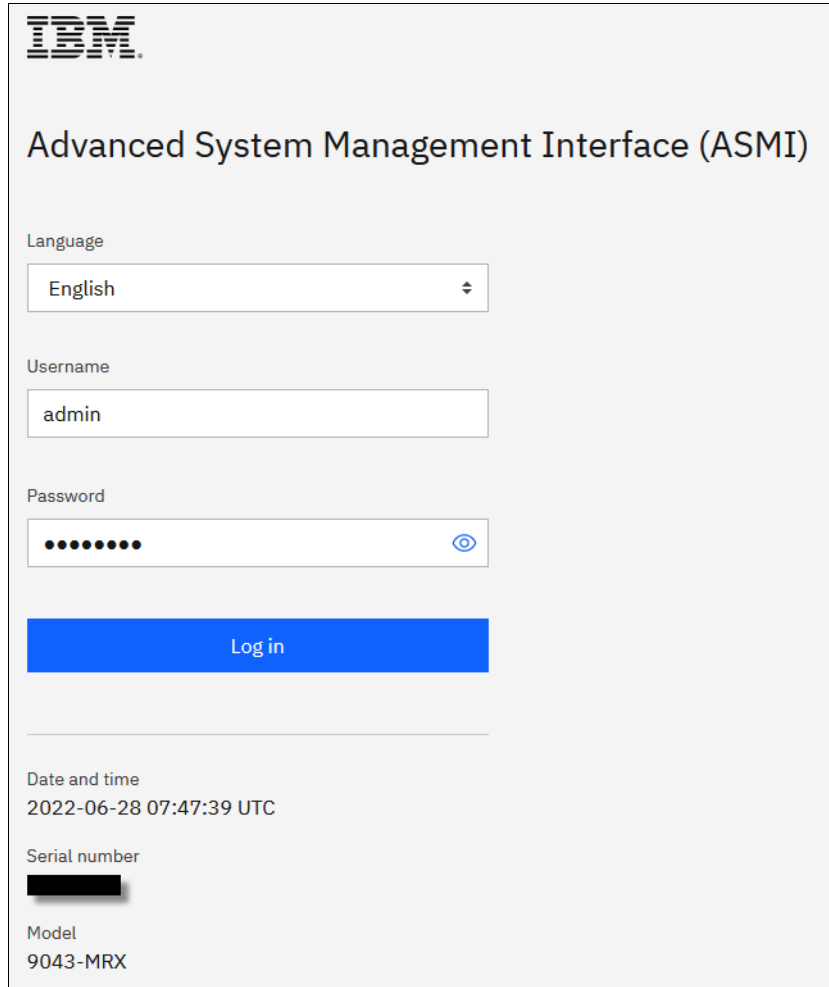
2.5.1 Managing the system by using the ASMI GUI

The ASMI is the GUI to the eBMC. It is similar in terms of its function to the ASMI of FSP-managed servers (for example, the Power E1080 server), but it is a complete redesign of the UI driven through customer feedback of a Design Thinking workshop.

To enter the ASMI GUI, go to an HMC, select the server, and then select **Operations** → **Launch Advanced System Management**. A window opens and shows the name of the system, MTM and serial number, and the IP address of the service processor (eBMC). Click **OK**, and the ASMI window opens.

If the eBMC is not in a private but reachable network, you can connect directly by entering `https://<eBMC IP>` into your web browser.

Figure 2-16 show the login window for the ASMI.

Figure 2-16   ASMI login window

The default user when you log in the first time is `admin` and the default password is also `admin` but invalidated. After the first login, you must immediately change the admin password, which is also true after performing a factory reset of the system. This policy change ensures that the eBMC is not left in a state with a well-known password. The password needs to be a strong one, and not, for example, abcd1234. For more information about the password rules, see Setting the password.

After login, you see the Overview page with server, firmware, network, power, and status information, as shown in Figure 2-17.
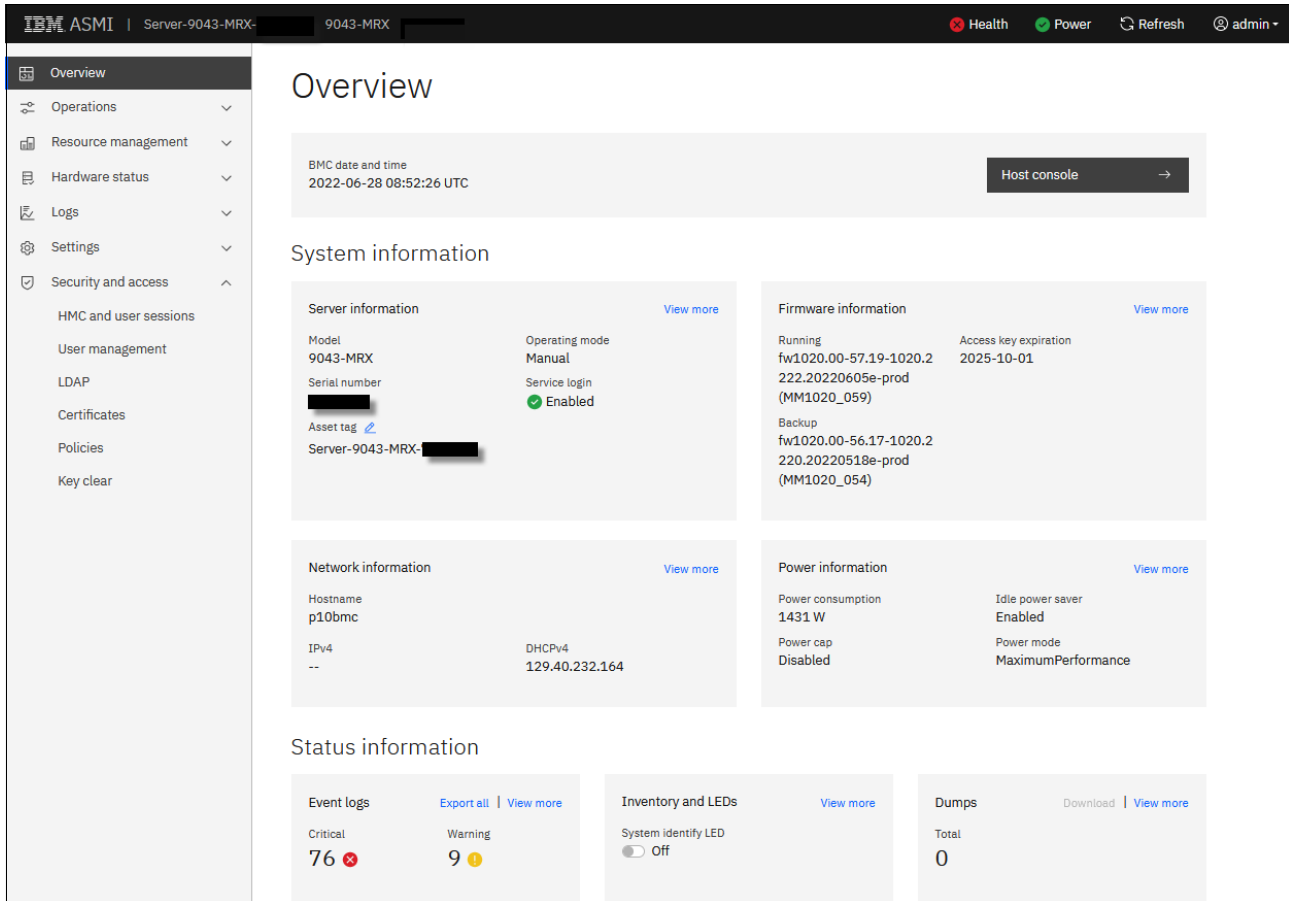
Figure 2-17   ASMI Overview window

The new ASMI for eBMC managed servers have some major differences and some valuable new features:

► System firmware updates

It is possible to install a firmware update for the server by using the ASMI GUI, even if the system is managed by an HMC. In this case, the firmware update is always disruptive. To install a concurrent firmware update, you must use the HMC.

► Download of dumps

Dumps can be downloaded by using the HMC, but if necessary, you also can download them from the ASMI menu.

It is also possible to initiate a dump from the ASMI by selecting **Logs** → **Dumps**, selecting the dump type, and clicking **Initiate dump**. The possible dump types are:

– Baseboard management controller (BMC) dump (nondisruptive)

– Resource dump

– System dump (disruptive)

► Network Time Protocol (NTP) server support

► Lightweight directory access protocol (LDAP) for user management

► Host console

By using the host console, you can watch the boot process of the server. The host console also can be used to access the OS if there is only a single LPAR using all resources.

> **Note:** The host console can also be accessed by using an SSH client over port 2200 and logging in with the admin user.

► User management

In the eBMC, it is possible to create you own users. This feature also can be used to create an individual user that can be used for the HMC to access the server. There are two types of privileges for a user: Administrator or ReadOnly. As the name indicates, with the ReadOnly privileges, you cannot modify something (except the password of that user), and a user with that privilege cannot be used for HMC access to the server.

► IBM Security® through Access Control Files (ACFs)

In FSP-managed servers, IBM Support generates a password by using the serial number and the date to get "root access" to the service processor by using the user celogin. In eBMC managed systems, IBM support generates an ACF. This file must be uploaded to the server to get access. This procedure is, for example, needed if you lost the admin password and want to reset it.

► Jumper reset

It is possible to reset everything on the server by using a jumper. This reset is a factory reset that resets everything, like LPARs, eBMC settings, or the NVRAM.

Real-time progress indicator

The ASMI of an eBMC server also provides a real-time progress indicator to see the operator panel codes. To open the windows that show the codes, select **Logs** → **Progress logs**, and then click **View code in real time**.

Inventory and LEDs

Under the **Hardware status** → **Inventory and LEDs** menu, you find most of the hardware components with their state and an identification LED, and the system identify LED and the System attention LED. You can turn on and off all identification LEDs. Only the System attention LED can be turned off.

It is also possible to display the details of a component, which is helpful to see details such as the size of a DIMM or the part numbers if something must be exchanged.

Sensors

The ASMI has many sensors for the server, which you can access by selecting **Hardware status** → **Sensors**. The loading of the sensors takes some time, and during that time you see a progress bar on the top of the window.

> **Note:** Although the progress bar might be finished, it might take some extra time until the sensors appear.

Figure 2-18 shows an example of the Sensors window.
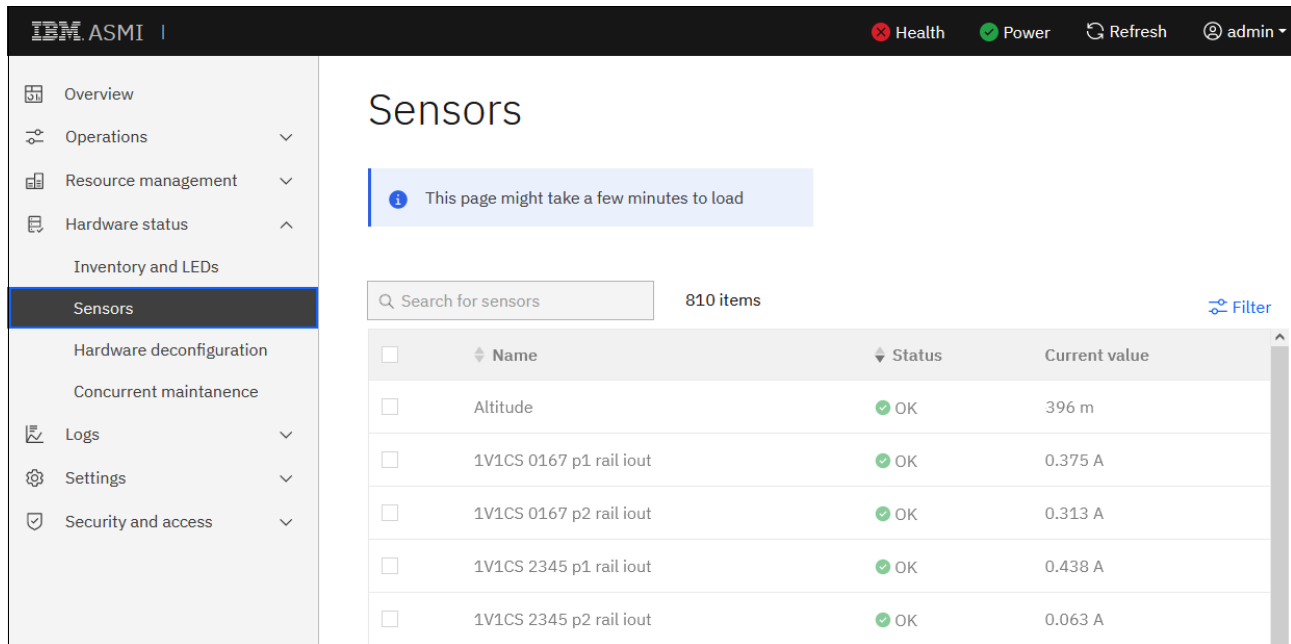
Figure 2-18   ASMI Sensors window

Network settings

The default network setting for the two eBMC ports is DHCP. Therefore, when you connect a port to a private HMC network where the HMC is connected to a DHCP server, the new system should get its IP address from the HMC during the startup of the firmware. Then, the new system automatically appears in the HMC and can be configured. DHCP is the IBM recommended way to attach a server to the HMC.

If you do not use DHCP and want to use a static IP address, you can set the IP address in the ASMI GUI. However, because there are no default IP addresses that are the same for every server, you first must discover the configured IP address.

To discover the configured IP address, use the operator panel and complete the following steps:

1. Use the Increment or Decrement buttons to scroll to function 02.

2. Press Enter until the value changes from N to M, which activates access to function 30.

3. Scroll to function 30 and press Enter. Function 30** appears.

4. Scroll to 3000 and press Enter, which shows you the IP address of the eth0 port.

5. If you scroll to 3001 and press Enter, you see the IP address of eth1.

6. After you discover the IP address, scroll again to function 02 and set the value back from M to N.

For more information about function 30 in the operator panel, see IBM Documentation.

Now that you have discovered the IP address, you can connect any computer with a web browser to an IP address in the same subnet (class C) and connect the computer with the correct Ethernet port of the Power E1050 server.

Hint: Most connections work by using a standard Ethernet cable. If you do not see a link when using a standard Ethernet cable, try a crossover cable, where the send and receive wires are crossed.

After connecting the cable, you can use a web browser to access the ASMI by using the URL `https://<IP address>`, and then you can configure the network ports. To configure the network ports, select **Settings** → **Network** and select the correct adapter to configure. Figure 2-19 shows an example of changing eth1. Before you can configure a static IP address, turn off DHCP. It is possible to configure several static IP addresses on one physical Ethernet port.
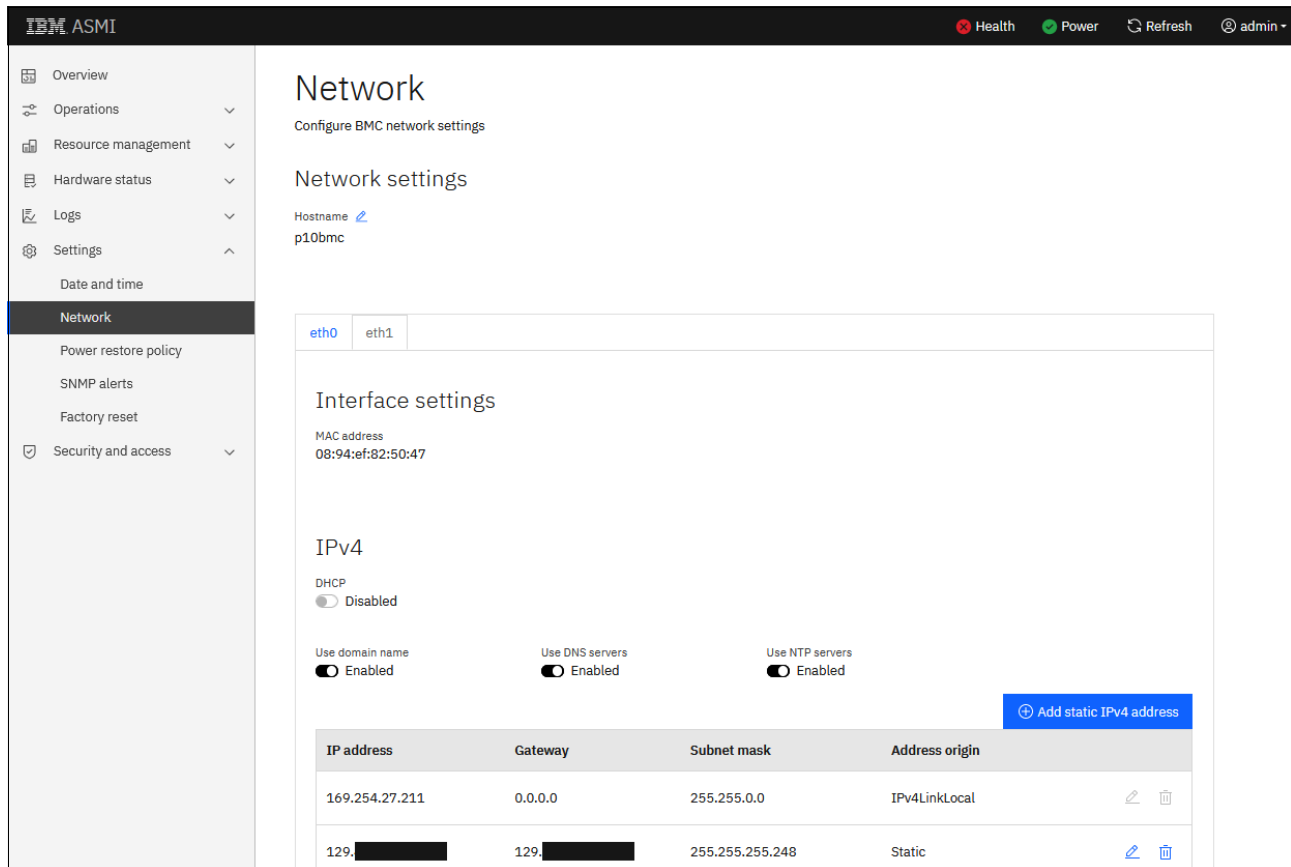
Figure 2-19   ASMI network settings

You cannot configure the Virtualization Management Interface (VMI) address in the ASMI network settings. The VMI address is another IP address that is configured on the physical eBMC Ethernet port of the server to manage the virtualization of the server. The VMI address can be configured only in the HMC.

Using an Access Control File

If you lose the access password for the ASMI service user, you can access the ASMI by using an ACF. An ACF is a digital certificate that is provided by IBM Support when you open a support case. To use the ACF, the system must be enabled at the server by using the operator panel.

Complete the following steps:

1. On the operator panel, use the Increment or Decrement buttons to scroll to function 74.

2. Press Enter, and select 00 to accept the function (FF rejects it).

3. The ACF function now is active for 30 minutes. To use it, access the ASMI login window.

4. To upload the ACF into the system and access the ASMI, click **Upload service login certificate**.

For more information, see IBM Documentation.

Policies

In the **Security and access** → **Policies** menu, you can turn on and off security-related functions, for example, whether you can manage your server by using Intelligent Platform Management Interface (IPMI).

Some customers require that the USB ports of the server should be disabled. You can accomplish this task by using policies. To do so, clear **Host USB enablement**, as shown in Figure 2-20.
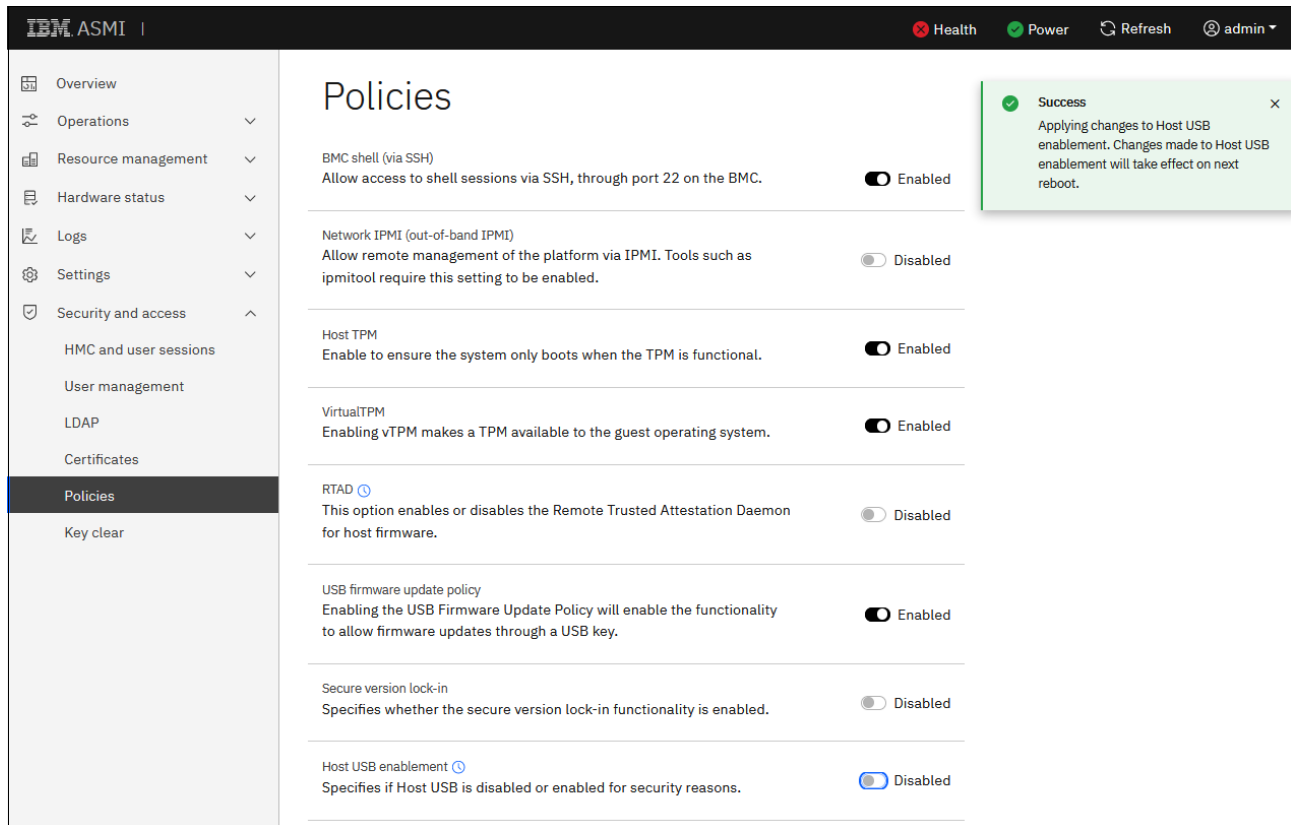
Figure 2-20   How to turn off the USB ports of the server

2.5.2 Managing the system by using DMTF Redfish

eBMC-based systems can be managed by using the DMTF Redfish application programming interfaces (APIs). Redfish is a Representational State Transfer (REST) API that is used for platform management and is standardized by the Distributed Management Task Force, Inc.

There are multiple ways of working with Redfish. One possibility is to use the OS command **curl**. The following examples show how to work with **curl** and Redfish.

Before you can get data from the server or run systems management tasks by using Redfish, you must authenticate against the server. After you authenticate by using a user ID and password, you receive a token from the serve that you use later.

Example 2-1 show how to obtain a token.

Example 2-1   Obtaining a token from Redfish

```
# export eBMC=<IP>
# export USER=admin
# export PASSWORD=<PW>

# export TOKEN=`curl -k -H "Content-Type: application/json" -X POST https://${eBMC}/login
-d "{\"username\" : \"${USER}\", \"password\" : \"${PASSWORD}\"}" | grep token | awk
'{print $2;}' | tr -d '"'`
```

With the token, you can get data from the server. First, request the data of the Redfish root by using /Redfish/v1. You get data with additional branches in the Redfish tree, for example, Chassis (uppercase C). To dig deeper, use the newly discovered odata.id /Redfish/v1/Chassis, as shown in Example 2-2.

Example 2-2   Getting chassis data from Redfish

```
#curl -s -k -H "X-Auth-Token: $TOKEN" -X GET https://${eBMC}/redfish/v1
{
  "@odata.id": "/redfish/v1",
  "@odata.type": "#ServiceRoot.v1_12_0.ServiceRoot",
  "AccountService": {
    "@odata.id": "/redfish/v1/AccountService"
  },
  "Cables": {
    "@odata.id": "/redfish/v1/Cables"
  },
  "CertificateService": {
    "@odata.id": "/redfish/v1/CertificateService"
  },
  "Chassis": {
    "@odata.id": "/redfish/v1/Chassis"
  },

# curl -k -H "X-Auth-Token: $TOKEN" -X GET https://${eBMC}/Redfish/v1/Chassis
{
  "@odata.id": "/Redfish/v1/Chassis",
  "@odata.type": "#ChassisCollection.ChassisCollection",
  "Members": [
    {
      "@odata.id": "/Redfish/v1/Chassis/chassis"
    }
  ],
  "Members@odata.count": 1,
  "Name": "Chassis Collection"

# curl -k -H "X-Auth-Token: $TOKEN" -X GET https://${eBMC}/Redfish/v1/Chassis/chassis
{
  "@odata.id": "/Redfish/v1/Chassis/chassis",
  "@odata.type": "#Chassis.v1_16_0.Chassis",
  "Actions": {
```

```
...
"PCIeSlots": {
    "@odata.id": "/Redfish/v1/Chassis/chassis/PCIeSlots"
  },
...
"Sensors": {
    "@odata.id": "/Redfish/v1/Chassis/chassis/Sensors"
  },
...
```

Under Chassis (uppercase C), there is another chassis with a lowercase c. Use the tree with both (/Redfish/c1/Chassis/chassis). After running the chassis, you can see in Example 2-2 on page 80 for example, that there are PCIeSlots and Sensors, among other resources of the server.

In Example 2-3, you can see what is under Sensors. There you can find the same sensors as in the ASMI GUI (see Figure 2-18 on page 77). In the output, you find, for example, the sensor total_power. When you ask for details about that sensor, as shown in Example 2-3, you can see that the server needed 1.426 watts at the time of running the command.

Example 2-3   Getting sensor data from Redfish

```
# curl -k -H "X-Auth-Token: $TOKEN" -X GET
https://${eBMC}/Redfish/v1/Chassis/chassis/Sensors
{
  "@odata.id": "/Redfish/v1/Chassis/chassis/Sensors",
  "@odata.type": "#SensorCollection.SensorCollection",
  "Description": "Collection of Sensors for this Chassis",
  "Members": [
    {
      "@odata.id": "/Redfish/v1/Chassis/chassis/Sensors/Altitude"
    },
    {
      "@odata.id": "/Redfish/v1/Chassis/chassis/Sensors/1V1CS_0167_p1_rail_iout"
    },
...
    {
      "@odata.id": "/Redfish/v1/Chassis/chassis/Sensors/total_power"
    },

# curl -k -H "X-Auth-Token: $TOKEN" -X GET \
https://${eBMC}/Redfish/v1/Chassis/chassis/Sensors/total_power
{
  "@odata.id": "/Redfish/v1/Chassis/chassis/Sensors/total_power",
  "@odata.type": "#Sensor.v1_0_0.Sensor",
  "Id": "total_power",
  "Name": "total power",
  "Reading": 1426.0,
  "ReadingRangeMax": null,
  "ReadingRangeMin": null,
  "ReadingType": "Power",
  "ReadingUnits": "W",
  "Status": {
    "Health": "OK",
    "State": "Enabled"
  }
```

Some other useful strings to ask for are shown in Example 2-4.

Example 2-4   Other useful Redfish data

```
#Type and model of the server
   # curl -s -k -H "X-Auth-Token: $TOKEN" -X GET https://${eBMC}/Redfish/v1/Systems/system
   | grep Model | grep -v SubModel | grep \ -v \"\"
   "Model": "9043-MRX",
#Serial number
   # curl -s -k -H "X-Auth-Token: $TOKEN" -X GET https://${eBMC}/redfish/v1/Systems/system
   | grep SerialNumber \
   "SerialNumber": "<SN>",
#Type, model, and serial number
   #curl -s -k -H "X-Auth-Token: $TOKEN" -X GET https://${eBMC}/redfish/v1/Systems/system |
   grep AssetTag
   "AssetTag": "Server-9043-MRX-<SN>",
#System indicator LED
   #curl -s -k -H "X-Auth-Token: $TOKEN" -X GET https://${eBMC}/redfish/v1/Systems/system |
   grep IndicatorLED
   "IndicatorLED": "Off",
#Total memory
   # curl -s -k -H "X-Auth-Token: $TOKEN" -X GET https://${eBMC}/redfish/v1/Systems/system
   | grep TotalSystemMemoryGiB
   "TotalSystemMemoryGiB": 8192
#Power state
   # curl -s -k -H "X-Auth-Token: $TOKEN" -X GET https://${eBMC}/redfish/v1/Systems/system
   | grep PowerState
     "PowerState": "On",
```

It is also possible to run operations on the server by using the `POST` method with the Redfish API interface. In Example 2-5, you can see the `curl` commands that can start or stop the server.

Example 2-5   POST command examples for Redfish

```
#Power on server
   # curl -k -H "X-Auth-Token: $TOKEN" -X POST https://${eBMC}/redfish/v1/Systems/
   system/Actions/Reset -d '{"ResetType":"On"}'
#Power off server
   # curl -k -H "X-Auth-Token: $TOKEN" -X POST https://${eBMC}/redfish/v1/Systems/
   system/Actions/Reset -d '{"ResetType":"ForceOff"}'
```

For more information about Redfish, see Managing the system by using DMTF Redfish APIs. For more information about how to work with Redfish in IBM Power servers, see Managing Power Systems servers by using DMTF Redfish APIs.

2.5.3 Managing the system by using the IPMI

The Power E1050 server can also be managed by using the IPMI, but the IPMI is disabled by default on your server. Inherent security vulnerabilities are associated with the IPMI, so consider using Redfish APIs or the GUI to manage your system.

If you want to use IPMI, you first must enable it. To do so, select **Security and access** → **Policies**. Then, enable the policy **Network IPMI (out-of-band IPMI)**.

A list of common IPMI commands can be found at Common IPMI commands.

3

Available features and options

In this chapter, you learn about the various features and options for the IBM Power E1050 server, including the processor and memory features, and the Peripheral Component Interconnect Express (PCIe) adapter for storage, network, and other functions. This chapter also describes the server racks and their options.

3.1 Processor module features

This section describes all processor-related Feature Codes for the Power10 processor-based Enterprise Midrange Power E1050 server.

The Power E1050 supports 24 - 96 processor cores. A minimum of two and a maximum of four processor modules are required for each system. The modules can be added to a system later through a Miscellaneous Equipment Specification (MES) order, but they require scheduled downtime to install. All processor modules in one server must be at the same gigahertz frequency (that is, they must be the same processor module feature number).

Table 3-1 lists the processor card Feature Codes that are available at initial order for Power E1050 servers: the related module type, the number of function cores, the typical frequency range in maximum performance mode (MPM), and the socket options.

Table 3-1   Processor card features

E1050 Feature Code	Processor SMT8 cores	Maximum system cores	Typical frequency range	Socket options
#EPEU	12	48	3.35 - 4.00 GHz	2, 3, or 4
#EPEV	18	72	3.20 - 4.00 GHz	2, 3, or 4
#EPGW	24	96	2.95 - 3.90 GHz	2, 3, or 4
#EHC8[a]	24	96	2.95 - 3.90 GHz	2, 3, or 4

a. Healthcare solution Edition (North America (NA) only)

> **Note:** IBM intends to support SAP HANA on the Power E1050 after initial general availability (GA). SAP HANA on Power E1050 will be certified with four sockets (4S), 24 cores, and 16 TB of memory. All processor core options, including 12-core and 18-core, also will be offered after certification.

Permanent Capacity on Demand processor core activations

Permanent Capacity on Demand (CoD) processor core activations are required for the first processor module in the configuration and are optional for the second and fourth modules. Specifically:

► Two, three, or four 12-core typical 3.35 – 4.0 GHz (max) processor modules (#EPEU) require 12 processor core activations (#EPUR) at a minimum.

► Two, three, or four 18-core typical 3.20 – 4.0 GHz (max) processor modules (#EPEV) require 18 processor core activations (#EPUS) at a minimum.

► Two, three, or four 24-core, typical 2.95 – 3.90 GHz (max) processor modules (#EPGW) require 24 processor core activations (#EPYT) at a minimum.

For Linux, the following activation features are available:

► One-core Processor Activation for #EPEU Linux only (#EPUN)

► One-core Processor Activation for #EPEV Linux only (#EPUP)

► One-core Processor Activation for #EPEW Linux only (#EPUM)

For the Healthcare solution edition, the Processor Activation (24) for Healthcare Solution #EHC8 (#EHCA) activation feature is available.

Temporary CoD capabilities

Temporary CoD is used for processor cores that are not permanently activated.

A Hardware Management Console (HMC) is required for Temporary CoD.

Cloud solution and dynamic capacity

The IBM Power Private Cloud with Shared Utility Capacity solution is an alternative to the static processor activation use model. It can be used to enable cloud agility and cost optimization with pay-for-use pricing Power E1050 server support.

Both Elastic and Shared Utility Capacity options are available on all Power E1050 servers through the Virtual Capacity (4586-COD) machine type and model (MTM). For more information, see the IBM Entitled Systems Support (IBM ESS) website.

With Elastic Capacity on Power E1050 servers, you can deploy pay-for-use consumption of processors by the day across a collection of Power E1050 and Power E950 servers.

Shared Utility Capacity on Power E1050 servers provides enhanced multisystem resource sharing and by-the-minute tracking and consumption of compute resources across a collection of systems within a Power Enterprise Pools 2.0 (PEP2). Shared Utility Capacity on Power E1050 servers has the flexibility to tailor initial system configurations with the right mix of purchased and pay-for-use consumption of processor.

Table 3-2 lists the Feature Codes for processor activation with PEP2.

Table 3-2   Base Processor activations in a PEP2

Feature Code	Description
#EPRS	1 core Base Processor Activation (Pools 2.0) for EPEU
#EPRT	1 core Base Processor Activation (Pools 2.0) for #EPEV
#EPRV	1 core Base Processor Activation (Pools 2.0) for #EPGW
#EPRW	1 core Base Processor Activation (Pools 2.0) for #EPEU Linux only
#EPRX	1 core Base Processor Activation (Pools 2.0) for #EPEV Linux only
#EPRZ	1 core Base Processor Activation (Pools 2.0) for #EPGW Linux only

An initially ordered configuration based on static processor activations can be converted to the Shared Utility Capacity use model.

Table 3-3 provides a compilation of all processor-related Feature Codes for processor conversion for the Power E1050 servers.

Table 3-3   Base Processor activation conversion options in a PEP2

Feature Code	Description
#EPSA	1 core Base Processor Activation (Pools 2.0) for #EPEU (conv. from Static)
#EPSB	1 core Base Processor Activation (Pools 2.0) for #EPEV (conv. from Static)
#EPSD	1 core Base Processor Activation (Pools 2.0) for #EPGW (conv. from Static)
#EPSE	1 core Base Processor Activation (Pools 2.0) for #EPEU (conv. from Static) Linux only
#EPSF	1 core Base Processor Activation (Pools 2.0) for #EPEV (conv. from Static) Linux only

Feature Code	Description
#EPSH	1 core Base Processor Activation (Pools 2.0) for #EPGW (conv.from Static) Linux only
#ERQ0	1 core Base Processor Activation (Pools 2.0) for EPEU (from prev.)
#ERQ1	1 core Base Processor Activation (Pools 2.0) for #EPEV (from prev.)
#ERQ2	1 core Base Processor Activation (Pools 2.0) for #EPGW (from prev.)
#ERQ4	1 core Base Processor Activation (Pools 2.0) for #EPEU (from prev.) Linux only
#ERQ5	1 core Base Processor Activation (Pools 2.0) for #EPEV (from prev.) Linux only
#ERQ6	1 core Base Processor Activation (Pools 2.0) for #EPGW (from prev.) Linux only

3.2 Memory features

All available memory Feature Codes for the Power E1050 server are listed in Table 3-4. Each memory Feature Code relates to two Differential Dual Inline Memory Module (DDIMM) cards of identical specification.

Table 3-4   Power E1050 memory features

Feature Code	Capacity	Packaging	Dynamic RAM (DRAM) density	DRAM data rate
#EM75	64 GB	Two 32-GB DDIMMs	16 Gbit	3200 MHz
#EM76	128 GB	Two 64-GB DDIMMs	16 Gbit	3200 MHz
#EM77	256 GB	Two 128-GB DDIMMs	16 Gbit	2933 MHz
#EM7J[a]	512 GB	Two 256-GB DDIMMs	16 Gbit	2933 MHz

a. Available in 4Q2022

Permanent CoD memory activations
Permanent CoD memory activations are required for at least 50% of the physically installed memory or 256 GB of activations, whichever is larger. Use 1 GB activation (#EMCP) and 100 GB activation (#EMCQ) features to order permanent memory activations.

Temporary CoD memory activations
Temporary CoD memory activations are available for memory capacity that is not permanently activated:

► Delivery through Virtual Capacity MTM (4586-COD) by using the IBM ESS process.

► An HMC is required for Temporary CoD.

All the memory activation features are listed in Table 3-5, which also include the memory conversion Feature Codes.

Table 3-5   Memory activation features

Feature Code	Description
EMCP	1 GB Memory Activations for MRX
EMCQ	100 GB Memory Activations for MRX
EMCE	512 of #EMCP 1 GB Memory Activations for MRX
EHCB	512 GB Base Memory activation for MRX #EHC8
EPSJ	1 GB Base Memory activation (Pools 2.0)
EPSP	100 GB Base Memory activation (Pools 2.0)
EPSW	256 GB Base Memory Activation for Pools 2.0
EPSU	256 GB Base Memory Activation for Pools 2.0 - Linux only
EPSQ	512 GB Base Memory Activation for Pools 2.0 (from Static)
EPSR	1 GB Base Memory Activation for Pools 2.0 (from Static)
EPSS	100 GB Base Memory Activation for Pools 2.0 (from Static)
ERQM	256 GB Base Memory Activation (Pools 2.0) (from prev.)
ERQN	256 GB Base Linux only Memory Activation (Pools 2.0) (from prev.)
EPST	512 GB Base Memory Activation for Pools 2.0 (from Linux)

► Memory is ordered in a quantity of eight of the same memory feature.

► The minimum memory that is supported per two Power10 processors that are installed is 256 GB.

► The minimum memory that is supported per three Power10 processors that are installed is 384 GB.

► The minimum memory that is supported per four Power10 processors that are installed is 512 GB.

The E1050 9043-MXR server supports, as an option, Active Memory Expansion (AME) with Feature Code #EMBM, and also as an option Active Memory Mirroring (AMM) for the Hypervisor with Feature Code #EM81.

3.3 Power supply features

The Power E1050 server has four AC to DC 2300 W Titanium classed power supply units (PSUs). The PSUs are concurrently maintainable.

> **Note:** A Titanium-certified (best in class) power supply has an efficiency of 90% - 95% on load.
>
> Every 1% efficiency loss at the PSU requires more than 1% more power to accomplish the same output. Efficient power supplies help data centers to achieve their sustainability goals.

The Power E1050 power supplies have 2+2 redundancy. The failed PSU is hot-pluggable.

► When one PSU is defective during run time, the system continues to run with full performance.

► When two PSUs are defective during run time, the system continues to run, but might throttle as needed.

► During boot, the system requires three PSUs to be installed, but will power on with only two good PSUs.

3.4 Peripheral Component Interconnect adapters

This section covers the various types and functions of the PCIe adapters that are supported by the Power E1050 (9043-MRX). This list is based on the PCIe adapters that are available on the GA date of these systems, but it is subject to change as more PCIe adapters are tested and certified, or listed adapters are no longer available. The latest list of supported adapters is available at the Announcement Letter or Sales Manual page, which can be found at the IBM Offering Information website.

The following sections describe the supported adapters and provide tables of orderable and supported feature numbers. The tables indicate OS support (AIX and Linux) for each of the adapters.

Note: The maximum number of adapters in each case might require the server to have an external PCIe expansion drawer.

The Order type table column in the following subsections is defined as follows:

Initial   Denotes the orderability of a feature *only* with the purchase of a new system.

MES   Denotes the orderability of a feature *only* as part of an MES upgrade purchase for an existing system.

Both   Denotes the orderability of a feature as part of *both* new and MES upgrade purchases.

Supported   Denotes that feature is *not* orderable with a system, but is *supported*; that is, the feature can be migrated from existing systems, but cannot be ordered new.

3.4.1 Local area network adapters

To connect the IBM Power E1050 server models to a local area network (LAN), you can use the LAN adapters that are supported in the PCIe slots of the system. Several connection speeds and physical connections are supported.

Table 3-6 lists the LAN adapters that are supported within the Power E1050 servers.

Table 3-6   LAN adapters for the Power E1050 server

Feature Code	Description	OS support	Order type
#5899	PCIe2 LP 4-port 1 GbE Adapter	AIX and Linux	Supported
#EC21	PCIe3 LP 2-port 25/10 Gb NIC&ROCE SR/Cu Adapter	AIX and Linux	Supported
#EC2U	PCIe3 2-Port 25/10 Gb NIC&ROCE SR/Cu Adapter[a]	AIX and Linux[b]	Both
#EC66	PCIe4 LP 2-port 100 Gb ROCE EN adapter[c]	AIX and Linux[b]	Both
#EN0S	PCIe2 4-port (10 Gb+1 GbE) SR+RJ45 Adapter	AIX and Linux	Supported
#EN0U	PCIe2 4-port (10 Gb+1 GbE) Copper SFP+RJ45 Adapter	AIX and Linux	Supported
#EN0W	PCIe2 2-port 10/1 GbE BaseT RJ45 Adapter	AIX and Linux	Both

a. Requires one or two appropriate transceivers to provide 10 Gbps SFP+ (#EB46), 25 Gbps SFP28 (#EB47), or 1 Gbps RJ45 (#EB48) connectivity as required.
b. Linux support requires Red Hat Enterprise Linux (RHEL) 8.4 or later, Red Hat Enterprise Linux for SAP 8.4 or later, SUSE Linux Enterprise Server 15 Service Pack 3 or later, SUSE Linux Enterprise Server for SAP with SUSE Linux Enterprise Server 15 Service Pack 3 or later, or Red Hat OpenShift Container Platform 4.9 or later. All require Mellanox OFED 5.5 drivers or later.
c. To deliver the full performance of both ports, each 100 Gbps Ethernet adapter must be connected to a PCIe slot with 16 lanes (x16) of PCIe Gen4 connectivity.

3.4.2 Fibre Channel adapters

The Power10 processor-based Enterprise Midrange servers support connection to devices that use Fibre Channel (FC) connectivity, either directly or through a storage area network (SAN). There is a range of PCIe connected FC adapters that are available in both low-profile and full-height form factors.

All supported FC adapters have LC connections. If you are attaching a switch or a device with an SC type fiber connector, then an LC-SC 50-Micron Fiber Converter Cable (#2456) or an LC-SC 62.5-Micron Fiber Converter Cable (#2459) is required.

Table 3-7 lists the FC adapters that are supported within the Power E1050 servers.

Table 3-7   Fibre Channel adapters for the Power E1050 servers

Feature Code	Description	OS support	Order type
#EN1A	PCIe3 32 Gb 2-port Fibre Channel Adapter	AIX and Linux	Both
#EN1C	PCIe3 16 Gb 4-port Fibre Channel Adapter	AIX and Linux	Both
#EN1E	PCIe3 16 Gb 4-port Fibre Channel Adapter	AIX and Linux	Both
#EN1G	PCIe3 2-port 16 Gb Fibre Channel Adapter	AIX and Linux	Both
#EN1J	PCIe4 32 Gb 2-port Optical Fibre Channel Adapter	AIX and Linux	Both
#EN2A	PCIe3 16 Gb 2-port Fibre Channel Adapter	AIX and Linux	Both

3.4.3 Serial-attached SCSI adapters

The internal storage in the Power10 processor-based Enterprise Midrange servers is all based on Non-volatile Memory Express (NVMe) devices that are connected over PCIe directly. You also can connect further storage expansion drawers to the system by using serial-attached SCSI (SAS) connections.

Table 3-8 lists the SAS adapters that are supported within the Power E1050 servers and with the PCIe expansion drawer (EMX0) connected.

Table 3-8   SAS adapters for the Power E1050 servers

Feature Code	Description	OS support	Order type
#EJ0J	PCIe3 RAID SAS Adapter Quad-Port 6 Gb x8	AIX and Linux	Both
#EJ14	PCIe3 12 GB Cache RAID PLUS SAS Adapter Quad-port 6 Gb x8	AIX and Linux	Both

3.4.4 Universal Serial Bus adapters

Universal Serial Bus (USB) adapters are available to support the connection of devices, such as external DVD drives to the Power10 processor-based Enterprise Midrange servers.

Table 3-9 list the USB adapter that is supported within the Power E1050 with the PCIe expansion drawer (EMX0) connected to it.

Table 3-9   USB adapter for the Power E1050 server

Feature Code	Description	OS support	Order type
#EC6K	PCIe2 LP 2-port USB 3.0 Adapter	AIX and Linux	Both

3.4.5 Cryptographic co-processor adapters

Two different cryptographic co-processors and accelerators are supported by the Power10 processor-based Enterprise Midrange servers, both of which are full-height adapters. These adapters work with the IBM Common Cryptographic Architecture (CCA) to deliver acceleration of cryptographic workloads. For more information about the cryptographic co-processors, the associated software that is available, and the CCA that is available, see the IBM Systems cryptographic HSMs website.

PCIe Gen3 4767 Cryptographic Coprocessor

A secure-key adapter provides both cryptographic co-processor and cryptographic accelerator functions in a single PCIe card. The adapter is suited to applications requiring high-speed, security-sensitive Rivest-Shamir-Adleman (RSA) acceleration, cryptographic operations for data encryption and digital signing, secure management, and usage of cryptographic keys or custom cryptographic applications. The adapter provides secure storage of cryptographic keys in a tamper-resistant hardware security module that meets FIPS 140-2 level 4 security requirements. The adapter is a PCIe Gen 3 x4 full-height short card. The adapter runs in dedicated mode only (no PowerVM virtualization).

This adapter is available only as the full-height form factor, and it is available in two variations with two different Feature Codes:

► #EJ32 does not include a Blind-Swap Cassette (BSC), and can be installed only within the chassis of a Power E1050 server.

► #EJ33 includes a BSC housing, and can be installed only in a PCIe Gen3 I/O expansion drawer enclosure.

PCIe Gen3 4769 Cryptographic Coprocessor

The 4769 PCIe Cryptographic Coprocessor has a PCIe local bus compatible interface. The co-processor holds a security-enabled subsystem module and batteries for backup power. The hardened encapsulated subsystem contains redundant IBM PowerPC 476 processors; custom symmetric key and hashing engines to perform Advanced Encryption Standard (AES), Data Encryption Standard (DES), Triple DES (TDES), SHA-1 and SHA- 2, MD5 and Hash-Based Message Authentication Code (HMAC); and public key cryptographic algorithm support for RSA and Elliptic Curve Cryptography. Other hardware support includes a secure real-time clock, hardware random number generator, and a prime number generator. It also contains a separate service processor that is used to manage self-test and firmware updates. The secure module is protected by a tamper responding design that protects against many types of attacks against the system.

The 4769 PCIe Cryptographic Coprocessor includes acceleration for AES, DES, TDES, HMAC, Cipher-based Message Authentication Code (CMAC), MD5, multiple Secure Hash Algorithm (SHA) hashing methods, modular-exponentiation hardware, such as RSA and error correction code (ECC), and full-duplex direct memory access (DMA) communications.

A security-enabled code-loading arrangement allows control program and application program loading and refreshes after co-processor installation in your server. IBM offers an embedded subsystem control program and a cryptographic application programming interface (API) that implements the IBM Common Cryptographic Architecture.

The 4769 PCIe Cryptographic Coprocessor is verified by NIST at FIPS 140-2 Level 4, which is the highest level of certification currently achievable for commercial cryptographic devices.

Table 3-10 summarizes the cryptographic co-processor and accelerator adapters that are supported in the Power10 processor-based Enterprise Midrange servers.

Table 3-10   Crypto adapter features for the Power E 1050 servers

Feature Code	Description	OS support	Order type
#EJ35[a]	PCIe3 Crypto Coprocessor no BSC 4769 (E1050 chassis only)	AIX and Linux Direct only	Both
#EJ37[a]	PCIe3 Crypto Coprocessor BSC-Gen3 4769 (PCIe expansion drawer only)	AIX and Linux Direct only	Both

a. Feature Codes #EJ35 and #EJ37 are both Feature Codes representing the same physical card. #EJ35 indicates no BSC. #EJ37 indicates a Gen 3 BSC.

3.5 Internal storage

The Power E1050 server includes an NVMe U.2 disk backplane that provides 10 internal NVMe bays. The backplane is always present in the server. A U.3-U.2 configurable disk drive can run as NVMe U.2 mode only. The internal storage subsystem does not support the SAS protocol, so SAS disk drive options are available and supported only by adding one or more external expansion drawers to a Power E1050 configuration that supports up to 64 EXP24SX SFF Drawers (#ESLS).

Every Power E1050 configuration supports single-port NVMe mode only. Dual-port NVMe mode is not supported. Only the Power E1050 4-sockets configuration supports up to 10 NVMe disk drives. Configurations for three and two sockets support up to six NVMe disk drives. NVMe disk drives that are used as mirrors for operating system (OS) and dual Virtual I/O Server (VIOS) redundancy must be plugged according to the rules to ensure as many as possible separate hardware paths for the dual and mirrored pairs.

Table 3-11 shows the NVMe location codes inside the server.

Table 3-11   NVMe location codes

NVMe#	NVMe0	NVMe1	NVMe2	NVMe3	NVMe4	NVMe5	NVMe6	NVMe7	NVme8	NVMe9
Location codes	P1-C0	P1-C1	P1-C2	P1-C3	P1-C4	P1-C5	P1-C6	P1-C7	P1-C8	P1-C9

All NVMe disk drives are driven directly from the system backplane, so there is no need to have a PCIe card or cables that are dedicated to this purpose. The 7-mm NVMe disk drives from the IBM Power E950 are also supported on the Power E1050, but you must order the NVMe carrier conversion kit (#EC7X) to hold these drives.

3.5.1 Recommended NVMe plug rules

The following NVMe plug rules are recommended to provide the server with the most redundancy in hardware for OS mirror. To verify the sequence of the recommended disk drives positions according to their quantity and the Power E1050 processor configuration (2-, 3- or 4-socket), see Table 3-12.

Table 3-12   Recommended NVMe plug rules

Number of NVMe drives	2-socket configuration	3-socket configuration	4-socket configuration
2	NVMe3 and NVMe4, or NVMe8 and NVMe9	NVMe3 and NVMe4, or NVMe8 and NVMe9	NVMe3 and NVMe4, or NVMe8 and NVMe9
4	NVMe3, NVMe4, NVMe8, and NVMe9	NVMe3, NVMe4, NVMe8, and NVMe9	NVMe3, NVMe4, NVMe8, and NVMe9

Number of NVMe drives	2-socket configuration	3-socket configuration	4-socket configuration
6	NVMe3, NVMe4, NVMe8, NVMe9, NVMe2, and NVMe7	NVMe3, NVMe4, NVMe8, NVMe9, NVMe2, and NVMe7	NVMe3, NVMe4, NVMe8, NVMe9, NVMe2, and NVMe7
More than 6	Not supported	Not supported	► First group: NVMe3, NVMe4, NVMe8, and NVMe9 ► Second group: NVMe2 and NVMe7 ► Third group: NVMe5 and NVMe6 ► Forth group: NVMe0 and NVMe1

If you have an odd number of NVMe disk drives, the disks must be installed by following the same order in Table 3-12 on page 92. As example, if the server configuration provides three NVMe disk drives, the suggested order is NVMe3, NVMe4, and NVMe8.

Figure 3-1 shows the NVMe bays according to th NVMe location codes.
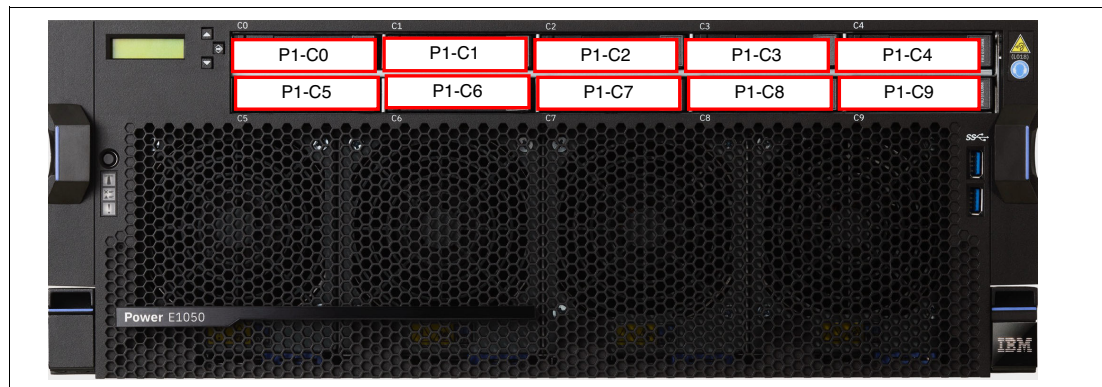
Figure 3-1   Power E1050 NVMe bays

3.5.2 NVMe options

Table 3-13 shows the Feature Codes that are related to NVMe disk drives that are supported on Power E1050.

Table 3-13   NVMe-supported Feature Codes

Feature Code	Description	Min	Max	OS support
#ECSJ	Mainstream 800 GB solid-state drive (SSD) NVMe U.2 module	0	10	AIX and Linux are supported.
#EC5K	Mainstream 1.6 TB SSD NVMe U.2 module	0	10	AIX and Linux are supported.

Feature Code	Description	Min	Max	OS support
#EC5L	Mainstream 3.2 TB SSD NVMe U.2 module	0	10	AIX and Linux are supported.
#EC5V	Enterprise 6.4 TB SSD PCIe4 NVMe U.2 module for AIX/Linux	0	10	AIX and Linux are supported.
#EC5X	Mainstream 800 GB SSD PCIe3 NVMe U.2 module for AIX/Linux	0	10	AIX and Linux are supported.
#EC7Q	800 GB Mainstream NVMe U.2 SSD 4k for AIX/Linux	0	10	AIX and Linux are supported.
#EC7T	800 GB Mainstream NVMe U.2 SSD 4k for AIX/Linux	0	10	AIX and Linux are supported.
#ES1E	Enterprise 1.6 TB SSD PCIe4 NVMe U.2 module for AIX/Linux	0	10	AIX and Linux are supported.
#ES1G	Enterprise 3.2 TB SSD PCIe4 NVMe U.2 module for AIX/Linux	0	10	AIX and Linux are supported.
#ES3B	Enterprise 1.6 TB SSD PCIe4 NVMe U.2 module for AIX/Linux	0	10	AIX and Linux are supported.
#ES3D	Enterprise 3.2 TB SSD PCIe4 NVMe U.2 module for AIX/Linux	0	10	AIX and Linux are supported.
#ES3F	Enterprise 6.4 TB SSD PCIe4 NVMe U.2 module for AIX/Linux	0	10	AIX and Linux are supported.

3.5.3 RAID support

The Power E1050 internal NVMe disk drives' volume and file systems can be protected in an OS-controlled RAID 0 or RAID 1 configuration. A hardware RAID option is not supported on the Power E1050 by using the internal NVMe drives.

3.6 External SAS ports

There are no external SAS ports on the Power E1050 server.

3.7 Media drawers

The IBM System Storage 7226 Model 1U3 Multi-Media Enclosure can accommodate up to two LTO tape drives, two RDX removable disk drive docking stations, or up to four DVD-RAM drives in a 1U rack-mountable dual bay enclosure. Up to two drives (or four DVD-RAM) can be installed in any combination in the 7226 Model 1U3 enclosure. The 7226 Model 1U3 offers SAS, USB, and FC electronic interface drive options for attachment to the Power E1050 server. Up to six of these multimedia drawers can be attached.

Table 3-14 on page 95 shows the drive features for the 7226 Model 1U3.

Table 3-14   Available drive features for the 7226 Model 1U3

Feature Code	Description
#1420	DVD-RAM SAS Sled with 1 DVD-RAM drive
#1422	DVD-RAM Slim SAS Drive (1 per #1420 Sled)
#5762	DVD-RAM USB Sled with 1 DVD-RAM Drive
#5757	DVD RAM Slim USB Drive (1 per #5762 Sled
#8441	LTO Ultrium 7 Half High SAS Tape Drive
#8541	LTO Ultrium 8 Half High SAS Tape Drive
#8641	LTO Ultrium 9 Half High SAS Tape Drive
#8446	LTO Ultrium 7 Half High Fibre Tape Drive
#8546	LTO Ultrium 8 Half High Fibre Tape Drive
#8646	LTO Ultrium 9 Half High Fibre Tape Drive
#EU03	RDX 3.0 Removable Disk Docking Station
#1107	500 GB Removable RDX Drive
#EU01	1.0 TB Removable RDX Drive
#EU2T	2.0 TB Removable RDX Drive

Note: Existing 32 GB (#EU08) and 1.5 TB (#EU15) removable disk drive cartridges are still supported on the RDX docking stations.

PCIe3 RAID SAS Tape/DVD Adapter Quad-port 6 Gb x8 (#EJ10) supports Tape/DVD with the following cable options:

► SAS AE1 Cable 4 m - HD Narrow 6 Gb Adapter to Enclosure (#ECBY)
► SAS YE1 Cable 3 m - HD Narrow 6 Gb Adapter to Enclosure (#ECBZ)

Note: Any of the existing 7216-1U2 and 7214-1U2 multimedia drawers also are supported.

3.7.1 External DVD drives

The Stand-alone USB DVD drive (#EUA5) is an optional, stand-alone external USB-DVD device. It requires high current at 5 V and must use the front USB 3.0 port on the Power E1050 server. This device comes with a USB cable. The cable provides the data path and power to this drive. A maximum of one external DVD drive can be attached. The USB DVD drive is supported by AIX, and it can be assigned to the VIOS.

3.7.2 RDX removable disk drives

The RDX USB External Docking Station (#EUA4) accommodates RDX removable disk cartridges of any capacity. The disk is in a protective rugged cartridge enclosure that plugs into the docking station. The docking station holds one removable rugged disk drive/cartridge at a time. The rugged removable disk cartridge and docking station can be used like a tape drive. This feature can be an excellent entry system save/restore option and a good alternative to DAT72, DAT160, 8 mm, and VXA-2 and VXA-320 tapes.

The RDX USB External Docking Station attaches to the Power E1050 server through an included USB cable that carries data and control information. It is not powered by the USB port on the Power server or a Power server USB adapter, but has a separate electrical power cord. Physically, the docking station is a stand-alone enclosure with the dimensions 2.0 x 7.0 x 4.25 inches, and the station can sit on a shelf or on top of equipment in a rack. Up to six of these devices can be attached. The USB DVD drive is supported by AIX and can be assigned to the VIOS.

Various disk drives are available, as shown in Table 3-15.

Table 3-15   RDX removable disk drives

Feature Code	Description
#1107	USB 500 GB Removable Disk Drive
#EU01	1 TB Removable Disk Drive Cartridge
#EU08[a]	RDX 320 GB Removable Disk Drive
#EU15[a]	1.5 TB Removable Disk Drive Cartridge
#EU2T	2 TB Removable Disk Drive Cartridge

a. Supported only. The feature can be migrated only from existing systems.

3.8 Disk and media features

NVMe SSDs are used for internal storage in the Power E1050 server. Up to six NVMe drives are supported on the Power E1050 with two or three dual-chip modules (DCMs), and up to 10 NVMe drives are supported on the Power E1050 with four DCMs installed. This NVMe backplane is a base component that is always present in the system.

Figure 3-2 shows the NVMe wiring to the DCMs.
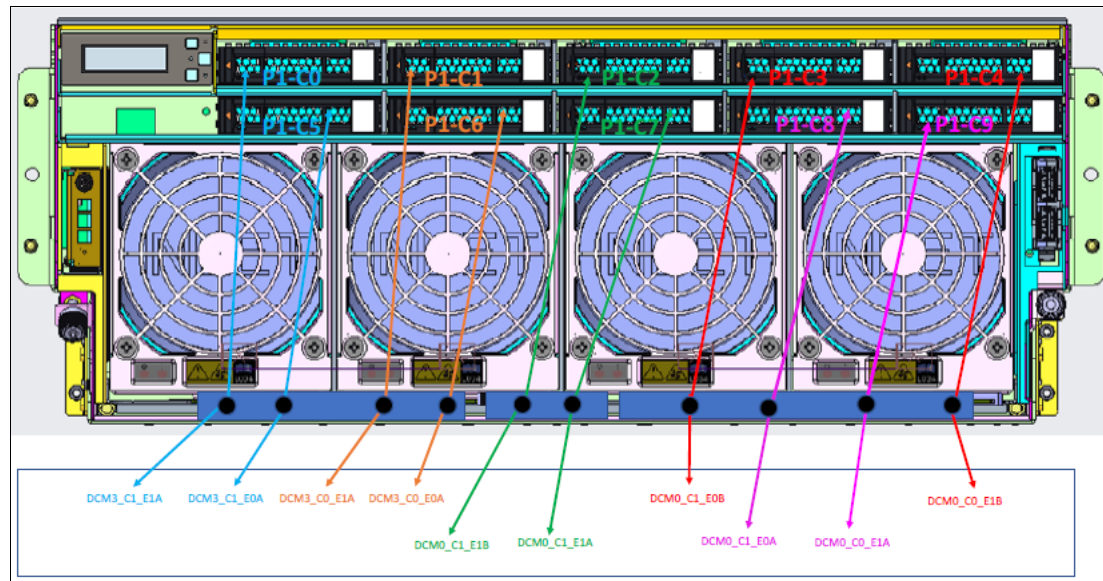
Figure 3-2   NVMe wiring to the DCMs

Table 3-16 shows the reference NVMe topology.

Table 3-16   Reference NVMe topology

DCM	NVMe#
DCM0	NVMe 2, 3, 4, 7, 8, and 9
DCM3 (only installed on four processors on the Power E1050)	NVMe 0, 1, 5, and 6

The NVMe plug rules that are shown in Table 3-17 are recommended for the Power E1050 to provide the most redundancy in hardware for OS mirroring.

Table 3-17   Recommended NVMe plug rules

# of NVMe drives	2 DCMs	3 DCMs	4 DCMs
2	NVMe 3 and 4, or NVMe 8 and 9	NVMe 3 and 4, or NVMe 8 and 9	NVMe 3 and 4, or NVMe 8 and 9
4	NVMe 3, 4, 8, and 9	NVMe 3, 4, 8, and 9	NVMe 3, 4, 8, and 9
6	NVMe 3, 4, 8, and 9 NVMe 2 and 7	NVMe 3, 4, 8, and 9 NVMe 2 and 7	NVMe 3, 4, 8, and 9 NVMe 2 and 7
More than 6	NVMe 3, 4, 8, and 9 NVMe 2 and 7 N/A (No NVMe 0, 1, 5, and 6 configuration is possible.)	NVMe 3, 4, 8, and 9 NVMe 2 and 7 N/A (No NVMe 0, 1, 5, and 6 configuration is possible.)	First NVMe 3, 4, 8, and 9 Next NVMe 2 and 7 Next NVMe 5 and 6 Finally, NVMe 0 and 1

For OS NVMe mirror pairs, only an OS-controlled RAID 0, 1 is supported. There is no support for hardware mirroring with the NVMe backplane.

Note: It is recommended (but not required) that the mirrored two NVMe drives of the pairs have the same capacity.

Here are the preferred NVMe pairs for mirroring:

► Mirrored OS: NVMe 3 and 4, or NVMe 8 and 9

► Mirrored dual-VIOS with four NVMe drives:

  – NVMe 3 for VIOS1, and NVMe 4 for VIOS2.

  – NVMe 9 mirrors NVMe 3, and NVMe 8 mirrors NVMe 4.

Note: In actual operation, there are many running partitions in the system, so the mirror pairs selections depend on the drives that are available or allocated to the partition.

Table 3-18 shows the internal storage option that is installed in the Power E1050 server.

Table 3-18   Internal storage option in the Power E1050 server

Feature Code	Description	Maximum
#EJ0Q[a]	Storage backplane with 10 NVMe U.2 drive slots	1

a. All drives are driven directly through Gen4 connectors from the system board (there is no PCIe card and no cable). Future Gen5 NVMe drives can be installed, but run with Gen4 speed.

Note: Power E1050 runs only in single-port NVMe mode. Dual-port NVMe mode is not supported.

Table 3-19 lists the available NVMe drive Feature Codes for the Power E1050 server.

Table 3-19   NVMe drive features in the Power E1050 server

Feature Code	Description	Min	Max	OS support
#EC5J	Mainstream 800 GB SSD NVMe U.2 module (7 mm)	0	10	AIX and Linux
#EC5K	Mainstream 1.6 TB SSD NVMe U.2 module (7 mm)	0	10	AIX and Linux
#EC5L	Mainstream 3.2 TB SSD NVMe U.2 module (7 mm)	0	10	AIX and Linux
#EC5V	Enterprise 6.4 TB SSD PCIe4 NVMe U.2 module for AIX/Linux (15 mm)	0	10	AIX and Linux
#EC5X	Mainstream 800 GB SSD PCIe3 NVMe U.2 module for AIX/Linux (7 mm)	0	10	AIX and Linux
#EC7Q	800 GB Mainstream NVMe U.2 SSD 4k for AIX/Linux (7 mm)	0	10	AIX and Linux
#EC7T	800 GB Mainstream NVMe U.2 SSD 4k for AIX/Linux (15 mm)	0	10	AIX and Linux
#ES1E	Enterprise 1.6 TB SSD PCIe4 NVMe U.2 module for AIX/Linux (15 mm)	0	10	AIX and Linux
#ES1G	Enterprise 3.2 TB SSD PCIe4 NVMe U.2 module for AIX/Linux (15 mm)	0	10	AIX and Linux
#ES3B	Enterprise 1.6 TB SSD PCIe4 NVMe U.2 module for AIX/Linux (15 mm)	0	10	AIX and Linux
#ES3D	Enterprise 3.2 TB SSD PCIe4 NVMe U.2 module for AIX/Linux (15 mm)	0	10	AIX and Linux
#ES3F	Enterprise 6.4 TB SSD PCIe4 NVMe U.2 module for AIX/Linux (15 mm)	0	10	AIX and Linux

Attention: A minimum quantity of one SSD NVMe drive must be ordered with an AIX or Linux OS if SAN Boot (#0837) or Remote Load Source (#EHR2) is *not* ordered:

► If SAN Boot (#0837) is ordered, then an adapter (FC or FCoE) that supports FC protocols to attach the system to a SAN must be ordered or present on the system instead.

► If Remote Load Source (#EHR2) is ordered, then at least one HDD or SSD drive must be present in the EXP24SX (#ESLS) drawer (or existing EXP12SX (#ESLL).

There is no SAS backplane that is supported on the Power E1050 server. SAS drives can be placed in the IBM EXP24SX SAS Storage Enclosure that is connected to the system only by using SAS adapters.

Important: Each NVMe slot has a 15-mm NVMe form factor. The carrier conversion KIT (#EC7X) feature can be used for converting 7-mm NVMe devices (this kit can be used with #EC5J, #EC5K, #EC5L, #EC5X, or #EC7Q drives) into 15-mm devices. This feature can be ordered with no NVMe drives in the carrier.

The default priority of slots should be 15-mm NVMe drives options as a higher priority than #EC7X.

If #EC7X is ordered without drives, customers must self-install their existing 7-mm NVMe devices into the conversion kit.

3.9 External IO subsystems

If more PCIe slots beyond the system node slots are required, the Power E1050 server supports adding I/O expansion drawers.

If you need more disks than are available with the internal disk bays, you can attach external disk subsystems that can be attached to the Power E1050 server, such as:

► EXP24SX SAS Storage Enclosure
► IBM System Storage

Note: The existing EXP12SX SAS Storage Enclosure (#ESLL) is still supported. Earlier storage enclosures like the EXP12S SAS Drawer (#5886) and EXP24 SCSI Disk Drawer (#5786) are *not* supported on the Power E1050.

3.9.1 PCIe Gen3 I/O expansion drawer

This 19-inch, 4U (4 EIA) enclosure provides PCIe Gen3 slots outside of the system unit. It has two module bays. One 6-slot fanout module (#EMXH) can be placed in each module bay. Two 6-slot modules provide a total of 12 PCIe Gen3 slots. Each fanout module is connected to a PCIe3 Optical Cable adapter that is in the system unit over an Active Optical Cable (AOC) pair or a CXP copper cable pair.

The PCIe Gen3 I/O Expansion Drawer has two redundant, hot-plug power supplies. Each power supply has its own separately ordered power cord. The two power cords plug in to a power supply conduit that connects to the power supply. The single-phase AC power supply is rated at 1030 W and can use 100 - 120 V or 200 - 240 V. If using 100 - 120 V, then the maximum is 950 W. It is a best practice that the power supply connects to a power distribution unit (PDU) in the rack. Power Systems PDUs are designed for a 200 - 240 V electrical source.

The drawer has fixed rails that can accommodate racks with depths 27.5" (69.9 cm) - 30.5" (77.5 cm).

Attention: #EMX0 has a cable management bracket at the rear of the drawer that swings up to provide service access to the PCIe adapters. 2U (2 EIA) of space is required to swing up the bracket. Thus, the drawer cannot be placed at the top 2U of a rack. There is a power cord access consideration with vertically mounted PDUs on the right side of the rack when viewed from the rear of the rack. The #EMX0 cable management bracket makes accessing some of the PDU outlets at the same rack height as the #EMX0 drawer more challenging. Using a horizontally mounted PDU or placing the PDU or #EMX0 at a different vertical location is recommended.

A BSC is used to house the full-height adapters that go into these slots. The BSC is the same BSC that is used with the previous generation server's 12X attached I/O drawers (#5802, #5803, #5877, and #5873). The drawer includes a full set of BSCs, even if the BSCs are empty.

Concurrent repair, and adding or removing PCIe adapters is done by HMC-guided menus or by OS support utilities.

Figure 3-3 shows a PCIe Gen3 I/O expansion drawer.

Figure 3-3   PCIe Gen3 I/O expansion drawer

Figure 3-4 shows the back view of the PCIe Gen3 I/O expansion drawer.
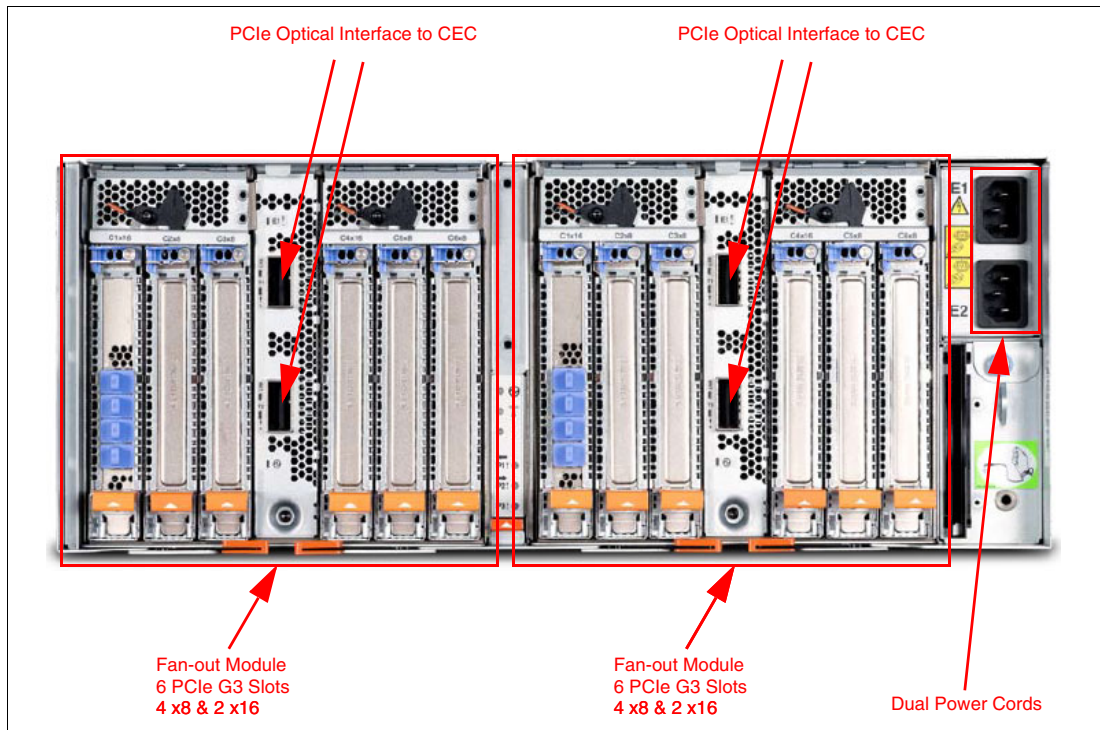
Figure 3-4   Rear view of the PCIe Gen3 I/O expansion drawer

I/O drawers and usable PCI slots

Figure 3-5 shows the rear view of the PCIe Gen3 I/O expansion drawer that is equipped with two PCIe3 6-slot fanout modules with the location codes for the PCIe adapter slots.
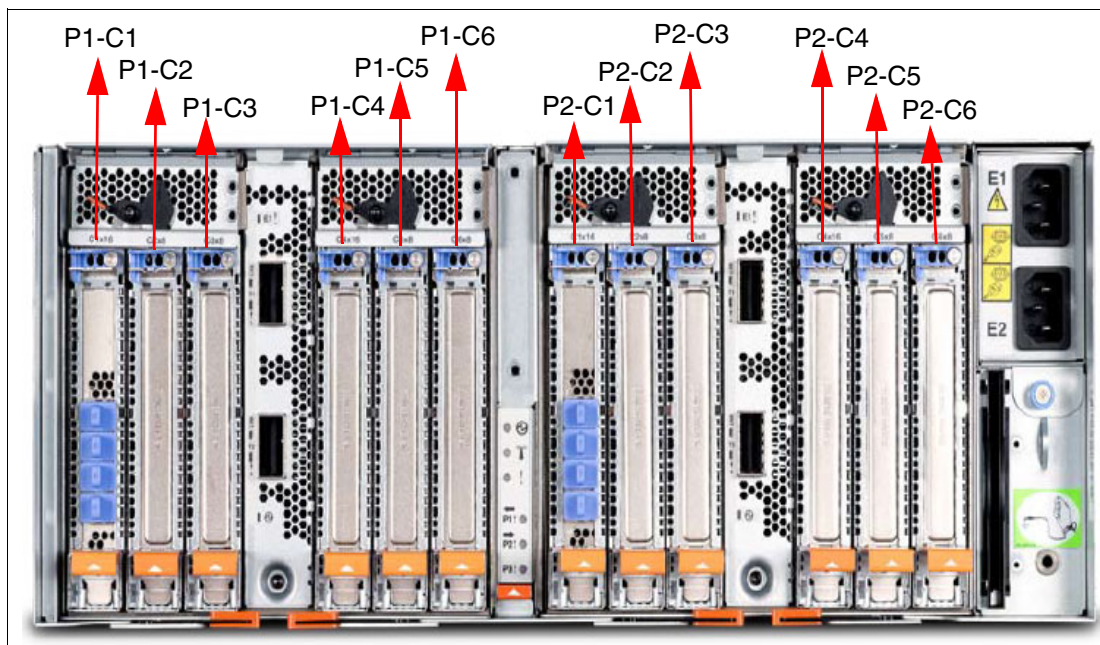
Figure 3-5   Rear view of a PCIe Gen3 I/O expansion drawer with PCIe slots location codes

Table 3-20 provides details about the PCI slots in the PCIe Gen3 I/O expansion drawer that is equipped with two PCIe3 6-slot fanout modules.

Table 3-20   PCIe slot locations for the PCIe Gen3 I/O expansion drawer with two fanout modules

Slot	Location code	Description
Slot 1	P1-C1	PCIe3, x16
Slot 2	P1-C2	PCIe3, x8
Slot 3	P1-C3	PCIe3, x8
Slot 4	P1-C4	PCIe3, x16
Slot 5	P1-C5	PCIe3, x8
Slot 6	P1-C6	PCIe3, x8
Slot 7	P2-C1	PCIe3, x16
Slot 8	P2-C2	PCIe3, x8
Slot 9	P2-C3	PCIe3, x8
Slot 10	P2-C4	PCIe3, x16
Slot 11	P2-C5	PCIe3, x8
Slot 12	P2-C6	PCIe3, x8

► All slots support full-length, full-height adapters or short (LP) adapters with a full-height tailstock in a single-wide, Gen3 BSC.

► Slots C1 and C4 in each PCIe3 6-slot fanout module are x16 PCIe3 buses, and slots C2, C3, C5, and C6 are x8 PCIe buses.

► All slots support enhanced error handling (EEH).

► All PCIe slots are hot-swappable and support concurrent maintenance.

Table 3-21 summarizes the maximum number of I/O drawers that are supported and the total number of PCI slots that are available.

Table 3-21   Maximum number of I/O drawers that are supported and total number of PCI slots

Server	Maximum number of I/O expansion drawers	Maximum number of I/O fanout modules	Maximum PCIe slots
Power E1050 (2-socket)	2	4	31
Power E1050 (3- and 4-socket)	4	8	51

PCIe x16 to CXP Converter Adapter (Gen4)

The PCIe x16 to CXP Converter Adapter provides two ports for the attachment of two expansion drawer cables. One adapter supports the attachment of one PCIe3 6-slot fanout module in an EMX0 PCIe Gen3 I/O expansion drawer by using a CXP 16X AOC Pair.

The cable adapter can be placed in any slot in the system node. However, if an I/O expansion drawer is present, the PCIe x16 to CXP Converter Adapter must be given the highest priority. The maximum number of adapters that is supported is eight per system.

Table 3-22 shows the plug sequence for the converter adapter that can be installed in the Power E1050 server.

Table 3-22   Plug sequence of converter adapter in the Power E1050

Feature Code	Slot priorities (Two processors) with 2 I/O drawers	Maximum number of adapters that are supported	Slot priorities (3 and 4 processors) with 4 I/O drawers	Maximum number of adapters that are supported
#EJ2A	C10, C7, C11, and C8	4	C10, C7, C11, C8, C5, C3, C4, and C2	8

PCIe Gen3 I/O expansion drawer optical cabling

I/O drawers are connected to the adapters in the system node through data transfer cables:

► 3M Active Optical Cable Pair for PCIe3 Expansion Drawer (#ECCX)

► 10M Active Optical Cable Pair for PCIe3 Expansion Drawer (#ECCY)

► 3M Copper CXP Cable Pair for PCIe3 Expansion Drawer (#ECCS)

Note: The 3.0-meter copper cable pair connects a PCIe3 fanout module (#EMXH) in the PCIe Gen3 I/O Expansion Drawer to a PCIe Optical Converter Adapter (#EJ2A) in the system unit. There are two identical copper cables in the cable pair, each with two CXP connectors.

The output of the adapter is a CXP interface that also can be used for this copper cable pair.

Although these cables are not redundant, the loss of one cable reduces the I/O bandwidth (that is, the number of lanes that are available to the I/O module) by 50%.

Cable lengths: Use the 3.0-m cables for intra-rack installations. Use the 10.0-m cables for inter-rack installations.

Limitation: You *cannot* mix copper and optical cables on the same PCIe Gen3 I/O drawer. Both fanout modules either both use copper cables or both use optical cables.

A minimum of one PCIe x16 to CXP Converter adapter for PCIe3 Expansion Drawer is required to connect to the PCIe3 6-slot fanout module in the I/O expansion drawer. The fanout module has two CXP ports. The top CXP port of the fanout module is cabled to the top CXP port of the PCIe x16 to CXP Converter adapter. The bottom CXP port of the fanout module is cabled to the bottom CXP port of the same PCIe x16 to CXP Converter adapter.

To set up the cabling correctly, complete the following steps:

1. Connect an optical cable or copper CXP cable to connector T1 on the PCIe x16 to CXP Converter adapter in your server.

2. Connect the other end of the optical cable or copper CXP cable to connector T1 on one of the PCIe3 6-slot fanout modules in your expansion drawer.

3. Connect another cable to connector T2 on the PCIe x16 to CXP Converter adapter in your server.

4. Connect the other end of the cable to connector T2 on the PCIe3 6-slot fanout module in your expansion drawer.

5. Repeat steps 1 on page 103 - 4 for the other PCIe3 6-slot fanout module in the expansion drawer, if required.

Drawer connections: Each fanout module in a PCIe3 Expansion Drawer can be connected only to a single PCIe x16 to CXP Converter adapter for PCIe3 Expansion Drawer.

Figure 3-6 shows the connector locations for the PCIe Gen3 I/O Expansion Drawer.
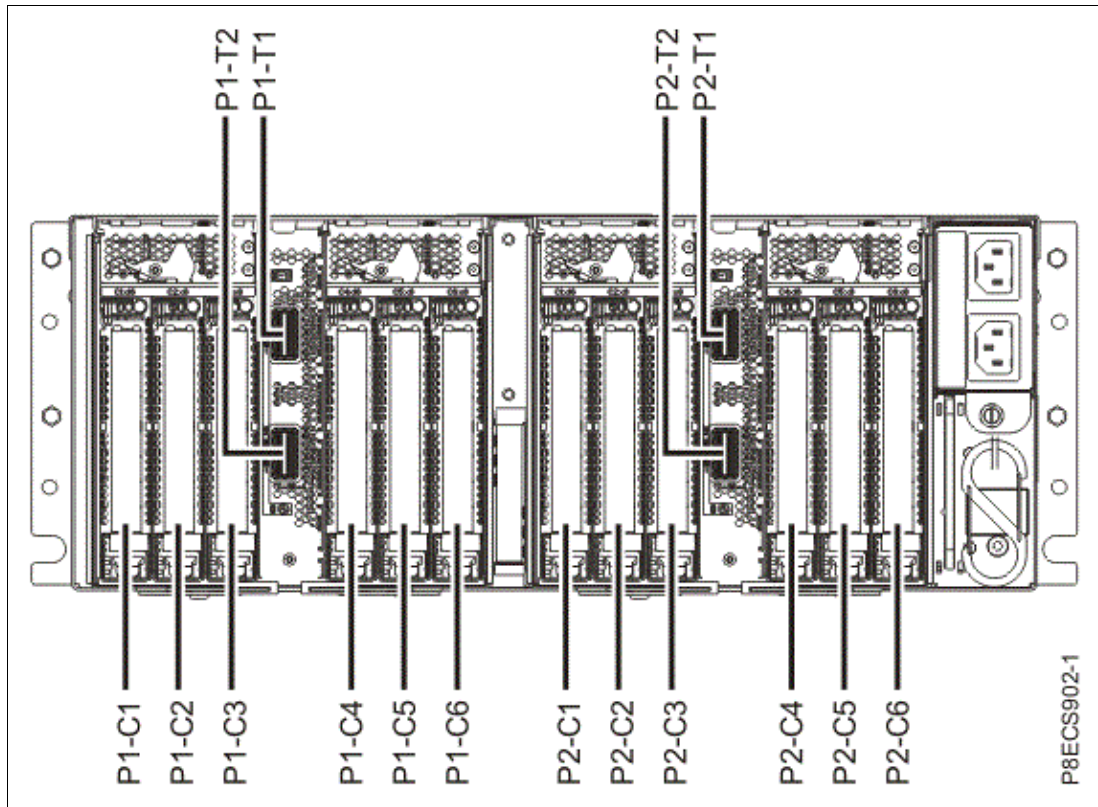
Figure 3-6   Connector locations for the PCIe Gen3 I/O expansion drawer

Figure 3-7 shows the typical optical cable connections.

Server   I/O Expansion Drawer

Optical cable adapter   Fanout module

T1   Optical cables   T1
T2   T2

Optical cable adapter   Fanout module

T1   Optical cables   T1
T2   T2
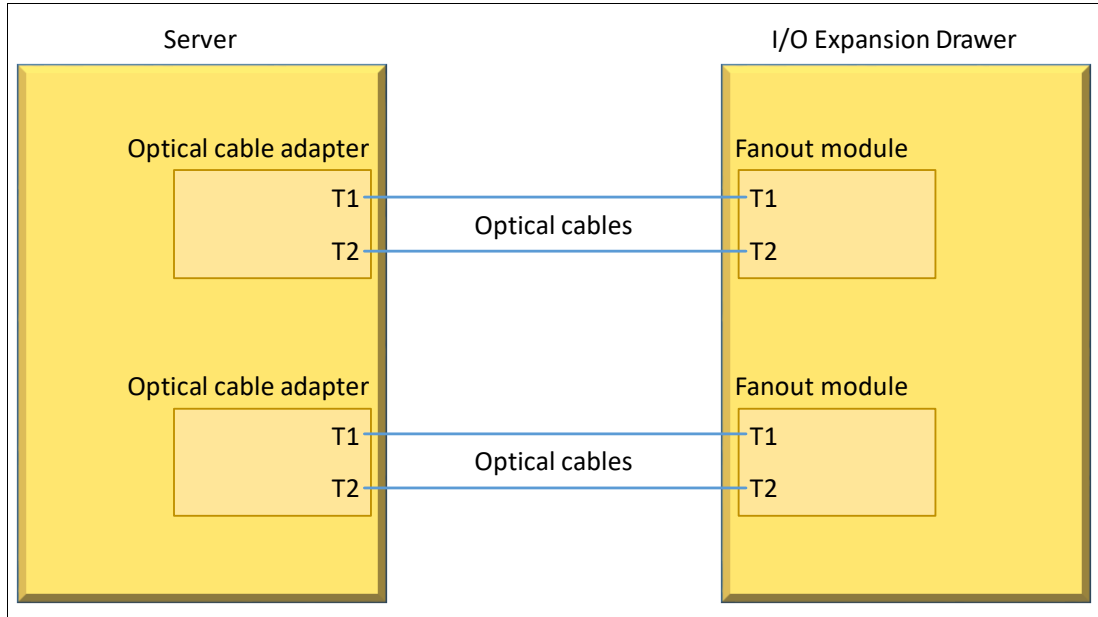
Figure 3-7   Typical optical cable connections

PCIe Gen3 I/O expansion drawer system power control network cabling

There is no system power control network (SPCN) that is used to control and monitor the status of power and cooling within the I/O drawer. SPCN capabilities are integrated into the optical cables.

Note: Feature conversions are available for earlier versions of the optical cables (#ECC7 - #ECCX, and #ECC8 - #ECCY) and fan-out modules (#EMXG - #EMXH).

Table 3-23 provides an overview of all the PCIe adapters for the I/O expansion drawer that are connected to the Power E1050 server. *Available* means that the adapter is available and is orderable. *Supported* means that the adapter is supported on the Power E1050 server during a model conversion, that is, the adapter works, but additional adapters cannot be ordered on the new system.

Table 3-23   Available and supported I/O adapters for the I/O expansion drawer

Type	Feature Code	Adapter	Available or Supported
Storage	#EC7B	PCIe4 1.6 TB NVMe Flash Adapter x8	Available
	#EC7D	PCIe4 3.2 TB NVMe Flash Adapter x8	Available
	#EC7F	PCIe4 6.4 TB NVMe Flash Adapter x8	Available

Type	Feature Code	Adapter	Available or Supported
FC	#EN1A	PCIe3 32 Gb 2-port Fibre Channel Adapter	Available
	#EN1C	PCIe3 16 Gb 4-port Fibre Channel Adapter	Available
	#EN1E	PCIe3 16 Gb 4-port Fibre Channel Adapter	Available
	#EN1G	PCIe3 2-Port 16 Gb Fibre Channel Adapter	Available
	#EN1J	PCIe4 32 Gb 2-port Optical Fibre Channel Adapter	Available
	#EN2A	PCIe3 16 Gb 2-port Fibre Channel Adapter	Available
SAS	#EJ10	PCIe3 SAS Tape/DVD Adapter Quad-port 6 Gb x8	Available
	#EJ14	12 GB Cache RAID PLUS SAS Adapter Quad-port 6 Gb x8	Available
	#EJ0J	PCIe3 RAID SAS Adapter Quad-port 6 Gb x8	Available
LAN	#EC2U	PCIe3 2-Port 25/10 Gb NIC&ROCE SR/Cu Adapter	Available
	#EN0W	PCIe2 2-port 10/1 GbE BaseT RJ45 Adapter	Available
	#EC66	PCIe4 2-port 100 Gb ROCE EN adapter	Available
	#EC76	PCIe4 2-port 100 Gb No Crypto Connectx-6 DX QFSP56	Available
Crypto	#EJ35	PCIe3 Crypto Coprocessor no BSC 4769	Available
	#EJ37	PCIe3 Crypto Coprocessor BSC-Gen3 4769	Available
USB	#EC6K	PCIe2 2-Port USB 3.0 Adapter	Available
RIO	#EJ2A	PCIe3 x16 to CXP Converter Adapter (support AOC)	Available
Storage	#EC5B	PCIe3 x8 1.6 TB NVMe Flash Adapter	Supported
	#EC5D	PCIe3 x8 3.2 TB NVMe Flash Adapter	Supported
	#EC5F	PCIe3 x8 6.4 TB NVMe Flash Adapter	Supported
SAS	#EJ0L	PCIe3 12 GB Cache RAID SAS Adapter Quad-port 6 Gb x8	Supported
LAN	#5899	PCIe2 4-port 1 GbE Adapter	Supported
	#EC2S	PCIe3 2-Port 10 Gb NIC&ROCE SR/Cu Adapter	Supported
	#EC2T	PCIe3 LP 2-Port 25/10Gb NIC&ROCE SR/Cu Adapter	Supported
	#EN0S	PCIe2 4-Port (10 Gb+1 GbE) SR+RJ45 Adapter	Supported
	#EN0U	PCIe2 4-port (10 Gb+1 GbE) Copper SFP+RJ45 Adapter	Supported
Crypto	#EJ32	PCIe3 Crypto Coprocessor no BSC 4767	Supported
	#EJ33	PCIe3 Crypto Coprocessor BSC-Gen3 4767	Supported

3.9.2 EXP24SX SAS storage enclosure

The EXP24SX SAS storage enclosure (#ESLS) is the only DASD drawer that is available for the Power E1050 that provides extra disk drives.

The EXP24SX drawer is a storage expansion enclosure with twenty-four 2.5-inch SFF SAS bays. It supports up to 24 hot-plug HDDs or SSDs in only 2 EIA of space in a 19-inch rack. The EXP24SX SFF bays use SFF Gen2 (SFF-2) carriers or trays.

The EXP24SX drawers feature the following high-reliability design points:

- ► SAS bays that support hot-swap.
- ► Redundant and hot-plug power and fan assemblies.
- ► Dual power cords.
- ► Redundant and hot-plug Enclosure Services Managers (ESMs).
- ► Redundant data paths to all drives.
- ► LED indicators on drives, bays, ESMs, and power supplies that support problem identification.
- ► Through the SAS adapters and controllers, drives that can be protected with RAID and mirroring and hot-spare capability.

The enclosure has adjustable depth rails and can accommodate rack depths from 59.5 - 75 cm (23.4 - 29.5 inches). Slot filler panels are provided for empty bays when initially shipped from IBM.

Figure 3-8 shows the EXP24SX drawer.

Figure 3-8   EXP24SX drawer

With AIX, Linux, or VIOS, the EXP24SX can be ordered with four sets of six bays (mode 4), two sets of 12 bays (mode 2), or one set of 24 bays (mode 1). It is possible to change the mode setting in the field by using software commands along with a documented procedure.

Table 3-24 provides an overview of all the disks for the EXP24SX that are connected to the Power E1050. *Available* means that the device is available and is orderable. *Supported* means that the device is supported on the Power E1050 during a model conversion, that is, the device works, however additional devices cannot be ordered on the new system.

Table 3-24   Available and supported SAS drives for EXP24SX

Feature Code	Device	Available or Supported
#ESEV	600 GB 10K RPM SAS SFF-2 HDD 4 K	Available
#ESF3	1.2 TB 10K RPM SAS SFF-2 HDD 4 K	Available
#ESFT	1.8 TB 10K RPM SAS SFF-2 HDD 4 K	Available
#ESK1	387 GB Enterprise SAS 5xx SFF-2 SSD	Available
#ESK3	775 GB Enterprise SAS 5xx SFF-2 SSD	Available
#ESK8	387 GB Enterprise SAS 4k SFF-2 SSD	Available
#ESKC	775 GB Enterprise SAS 4k SFF-2 SSD	Available
#ESKG	1.55 TB Enterprise SAS 4k SFF-2 SSD	Available
#ESKK	931 GB Mainstream SAS 4k SFF-2 SSD	Available
#ESKP	1.86 TB Mainstream SAS 4k SFF-2 SSD	Available
#ESKT	3.72 TB Mainstream SAS 4k SFF-2 SSD	Available
#ESKX	7.44 TB Mainstream SAS 4k SFF-2 SSD	Available
#ESMB	931 GB Mainstream SAS 4k SFF-2 SSD	Available
#ESMF	1.86 TB Mainstream SAS 4k SFF-2 SSD	Available
#ESMK	3.72 TB Mainstream SAS 4k SFF-2 SSD	Available
#ESMV	7.44 TB Mainstream SAS 4k SFF-2 SSD	Available
#ESNM	300 GB 15 K RPM SAS SFF-2 4k Block Cached Disk Drive	Available
#ESNR	600 GB 15 K RPM SAS SFF-2 4k Block Cached Disk Drive	Available
#ETK1	387 GB Enterprise SAS 5xx SFF-2 SSD	Available
#ETK3	775 GB Enterprise SAS 5xx SFF-2 SSD	Available
#ETK8	387 GB Enterprise SAS 4k SFF-2 SSD	Available
#ETKC	775 GB Enterprise SAS 4k SFF-2 SSD	Available
#ETKG	1.55 TB Enterprise SAS 4k SFF-2 SSD	Available
#1953	300 GB 15k RPM SAS SFF-2 Disk Drive	Supported
#1964	600 GB 10k RPM SAS SFF-2 HDD	Supported
#ES94	387 GB Enterprise SAS 4k SFF-2 SSD	Supported
#ESB2	387 GB Enterprise SAS 5xx SFF-2 SSD	Supported
#ESB6	775 GB Enterprise SAS 5xx SFF-2 SSD	Supported

Feature Code	Device	Available or Supported
#ESBA	387 GB Enterprise SAS 4k SFF-2 SSD	Supported
#ESBG	775 GB Enterprise SAS 4k SFF-2 SSD	Supported
#ESBL	1.55 TB Enterprise SAS 4k SFF-2 SSD	Supported
#ESGV	387 GB Enterprise SAS 5xx SFF-2 SSD	Supported
#ESGZ	775 GB Enterprise SAS 5xx SFF-2 SSD	Supported
#ESJ0	931 GB Mainstream SAS 4k SFF-2 SSD	Supported
#ESJ2	1.86 TB Mainstream SAS 4k SFF-2 SSD	Supported
#ESJ4	3.72 TB Mainstream SAS 4k SFF-2 SSD	Supported
#ESJ6	7.45 TB Mainstream SAS 4k SFF-2 SSD	Supported
#ESJJ	931 GB Mainstream SAS 4k SFF-2 SSD	Supported
#ESJL	1.86 TB Mainstream SAS 4k SFF-2 SSD	Supported
#ESJN	3.72 TB Mainstream SAS 4k SFF-2 SSD	Supported
#ESJQ	7.44 TB Mainstream SAS 4k SFF-2 SSD	Supported
#ESNA	775 GB Enterprise SAS 4k SFF-2 SSD	Supported
#ESNE	1.55 TB Enterprise SAS 4k SFF-2 SSD	Supported

Note: For the EXP24SX drawer, a maximum of twenty-four 2.5-inch SSDs or 2.5-inch HDDs are supported in the #ESLS 24 SAS bays. HDDs and SSDs cannot be mixed in the same mode-1 drawer. HDDs and SSDs can be mixed in a mode-2 or mode-4 drawer, but they cannot be mixed within a logical split of the drawer. For example, in a mode-2 drawer with two sets of 12 bays, one set can hold SSDs and one set can hold HDDs, but you cannot mix SSDs and HDDs in the same set of 12 bays.

Important: When changing modes, a skilled, technically qualified person must follow the special documented procedures. Improperly changing modes can destroy RAID sets, which prevent access to data, or allow other partitions to access another partition's data.

The front view of the ESLS storage enclosure with mode group and drive locations is shown in Figure 3-9.
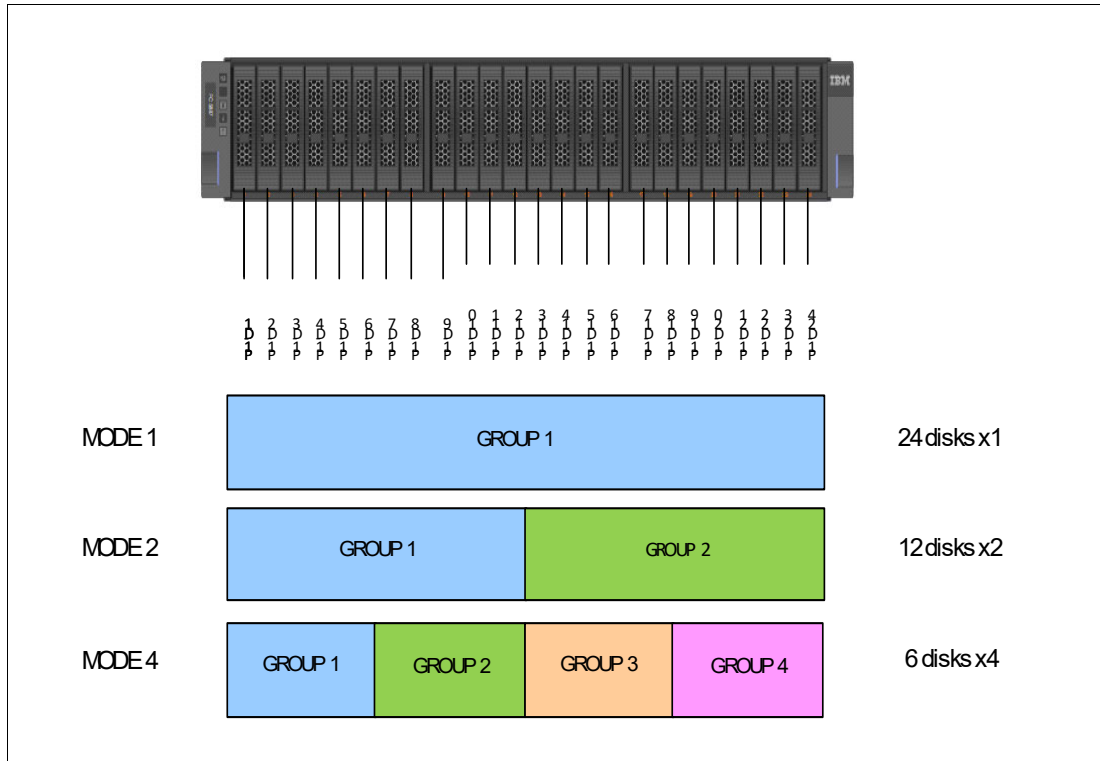
Figure 3-9   Front view of the ESLS storage enclosure with mode groups and drive locations

Four mini-SAS HD ports on the EXP24SX are attached to PCIe Gen3 SAS adapters. The following PCIe3 SAS adapters support the EXP24SX:

► PCIe3 RAID SAS Adapter Quad-port 6 Gb x8 (#EJ0J)

► PCIe3 12 GB Cache RAID SAS Adapter Quad-port 6 Gb x8 (#EJ0L)

► PCIe3 LP RAID SAS Adapter Quad-port 6 Gb x8 (#EJ0M)

► PCIe3 12 GB Cache RAID Plus SAS Adapter Quad-port 6 Gb x8 (#EJ14)

The attachment between the EXP24SX drawer and the PCIe Gen 3 SAS adapter is through SAS YO12 or X12 cables. The PCIe Gen 3 SAS adapters support 6 Gb throughput. The EXP24SX drawer can support up to 12 Gb throughput if future SAS adapters support that capability. Thee Cable options are:

► 3.0M SAS X12 Cable (Two Adapter to Enclosure) (#ECDJ)

► 4.5M SAS X12 Active Optical Cable (Two Adapter to Enclosure) (#ECDK)

► 10M SAS X12 Active Optical Cable (Two Adapter to Enclosure) (#ECDL)

► 1.5M SAS YO12 Cable (Adapter to Enclosure) (#ECDT)

► 3.0M SAS YO12 Cable (Adapter to Enclosure) (#ECDU)

► 4.5M SAS YO12 Active Optical Cable (Adapter to Enclosure) (#ECDV)

► 10M SAS YO12 Active Optical Cable (Adapter to Enclosure) (#ECDW)

Six SAS connectors are at the rear of the EXP24SX drawers to which SAS adapters or controllers are attached. They are labeled T1, T2, and T3: two T1s, two T2s, and two T3s connectors. Consider the following points:

► In mode 1, two or four of the six ports are used. Two T2 ports are used for a single SAS adapter, and two T2 and two T3 ports are used with a paired set of two adapters or a dual adapters configuration.

► In mode 2 or mode 4, four ports are used (two T2s and two T3 connectors) to access all the SAS bays.

► The T1 connectors are not used.

Figure 3-10 shows the connector locations for the EXP24SX storage enclosure.
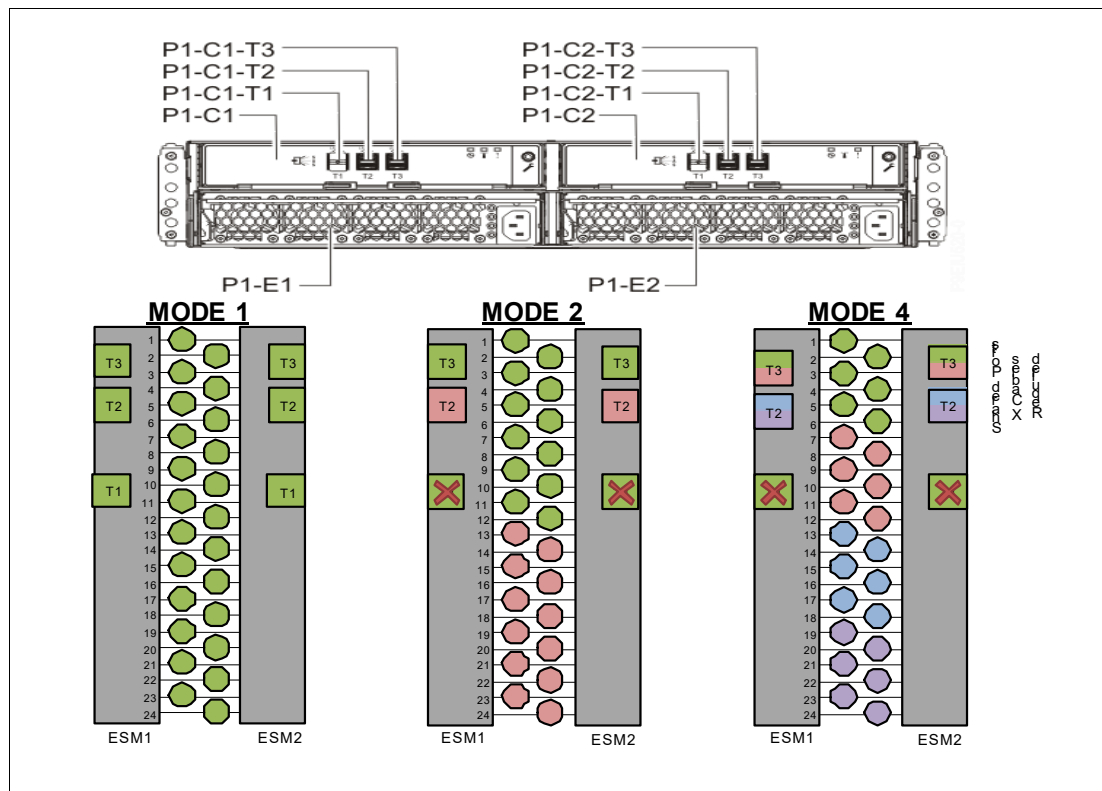
Figure 3-10   Rear view of the EXP24SX with location codes and different split modes

For more information about SAS cabling and cabling configurations, see SAS cabling for the ESLS storage enclosures.

Note: Up to 64 EXP24SX or EXP12SX storage drawers can be attached to a Power E1050 server. With 24 drives per drawer, up to 1,536 SFF-2 drives are supported. A maximum of 16 EXP24SX per PCIe Gen3 I/O drawers are supported due to cable management considerations.

3.9.3 IBM storage

The IBM System Storage Disk Systems products and offerings provide compelling storage solutions with superior value for all levels of business, from entry-level to high-end storage systems. IBM Storage simplifies data infrastructure by using an underlying software foundation to strengthen and streamline the storage in the hybrid cloud environment, which uses a simplified approach to containerization, management, and data protection.

For more information about the various offerings, see Data Storage Solutions.

The following sections highlight a few of these offerings.

IBM Elastic Storage System

IBM Elastic Storage® System is a modern implementation of software-defined storage (SDS). The IBM Elastic Storage System 3200 and IBM Elastic Storage System 5000 make it easier for you to deploy fast, highly scalable storage for artificial intelligence (AI) and big data. With the low latency and high-performance NVMe storage technology and 8Y global file system and global data services of IBM Spectrum® Scale, the IBM Elastic Storage System 3200 and 5000 nodes can grow to over yottabyte (YB) configurations and be integrated into a federated global storage system. For more information, see IBM Elastic Storage System.

IBM FlashSystem: Flash data storage

The IBM FlashSystem® family is a portfolio of cloud-enabled storage systems that can be easily deployed and quickly scaled to help optimize storage configurations, streamline issue resolution, and reduce storage costs. IBM FlashSystem is built with IBM Spectrum Virtualize software to help deploy sophisticated hybrid cloud storage solutions, accelerate infrastructure modernization, address security needs, and maximize value by using the power of AI. The new IBM FlashSystem models provide enterprise-grade functions and deliver the performance to facilitate cybersecurity without compromising production workloads. They offer the advantages of end-to-end NVMe, the innovation of IBM FlashCore® technology, and single-chip module (SCM) for ultra-low latency. For more information, see IBM FlashSystem.

IBM DS8000 storage system

IBM DS8900F is the next generation of enterprise data systems that are built with the most advanced Power processor-based technology and feature ultra-low application response times. Designed for data-intensive and mission-critical workloads, DS8900F adds next-level performance, data protection, resiliency, and availability across hybrid cloud solutions. This outcome is made possible through ultra-low latency, better than seven 9s (99.99999) availability, transparent cloud, tiering, and advanced data protection against malware and ransomware. Additionally, this enterprise class storage solution provides superior performance and higher capacity, which enables the consolidation of all mission-critical workloads in one place. IBM DS8900F can provide 100% data encryption at-rest, in-flight, and in the cloud. For more information, see IBM DS8000 Storage system.

IBM Tape Solutions

A simple and inexpensive data resilience solution virtually impervious to cyberattacks exists from one of the oldest technologies in the data center: tape. The solution requires users to simply remove the tapes storing their data from their networks and stack the tapes on the nearest shelf. This air gap that is created between the data and troublemakers provides a complete cyber-resilient defense that effectively prevents penetration by hackers. Air gaps are just one of several types of data protection that tape can offer. IBM tape-based data storage solutions provide many data protection features, including data encryption and compression, cloud-based disaster recovery, key management, and write once, read many (WORM) technology. For more information, see IBM Tape Solutions.

IBM SAN Volume Controller

IBM SAN Volume Controller (SVC) is an enterprise-class system that consolidates storage from over 500 IBM and third-party storage systems to improve efficiency, simplify management and operations, modernize storage with new capabilities, and enable a common approach to hybrid cloud regardless of storage system type. IBM SVC provides a complete set of data resilience capabilities with high availability (HA), business continuance, and data security features. By helping effectively maximize the economics of massive volumes of data, the IBM SVC helps improve data value, increase data security, enhance data simplicity, and promote 100% availability with IBM HyperSwap®. The IBM SVC also provides IBM Easy Tier® AI-driven automated tiering to improve performance at a lower cost. For more information, see IBM SVC.

3.10 System racks

The Power E1050 server fits a standard 19-inch rack. The server is certified and tested in the IBM Enterprise racks (7965-S42 and 7014-T42). Customers can choose to place the server in other racks if they are confident that those racks have the strength, rigidity, depth, and hole pattern characteristics that are needed. Contact IBM Support to determine whether other racks are suitable.

> **Order information:** It is a best practice that the Power E1050 server be ordered with an IBM 42U enterprise rack #ECR0 (7965-S42). This rack provides a complete and higher-quality environment for IBM Manufacturing system assembly and testing, and provides a complete package.

If a system is installed in a rack or cabinet that is not from IBM, ensure that the rack meets the requirements that are described in 3.10.6, "Original equipment manufacturer racks" on page 121.

> **Responsibility:** The customer is responsible for ensuring 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.

3.10.1 New rack considerations

Consider the following points when racks are ordered:

► The new IBM Enterprise 42U Slim Rack 7965-S42 offers 42 EIA units (U) of space in a slim footprint.

► The 7014-T42 rack is no longer available to purchase with a Power E1050 server. Installing a Power E1050 server in this rack is still supported.

> **Vertical PDUs:** All PDUs that are installed in a rack that contains a Power E1050 server must be installed horizontally to allow for cable routing in the sides of the rack.

3.10.2 IBM Enterprise 42U Slim Rack 7965-S42

The 2.0-meter (79-inch) Model 7965-S42 is compatible with past and present IBM Power 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 placed easily on standard 24-inch floor tiles.

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

The 7965-S42 rack includes space for up to four PDUs in side pockets. Extra PDUs beyond four are mounted horizontally and each uses 1U of rack space.

The Enterprise Slim Rack comes with options for the installed front door:

► Basic Black/Flat (#ECRM)

► High-End appearance (#ECRF)

► OEM Black (#ECRE)

All options include perforated steel, which provides ventilation, physical security, and visibility of indicator lights in the installed equipment within. All options come with a lock and mechanism included that is identical to the lock on the rear doors. Only one front door must be included for each rack ordered. The basic door (#ECRM) and OEM door (#ECRE) can be hinged on the left or right side.

> **Orientation:** #ECRF must not be flipped because the IBM logo would be upside down.

At the rear of the rack, either a perforated steel rear door (#ECRG) or a Rear Door Heat Exchanger (RDHX) can be installed. The only supported RDHX is IBM machine type 1164-95X, which can remove up to 30,000 watts (102,000 BTU) of heat per hour by using chilled water. The no additional charge Feature Code #ECR2 is included with the Enterprise Slim Rack as an indicator when ordering the RDHX.

The basic door (#ECRG) can be hinged on the left or right side, and includes a lock and mechanism identical to the lock on the front door. Either the basic rear door (#ECRG) or the RDHX indicator (#ECR2) must be included with the order of a new Enterprise Slim Rack.

Due to the depth of the Power E1050 server model, the 5-inch rear rack extension (#ECRK) is required for the Enterprise Slim Rack to accommodate this system. This extension expands the space that is available for cable management and allows the rear door to close safely.

Lifting considerations

Three to four service personnel are required to manually remove or insert a system unit into a rack, given its dimensions, weight, and content. To avoid the need for this many people to assemble at a client site for a service action, a lift tool can be useful. Similarly, if the client has chosen to install this customer setup (CSU) system, similar lifting considerations apply.

The Power E1050 server has a maximum weight of 70.3 kg (155 lb). However, by temporarily removing the power supplies, fans, and RAID assembly, the weight is easily reduced to a maximum of 55 kg (121 lb).

When lowering the Power E1050 server onto its rails in the rack, the server must be tilted on one end about 15 degrees so that the pins on the server enclosure fit onto the rails. This equates to lifting one end of the server 4 cm (1.6 in.). This task can be done by using a tip plate on a lift tool, manually adjusting the load on a lift tool, or tilting during the manual lift. Consider the optional feature #EB2Z lift tool.

3.10.3 AC power distribution unit and rack content

The IBM high-function PDUs provide more electrical power per PDU as earlier IBM PDUs, and they offer better PDU footprint efficiency. In addition, they are intelligent PDUs (iPDUs) that provide insight to actual power usage by receptacle and also provide remote power on/off capability for easier support by individual receptacle. The latest PDUs can be ordered as #ECJJ, #ECJL, #ECJN, and #ECJQ.

IBM Manufacturing integrates only the newer PDUs with the Power E1050 server. IBM Manufacturing does not support integrating earlier PDUs, such as #7188, #7109, or #7196. Clients can choose to use older IBM PDUs in their racks, but must install those earlier PDUs at their site.

Table 3-25 summarizes the high-function PDU FCs for 7965-S42 followed by a descriptive list.

Table 3-25   High-function PDUs that are available with IBM Enterprise Slim Rack (7965-S42)

PDUs	1-phase or 3-phase wye depending on country wiring standards	3-phase 208 V depending on country wiring standards
Nine C19 receptacles[a]	#ECJJ	#ECJL
Twelve C13 receptacles	#ECJN	#ECJQ

a. The Power E1050 server has an AC power supply with a C19/C20 connector.

Power sockets: The Power E1050 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.

▶ High Function 9xC19 PDU plus (#ECJJ)

This intelligent, switched 200 - 240 V AC PDU includes nine C19 receptacles on the front of the PDU. The PDU is mounted on the rear of the rack, which makes the nine C19 receptacles easily accessible. For comparison, this PDU is most like the earlier generation #EPTJ PDU.

▶ High Function 9xC19 PDU plus 3-Phase (#ECJL)

This intelligent, switched 208 V 3-phase AC PDU includes nine C19 receptacles on the front of the PDU. The PDU is mounted on the rear of the rack, which makes the nine C19 receptacles easily accessible. For comparison, this PDU is most like the earlier generation #EPTL PDU.

► High Function 12xC13 PDU plus (#ECJN)

This intelligent, switched 200 - 240 V AC PDU includes 12 C13 receptacles on the front of the PDU. The PDU is mounted on the rear of the rack, which makes the 12 C13 receptacles easily accessible. For comparison, this PDU is most like the earlier generation #EPTN PDU.

► High Function 12xC13 PDU plus 3-Phase (#ECJQ)

This intelligent, switched 208 V 3-phase AC PDU includes 12 C13 receptacles on the front of the PDU. The PDU is mounted on the rear of the rack, which makes the 12 C13 receptacles easily accessible. For comparison, this PDU is most like the earlier generation #EPTQ PDU.

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 3-26 shows the available wall power cord options for the PDU features, which must be ordered separately.

Table 3-26   PDU-to-wall power cord options for the PDU features

Feature Code	Wall plug	Rated voltage (V AC)	Phase	Rated amperage	Geography
#6489	IEC309, 3P+N+G, 32 A	230	3	32 amps per phase	Europe, Middle East, and Asia (EMEA)
#6491	IEC 309, P+N+G, 63 A	230	1	63 amps	EMEA
#6492	IEC 309, 2P+G, 60 A	200 - 208, 240	1	48 amps	US, Canada, Latin America (LA), and Japan
#6653	IEC 309, 3P+N+G, 16 A	230	3	16 amps per phase	Internationally available
#6654	NEMA L6-30	200 - 208, 240	1	24 amps	US, Canada, LA, and Japan
#6655	RS 3750DP (watertight)	200 - 208, 240	1	24 amps	US, Canada, LA, and Japan
#6656	IEC 309, P+N+G, 32 A	230	1	24 amps	EMEA
#6657	PDL	230 - 240	1	32 amps	Australia and New Zealand
#6658	Korean plug	220	1	30 amps	North and South Korea
#6667	PDL	380 - 415	3	32 amps	Australia and New Zealand

Notes: Ensure that the suitable 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, the Power E1050 server has four power supplies that must be connected to separate PDUs, which are not included in the base order.

For maximum availability, a best practice 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 more information about power requirements of and the power cord for the 7965-94Y rack, see IBM Documentation.

3.10.4 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 the rack from the bottom.
► IBM recommends the use of an IBM approved lift tool for installation of systems into any IBM or non-IBM rack.
► IBM does not support installation of the server nodes higher than the 29U position.
► Any remaining space in the rack can be used to install other systems or peripheral devices. Ensure that 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, follow the rack manufacturer's safety instructions regarding rack stability.

Order information: The racking approach for the initial order must be 7965-S42 or #ECR0. If an extra rack is required for I/O expansion drawers, an MES to a system or an #0553 must be ordered.

3.10.5 Useful rack additions

This section highlights several rack addition solutions for IBM Power rack-based systems.

IBM System Storage 7226 Model 1U3 Multi-Media Enclosure

The IBM System Storage 7226 Model 1U3 Multi-Media Enclosure can accommodate up to two tape drives, two RDX removable disk drive docking stations, or up to four DVD-RAM drives.

The IBM System Storage 7226 Multi-Media Enclosure supports LTO Ultrium and DAT160 Tape technology, DVD-RAM, and RDX removable storage requirements on the following IBM systems:

► IBM Power6 processor-based systems
► IBM Power7 processor-based systems
► IBM Power8 processor-based systems
► IBM Power9 processor-based systems

► IBM Power10 processor-based systems

The IBM System Storage 7226 Multi-Media Enclosure offers an expansive list of drive feature options, as listed in Table 3-27.

Table 3-27   Supported drive features for the 7226-1U3

Feature Code	Description	Status
#1420	DVD-RAM SAS Optical Drive	Available
#1422	DVD-RAM Slim SAS Optical Drive	Available
#5762	DVD-RAM USB Optical Drive	Available
#5763	DVD Front USB Port Sled with DVD-RAM USB Drive	Available
#5757	DVD RAM Slim USB Optical Drive	Available
#8348	LTO Ultrium 6 Half High Fibre Tape Drive	Available
#8341	LTO Ultrium 6 Half High SAS Tape Drive	Available
#8441	LTO Ultrium 7 Half High SAS Tape Drive	Available
#8546	LTO Ultrium 8 Half High Fibre Tape Drive	Available
#EU03	RDX 3.0 Removable Disk Docking Station	Available

The following options are available:

► LTO Ultrium 6 Half-High 2.5 TB SAS and FC Tape Drive: With a data transfer rate up to 320 MBps (assuming a 2.5:1 compression), the LTO Ultrium 6 drive is read/write compatible with LTO Ultrium 6 and 5 media, and read-only compatibility with LTO Ultrium 4. By using data compression, an LTO-6 cartridge can store up to 6.25 TB of data.

► The LTO Ultrium 7 drive offers a data rate of up to 300 MBps with compression. It also provides read/write compatibility with Ultrium 7 and Ultrium 6 media formats, and read-only compatibility with Ultrium 5 media formats. By using data compression, an LTO-7 cartridge can store up to 15 TB of data.

► The LTO Ultrium 8 drive offers a data rate of up to 300 MBps with compression. It also provides read/write compatibility with Ultrium 8 and Ultrium 7 media formats. It is not read/write compatible with other Ultrium media formats. By using data compression, an LTO-8 cartridge can store up to 30 TB of data.

► DVD-RAM: The 9.4 GB SAS Slim Optical Drive with an SAS and USB interface option is compatible with most standard DVD disks.

► RDX removable disk drives: The RDX USB docking station is compatible with most RDX removable disk drive cartridges when it is used in the same OS. The 7226 offers the following RDX removable drive capacity options:

  – 500 GB (#1107)
  – 1.0 TB (#EU01)
  – 2.0 TB (#EU2T)

Removable RDX drives are in a rugged cartridge that inserts in to an RDX removable (USB) disk docking station (#1103 or #EU03). RDX drives are compatible with docking stations, which are installed internally in Power8, Power9, and Power10 processor-based servers, where applicable.

Figure 3-11 shows the IBM System Storage 7226 Multi-Media Enclosure.

Figure 3-11   IBM System Storage 7226 Multi-Media Enclosure

The IBM System Storage 7226 Multi-Media Enclosure offers a customer-replaceable unit (CRU) maintenance service to help make the installation or replacement of new drives efficient. Other 7226 components also are designed for CRU maintenance.

The IBM System Storage 7226 Multi-Media Enclosure is compatible with most Power8, Power9, and Power10 processor-based systems that offer current level AIX and Linux OSs.

For a complete list of host software versions and release levels that support the IBM System Storage 7226 Multi-Media Enclosure, see IBM System Storage Interoperation Center (SSIC).

> **Note:** Any of the existing 7216-1U2, 7216-1U3, and 7214-1U2 multimedia drawers are also supported.

Flat panel display options

The IBM 7316 Model TF5 is a rack-mountable flat panel console kit that also can be configured with the tray pulled forward and the monitor folded up, which provides full viewing and keying capability for the HMC operator.

The Model TF5 is a follow-on product to the Model TF4 and offers the following features:

► A slim, sleek, and lightweight monitor design that occupies only 1U (1.75 in.) in a 19-inch standard rack

► A 18.5 inch (409.8 mm x 230.4 mm) flat panel TFT monitor with truly accurate images and virtually no distortion

► The ability to mount the IBM Travel Keyboard in the 7316-TF5 rack keyboard tray

► Support for the IBM 1x8 Rack Console Switch (#4283) IBM Keyboard/Video/Mouse (KVM) switches

   The #4283 is a 1x8 Console Switch that fits in the 1U space behind the TF5. It is a CAT5-based switch. It contains eight analog rack interface (ARI) ports for connecting PS/2 or USB console switch cables. It supports chaining of servers that use an IBM Conversion Options switch cable (#4269). This feature provides four cables that connect a KVM switch to a system, or can be used in a daisy-chain scenario to connect up to 128 systems to a single KVM switch. It also supports server-side USB attachments.

3.10.6 Original equipment manufacturer racks

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

IBM Documentation provides the general rack specifications, including the following information:

► The rack or cabinet must meet the EIA Standard EIA-310-D for 19-inch racks, which was published August 24, 1992. The EIA-310-D standard specifies internal dimensions, for example, the width of the rack opening (width of the chassis), the width of the module mounting flanges, and the mounting hole spacing.

► The front rack opening must be a minimum of 450 mm (17.72 in.) wide, and the rail-mounting holes must be 465 mm plus or minus 1.6 mm (18.3 in. plus or minus 0.06 in.) apart on center (horizontal width between vertical columns of holes on the two front-mounting flanges and on the two rear-mounting flanges).

Figure 3-12 is a top view showing the rack specification dimensions.
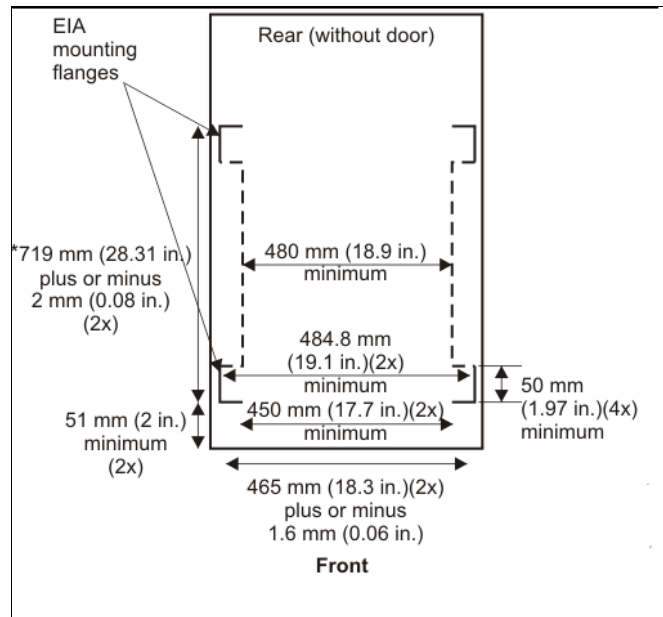
Figure 3-12   Rack specifications (top-down view)

► The vertical distance between mounting holes must consist of sets of three holes that are spaced (from bottom to top) 15.9 mm (0.625 in.), 15.9 mm (0.625 in.), and 12.7 mm (0.5 in.) on center, which makes each three-hole set of vertical hole spacing 44.45 mm (1.75 in.) apart on center.

Figure 3-13 shows the vertical distances between the mounting holes.
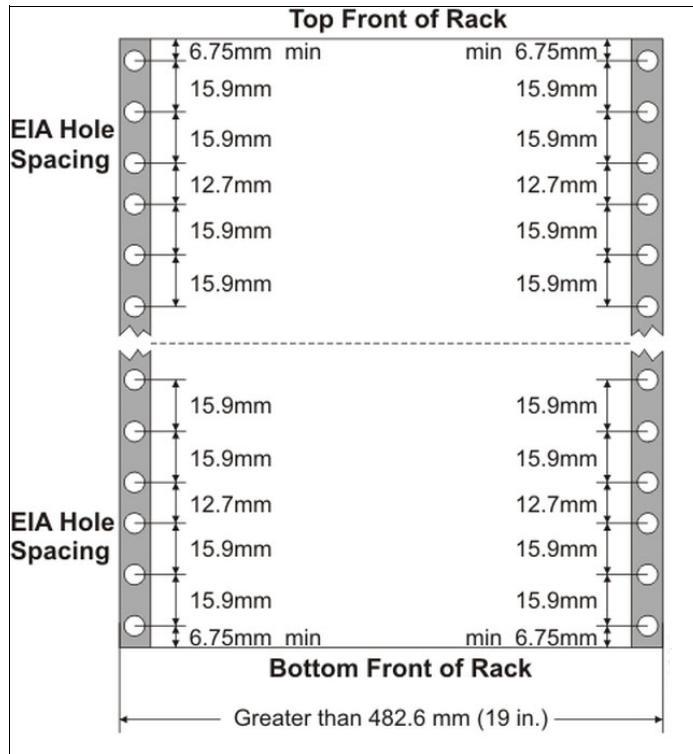
Figure 3-13   Vertical distances between mounting holes

► The following rack hole sizes are supported for racks where IBM hardware is mounted:

– 7.1 mm (0.28 in.) plus or minus 0.1 mm (round)
– 9.5 mm (0.37 in.) plus or minus 0.1 mm (square)

The rack or cabinet must be capable of supporting an average load of 20 kg (44 lb.) of product weight per EIA unit. For example, a four EIA drawer has a maximum drawer weight of 80 kg (176 lb.).

4

Resiliency

The reliability of systems starts with components, devices, and subsystems that are highly reliable. During the design and development process, subsystems go through rigorous verification and integration testing processes. During system manufacturing, systems go through a thorough testing process to help ensure the highest level of product quality.

Mathematically, reliability is defined in terms of how infrequently something fails.

At a system level, availability is about how infrequent failures cause workload interruptions. The longer the interval between interruptions, the more available a system is.

Serviceability is about how efficiently failures are identified and dealt with, and how application outages are minimized during repair.

The Power10 E1050 comes with the following reliability, availability, and serviceability (RAS) characteristics:

► Enterprise baseboard management controller (BMC) service processor for system management and service

► Open Memory Interface (OMI) and Differential Dual Inline Memory Modules (DDIMMS) RAS

► Power10 processor RAS

► I/O subsystem RAS

► Serviceability

These points are described in this chapter.

Additionally, the following points also are part of the Power E1050 RAS:

► Redundant and hot-plug cooling

► Redundant and hot-plug power

► Redundant voltage regulators

► Time of day battery concurrent maintenance

Table 4-1 shows the differences, from a RAS perspective, in the Power10 processor-based family.

Table 4-1   Power10 Servers RAS highlights comparison

Item	IBM Power S10xx	IBM Power E1050	IBM Power E1080
Base Power10 processor RAS features: ▶ First Failure Data Capture (FFDC) ▶ Processor Instruction Retry (PIR) ▶ L2/L3 Cache error correction code (ECC) protection with cache line-delete ▶ Power/cooling monitor function integrated into on-chip controllers of the processors	Yes.	Yes.	Yes.
Power10 Enterprise RAS feature: Core Checkstops	Yes.	Yes.	Yes.
Multi-node symmetric multiprocessing Fabric RAS: cyclic redundancy check (CRC)-checked processor fabric bus retry with spare data lane or bandwidth reduction	N/A.	N/A.	Yes. The Power10 processor-based design removes active components on cable and introduces internal cables to reduce backplane replacements.
Hot-plug with processor integrated	Yes.	Yes.	Yes.
Memory DIMM ECC supporting x4	Yes.	Yes.	Yes.
Dynamic Memory Row repair and spare dynamic RAM (DRAM) capability	2U DDIMM: No spare DRAM. Yes: Dynamic Row Repair.	Yes: Base. 4U DDIMM with 2 spare DRAMs per rank. Yes: Dynamic Row Repair.	Yes: Base. 4U DDIMM with 2 spare DRAMs per rank. Yes: Dynamic Row Repair.
Active Memory Mirroring (AMM) for Hypervisor	Yes: Base. New to scale-out.	Yes: Base.	Yes: Base.
Redundant or spare voltage phases on voltage converters for levels feeding processor	No.	Yes: N+1.	Yes: N+2.
Redundant On-board Power Management Integrated Circuits (PMICs) memory DDIMMs	No. 2U DDIMM.	Yes: Base. 4U DDIMM.	Yes: Base.
Service processor type	Enterprise Baseboard Management Controller (eBMC). Open standard with Redfish support.	eBMC. Open standard with Redfish support.	Flexible Service Processor (FSP). IBM proprietary.
Processor clocks redundancy/sparing	No.	No.	Redundant.
Redundant service processor and related boot facilities	No.	No.	Yes.

Item	IBM Power S10xx	IBM Power E1050	IBM Power E1080
Redundant trusted platform module (TPM) capability	No.	No.	Yes.
Multi-node support	N/A.	N/A.	Yes.

4.1 Service processor

The Power10 E1050 comes with a redesigned service processor that is based on a BMC design with firmware that is accessible through open-source industry-standard application programming interfaces (APIs), such as Redfish. An upgraded Advanced System Management Interface (ASMI) web browser user interface preserves the required RAS functions while allowing the user to perform tasks in a more intuitive way.

Diagnostic monitoring of recoverable errors from the processor chipset is performed on the system processor itself, and the unrecoverable diagnostic monitoring of the processor chipset is performed by the service processor. The service processor runs on its own power boundary and does not require resources from a system processor to be operational to perform its tasks.

The service processor supports surveillance of the connection to the Hardware Management Console (HMC) and to the system firmware (hypervisor). It also provides several remote power control options, environmental monitoring, reset, restart, remote maintenance, and diagnostic functions, including console mirroring. The BMC service processors menus (ASMI) can be accessed concurrently during system operation, allowing nondisruptive abilities to change system default parameters, view and download error logs, and check system health.

Redfish, an industry-standard API for server management, enables IBM Power servers to be managed individually or in a large data center. Standard functions such as inventory, event logs, sensors, dumps, and certificate management are all supported by Redfish. In addition, new user management features support multiple users and privileges on the BMC through Redfish or ASMI. User management through lightweight directory access protocol (LDAP) also is supported. The Redfish events service provides a means for notification of specific critical events such that actions can be taken to correct issues. The Redfish telemetry service provides access to a wide variety of data (such as power consumption, and ambient, core, DIMMs, and I/O temperatures) that can be streamed at periodic intervals.

The service processor monitors the operation of the firmware during the boot process and also monitors the hypervisor for termination. The hypervisor monitors the service processor and reports a service reference code when it detects surveillance loss. In the PowerVM environment, it performs a reset/reload if it detects the loss of the service processor.

4.2 Memory subsystem RAS

The Power10 E1050 server introduces a new 4U tall (DDIMM) subsystem, which has a new OpenCAPI memory interface that is called OMI, for resilient and fast communication to the processor.
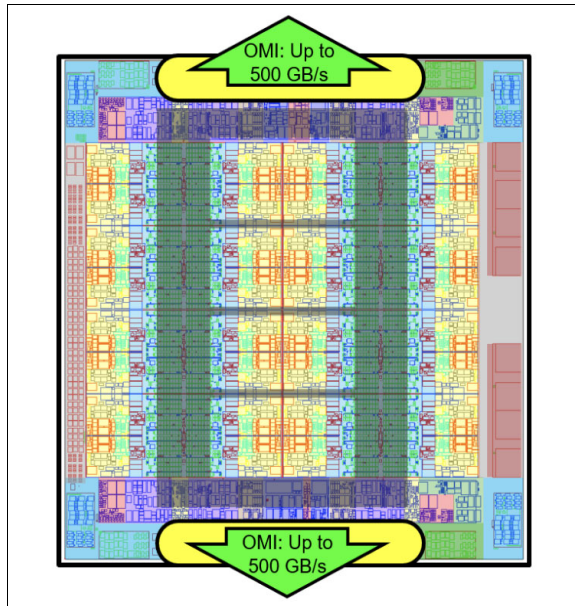
Figure 4-1   Power10 E1050 OMI

This new memory subsystem design delivers solid RAS. Unlike the processor RAS characteristics, the E1050 memory RAS varies significantly from that of the IBM Power E950. The Power E1050 supports the same 4U DDIMM height as the IBM Power E1080.

Table 4-2 compares memory DIMMs, and highlights the differences between the Power E950 DIMM and the Power E1050 DDIMM. It also provides the RAS impacts of the DDIMMs, which are applicable to the Power E1080 servers.

Table 4-2   Power E950 DIMMs versus Power E1050 DDIMMs RAS comparison

Item	Power E950 memory	Power E1050 memory	RAS impact
DIMM form factor	Riser card plus industry-standard DIMMs	4U DIMM	▶ Power E1050 4U DDIMM: Single field-replaceable unit (FRU) or fewer components to replace. ▶ Power E950 DIMM: A separate FRU is used for a memory buffer on a riser card and the industry-standard DIMMs.
Symbol correction	Single symbol correction	Dual symbol correction	▶ Power E1050 4U DDIMM: A data pin failure (1 symbol) that lines up with single cell failure on another DRAM is still correctable. ▶ Power E950 DIMM: A data pin failure (1 symbol) that lines up with single cell failure on another DRAM is uncorrectable.

Item	Power E950 memory	Power E1050 memory	RAS impact
X4 Chipkill	One spare DRAM across a DIMM repair	Single DRAM Chipkill correction plus	► Power E1050 4U DDIMM: – First Chipkill is fixed with a spare. – Second Chipkill is fixed with a spare. – Third Chipkill is fixed with ECC. – Fourth Chipkill is uncorrectable. ► Power E950 DIMM: – First Chipkill is fixed with a spare. – Second Chipkill is fixed with ECC. – Third Chipkill is uncorrectable.
DRAM row repair	Static	Dynamic	► Power E1050 4U DDIMM: Detect, fix, and restore at run time without a system outage. ► Power E950 DIMM: Detect at run time, but a fix and restore requires a system restart.
L4 cache	Yes	No	► Power E1050 4U DDIMM: Avoids L4 cache failure modes. ► Power E950 DIMM: The L4 cache fail rate is a significant contributor to DIMM replacements.
Voltage regulation redundancy	No	Yes	► Power E1050 4U DDIMM: Can survive a voltage regulation component failure. ► Power E950 DIMM: Voltage regulation and associated components are a single point of failure.

4.2.1 Memory buffer

The DDIMM contains a memory buffer with key RAS features, including protection of critical data/address flows by using CRC, ECC, and parity; a maintenance engine for background memory scrubbing and memory diagnostics; and a Fault Isolation Register (FIR) structure, which enables firmware attention-based fault isolation and diagnostics.

4.2.2 Open Memory Interface

The OMI interface between the memory buffer and processor memory controller is protected by dynamic lane calibration, and a CRC retry/recovery facility to retransmit lost frames to survive intermittent bit flips. A lane fail also can be survived by triggering a dynamic lane reduction from 8 to 4, independently for both up and downstream directions. A key advantage of the OMI interface is that it simplifies the number of critical signals that must cross connectors from processor to memory compared to a typical industry-standard DIMM design.

4.2.3 Memory ECC

The DDIMM includes a robust 64-byte Memory ECC with 8-bit symbols, which can correct up to five symbol errors (one x4 chip and one additional symbol), and retry for data and address uncorrectable errors.

4.2.4 Dynamic row repair

To further extend the life of the DDIMM, the dynamic row repair feature can restore full use of a DRAM for a fault that is contained to a DRAM row while the system continues to operate.

4.2.5 Spare temperature sensors

Each DDIMM has spare temperature sensors so that the failure of one does not require a DDIMM replacement.

4.2.6 Spare DRAMs

4U DDIMMs include two spare x4 memory modules (DRAMs) per rank, which can be substituted for failed DRAMs during a runtime operation. Combined with ECC correction, the two spares allow a 4U DDIMM to continue to function with three bad DRAMs per rank, compared to 1 (single device data correct) or 2 (double device data correct) bad DRAMs in a typical industry-standard DIMM design. This setup extends self-healing capabilities beyond what is provided with the dynamic row repair capability.

4.2.7 Spare Power Management Integrated Circuits

4U DDIMMs include spare PMICs so that the failure of one PMIC does not require a DDIMM replacement.

4.3 Power10 processor RAS

Although there are many differences internally in the Power10 processor compared to the Power9 processor that relate to performance, number of cores, and other features, the general RAS philosophy for how errors are handled has remains the same. Therefore, information about Power9 processor-based subsystem RAS can still be referenced to understand the design. For more information, see *Introduction to IBM Power Reliability, Availability, and Serviceability for Power9 processor-based systems using IBM PowerVM*.

The Power E1050 processor module is a dual-chip module (DCM) that differs from that of the Power E950, which has single-chip module (SCM). Each DCM has 30 processor cores, which is 120 cores for a 4-socket (4S) Power E1050. In comparison, a 4S Power E950 supports 48 cores. The internal processor buses are twice as fast with the Power E1050 running at 32 Gbps.

Despite the increased cores and the faster high-speed processor bus interfaces, the RAS capabilities are equivalent, with features like PIR, L2/L3 Cache ECC protection with cache line delete, and the CRC fabric bus retry that is a characteristic of Power9 and Power10 processors. As with the Power E950, when an internal fabric bus lane encounters a hard failure in a Power E1050, the lane can be dynamically spared out.

Figure 4-2 shows the Power10 DCM.
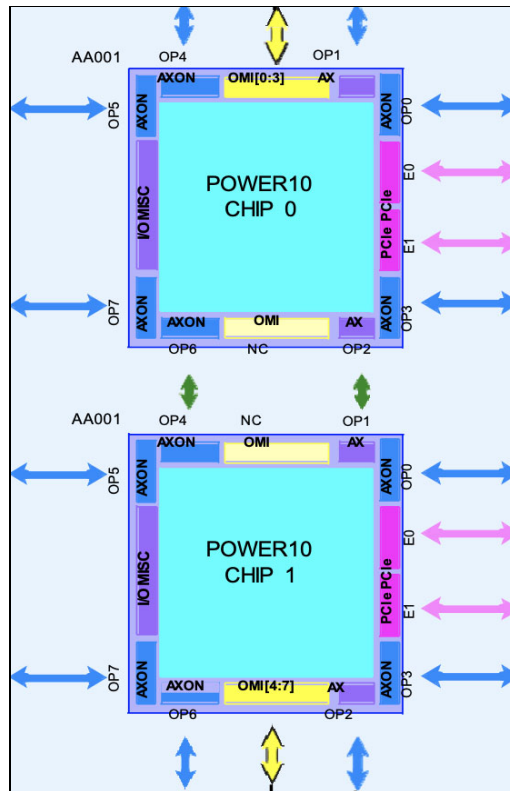
Figure 4-2   Power10 Dual-Chip Module

4.3.1 Cache availability

The L2/L3 caches in the Power10 processor in the memory buffer chip are protected with double-bit detect, single-bit correct ECC. In addition, a threshold of correctable errors that are detected on cache lines can result in the data in the cache lines being purged and the cache lines removed from further operation without requiring a restart in the PowerVM environment. Modified data is handled through Special Uncorrectable Error (SUE) handling. L1 data and instruction caches also have a retry capability for intermittent errors and a cache set delete mechanism for handling solid failures.

4.3.2 Special Uncorrectable Error handling

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

4.3.3 Uncorrectable error recovery

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

4.4 I/O subsystem RAS

The Power E1050 provides 11 general-purpose Peripheral Component Interconnect Express (PCIe) slots that allow for hot-plugging of I/O adapters, which makes the adapters concurrently maintainable. These PCIe slots operate at Gen4 and Gen5 speeds. Some of the PCIe slots support OpenCAPI and I/O expansion drawer cable cards.

Unlike the Power E950, the Power E1050 location codes start from index 0, as with all Power 10 systems. However, slot c0 is not a general-purpose PCIe slot because it is reserved for the eBMC Service Processor card.

Another difference between the Power E950 and the Power E1050 is that all the Power E1050 slots are directly connected to a Power10 processor. In the Power E950, some slots are connected to the Power9 processor through I/O switches.

All 11 PCIe slots are available if 3-socket or 4-socket DCMs are populated. In the 2-socket DCM configuration, only seven PCIe slots are functional.

DASD options

The Power E1050 provides 10 internal Non-volatile Memory Express (NVMe) drives at Gen4 speeds, which means that they are concurrently maintainable. The NVMe drives are connected to DCM0 and DCM3. In a 2-socket DCM configuration, only six of the drives are available. To access to all 10 internal NVMe drives, you must have a 4S DCM configuration. Unlike the Power E950, the Power E1050 has no internal serial-attached SCSI (SAS) drives You can use an external drawer to provide SAS drives.

The internal NVMe drives support OS-controlled RAID 0 and RAID 1 array, but no hardware RAID. For best redundancy, use an OS mirror and dual Virtual I/O Server (VIOS) mirror. To ensure as much separation as possible in the hardware path between mirror pairs, the following NVMe configuration is recommended:

► Mirrored OS: NVMe3 and NVMe4 pairs, or NVMe8 and NVMe9 pairs

► Mirrored dual VIOS:

  – Dual VIOS: NVMe3 for VIOS1, NVMe4 for VIOS2.

  – Mirrored dual VIOS: NVMe9 mirrors NVMe3, and NVMe8 mirrors NVMe4.

4.5 Serviceability

The purpose of serviceability is to efficiently repair the system while attempting to minimize or eliminate any impact to system operation. Serviceability includes system installation, Miscellaneous Equipment Specification (MES) (system upgrades/downgrades), and system maintenance or repair. Depending on the system and warranty contract, service may be performed by the client, an IBM representative, or an authorized warranty service provider. The serviceability features that are delivered in this system help provide a highly efficient service environment by incorporating the following attributes:

► Designed for IBM System Services Representative (IBM SSR) setup, install, and service.

► Error Detection and Fault Isolation (ED/FI).

► FFDC.

► Light path service indicators.

► Service and FRU labels that are available on the system.

► Service procedures are documented in IBM Documentation or available through the HMC.

► Automatic reporting of serviceable events to IBM through the Electronic Service Agent (ESA) Call Home application.

4.5.1 Service environment

In the PowerVM environment, the HMC is a dedicated server that provides functions for configuring and managing servers for either partitioned or full-system partition by using a GUI, command-line interface (CLI), or Representational State Transfer (REST) API. An HMC that is attached to the system enables support personnel (with client authorization) to remotely or locally (by using the physical HMC that is in proximity of the server being serviced) log in to review error logs and perform remote maintenance if required.

The Power10 processor-based servers support several service environments:

► Attachment to one or more HMCs or virtual HMCS (vHMCs) is a supported option by the system with PowerVM. This configuration is the default one for servers supporting logical partitions (LPARs) with dedicated or virtual I/O. In this case, all servers have at least one LPAR.

► No HMC. There are two service strategies for non-HMC systems:

  – Full-system partition with PowerVM: A single partition owns all the server resources and only one operating system (OS) may be installed. The primary service interface is through the OS and the service processor.

  – Partitioned system with NovaLink: In this configuration, the system can have more than one partition and can be running more than one OS. The primary service interface is through the service processor.

4.5.2 Service interface

Support personnel can use the service interface to communicate with the service support applications in a server by using an operator console, a GUI on the management console or service processor, or an OS terminal. The service interface helps to deliver a clear, concise view of available service applications, helping the support team to manage system resources and service information in an efficient and effective way. Applications that are available through the service interface are carefully configured and placed to grant service providers access to important service functions. Different service interfaces are used, depending on the state of the system, hypervisor, and operating environment. The primary service interfaces are:

► LEDs

► Operator panel

► BMC Service Processor menu

► OS service menu

► Service Focal Point (SFP) on the HMC or vHMC with PowerVM

In the light path LED implementation, the system can clearly identify components for replacement by using specific component-level LEDs and also can guide the servicer directly to the component by signaling (turning on solid) the enclosure fault LED and component FRU fault LED. The servicer also can use the identify function to flash the FRU-level LED. When this function is activated, a roll-up to the blue enclosure locate occurs. These enclosure LEDs turn on solid and can be used to follow the light path from the enclosure and down to the specific FRU in the PowerVM environment.

4.5.3 First Failure Data Capture and error data analysis

FFDC is a technique that helps ensure that when a fault is detected in a system, the root cause of the fault is captured without the need to re-create the problem or run any sort of extending tracing or diagnostics program. For most faults, a good FFDC design means that the root cause also can be detected automatically without servicer intervention.

FFDC information, error data analysis, and fault isolation are necessary to implement the advanced serviceability techniques that enable efficient service of the systems and to help determine the failing items.

In the rare absence of FFDC and Error Data Analysis, diagnostics are required to re-create the failure and determine the failing items.

4.5.4 Diagnostics

The general diagnostic objectives are to detect and identify problems so that they can be resolved quickly. Elements of th IBM diagnostics strategy include:

► Provides a common error code format equivalent to a system reference code with a PowerVM, system reference number, checkpoint, or firmware error code.

► Provides fault detection and problem isolation procedures. Supports remote connection, which can be used by the IBM Remote Support Center or IBM Designated Service.

► Provides interactive intelligence within the diagnostics with detailed online failure information while connected to the IBM back-end system.

4.5.5 Automatic diagnostics

The processor and memory FFDC technology is designed to perform without re-creating diagnostics or user intervention. Solid and intermittent errors are designed to be correctly detected and isolated at the time the failure occurs. Runtime and boot-time diagnostics fall into this category.

4.5.6 Stand-alone diagnostics

As the name implies, stand-alone or user-initiated diagnostics requires user intervention. The user must perform manual steps, including:

► Booting from the diagnostics CD, DVD, Universal Serial Bus (USB), or network

► Interactively selecting steps from a list of choices

4.5.7 Concurrent maintenance

The determination of whether a firmware release can be updated concurrently is identified in the readme file that is released with the firmware. An HMC is required for a concurrent firmware update with PowerVM. In addition, concurrent maintenance of PCIe adapters and NVMe drives is supported by PowerVM. Power supplies, fans, and operating panel LCDs are hot-pluggable.

4.5.8 Service labels

Service providers use these labels to assist them in performing maintenance actions. Service labels are found in various formats and positions are intended to transmit readily available information to the servicer during the repair process. Here are some of these service labels and their purposes:

- ► Location diagrams: Location diagrams are on the system hardware, relating information regarding the placement of hardware components. Location diagrams might include location codes, drawings of physical locations, concurrent maintenance status, or other data that is pertinent to a repair. Location diagrams are especially useful when multiple components such as DIMMs, processors, fans, adapters, and power supplies are installed.

- ► Remove/replace procedures: Service labels that contain remove/replace procedures are often found on a cover of the system or in other spots accessible to the servicer. These labels provide systematic procedures, including diagrams detailing how to remove or replace certain serviceable hardware components.

- ► Arrows: Numbered arrows are used to indicate the order of operation and the serviceability direction of components. Some serviceable parts such as latches, levers, and touch points must be pulled or pushed in a certain direction and in a certain order for the mechanical mechanisms to engage or disengage. Arrows generally improve the ease of serviceability.

4.5.9 QR labels

QR labels are placed on the system to provide access to key service functions through a mobile device. When the QR label is scanned, it goes to a landing page for Power10 processor-based systems, which contains each machine type and model (MTM) service functions of interest while physically at the server. These functions include things installation and repair instructions, reference code lookup, and other items.

4.5.10 Packaging for service

The following service features are included in the physical packaging of the systems to facilitate service:

- ► Color coding (touch points): Blue-colored touch points delineate touch points on service components where the component can be safely handled for service actions, such as removal or installation.

- ► Tool-less design: Selected IBM systems support tool-less or simple tool designs. These designs require no tools or simple tools such as flathead screw drivers to service the hardware components.

► Positive retention: Positive retention mechanisms help to ensure proper connections between hardware components, such as cables to connectors, and between two cards that attach to each other. Without positive retention, hardware components run the risk of becoming loose during shipping or installation, preventing a good electrical connection. Positive retention mechanisms like latches, levers, thumbscrews, pop nylatches (U-clips), and cables are included to help prevent loose connections and aid in installing (seating) parts correctly. These positive retention items do not require tools.

4.5.11 Error handling and reporting

In a system hardware or environmentally induced failure, the system runtime error capture capability systematically analyzes the hardware error signature to determine the cause of failure. The analysis result is stored in system NVRAM. When the system can be successfully restarted either manually or automatically, or if the system continues to operate, the error is reported to the OS. Hardware and software failures are recorded in the system log. When an HMC is attached in the PowerVM environment, an ELA routine analyzes the error, forwards the event to the SFP application running on the HMC, and notifies the system administrator that it has isolated a likely cause of the system problem. The service processor event log also records unrecoverable checkstop conditions, forwards them to the SFP application, and notifies the system administrator.

The system can *call home* through the OS to report platform-recoverable errors and errors that are associated with PCI adapters or devices.

In the HMC-managed environment, a Call Home service request is initiated from the HMC, and the pertinent failure data with service parts information and part locations is sent to an IBM service organization. Customer contact information and specific system-related data, such as the MTM and serial number, along with error log data that is related to the failure, are sent to IBM Service.

4.5.12 Live Partition Mobility

With PowerVM Live Partition Mobility (LPM), users can migrate an AIX, IBM i, or Linux VM partition running on one IBM Power partition server to another IBM Power server without disrupting services. The migration transfers the entire system environment, including processor state, memory, attached virtual devices, and connected users. It provides continuous OS and application availability during planned partition outages for repair of hardware and firmware faults. The Power10 servers that use Power10 processor-based technology support secure LPM, where the VM image is encrypted and compressed before transfer. Secure LPM uses on-chip encryption and compression capabilities of the Power10 processor for optimal performance.

4.5.13 Call Home

Call Home refers to an automatic or manual call from a client location to the IBM support structure with error log data, server status, or other service-related information. Call Home invokes the service organization in order for the appropriate service action to begin. Call Home can be done through the ESA that is embedded in the HMC, or through a version of the ESA that is embedded in the OSs for non-HMC-managed or a version of ESA that runs as a stand-alone Call Home application. While configuring Call Home is optional, clients are encouraged to implement this feature to obtain service enhancements such as reduced problem determination and faster and potentially more accurate transmittal of error information. In general, using the Call Home feature can result in increased system availability.

4.5.14 IBM Electronic Services

ESA and Client Support Portal (CSP) comprise the IBM Electronic Services solution, which is dedicated to providing fast, exceptional support to IBM clients. ESA is a no-charge tool that proactively monitors and reports hardware events, such as system errors, and collects hardware and software inventory. ESA can help focus on the client's company business initiatives, save time, and spend less effort managing day-to-day IT maintenance issues. In addition, Call Home Cloud Connect Web and Mobile capability extends the common solution and offers IBM Systems related support information that is applicable to servers and storage.

For more information, see IBM Call Home Connect Cloud.

4.6 Reliability and availability

This section looks at the more general concept of RAS as it applies to any system in the data center.

The goal is to briefly define what RAS is and look at how reliability and availability are measured.

4.6.1 Reliability modeling

The prediction of system level reliability starts with establishing the failure rates of the individual components that make up the system. Then, by using the appropriate prediction models, the component-level failure rates are combined to provide a system-level reliability prediction in terms of a failure rate.

However, in documentation system-level reliability is often described in terms of mean time between failures (MTBF) for repairable systems rather than a failure rate, for example, 50 years MTBF.

A 50-year MTBF might suggest that a system runs 50 years between failures, but what it actually means is that among 50 identical systems, one per year will fail on average over a large population of systems.

4.6.2 Different levels of reliability

When a component fails, the impact of that failure can vary depending on the component.

A power supply failing in a system with a redundant power supply must be replaced. However, by itself a failure of a single power supply should not cause a system outage, and it should be a concurrent repair with no downtime.

Other components in a system might fail and cause a system-wide outage where concurrent repair is not possible. Therefore, it is typical to talk about different MTBF numbers:

► MTBF – Results in repair actions.
► MTBF – Requires concurrent repair.
► MTBF – Requires a non-concurrent repair.
► MTBF – Results in an unplanned application outage.
► MTBF – Results in an unplanned system outage.

4.6.3 Measuring availability

Mathematically speaking, availability is often expressed as a percentage of the time that something is available or in use over a period. An availability number for a system can be mathematically calculated from the expected reliability of the system if both the MTBF and the duration of each outage is known.

For example, consider a system that always runs exactly one week between failures and each time it fails it is down for 10 minutes. For 168 hours in a week, the system is down (10/60) hours. It is up 168 hrs – (10/60) hrs. As a percentage of the hours in the week, it can be said that the system is (168-(1/6))*100% = 99.9% available.

99.999% available means approximately 5.3 minutes of downtime in a year. On average, a system that failed once a year and was down for 5.3 minutes would be 99.999% available. This concept is often called "five 9s of availability".

When talking about modern server hardware availability, short weekly failures like in the example is not the norm. Rather, the failure rates are much lower, and the MTBF is often measured in terms of years, perhaps more years than a system will be kept in service.

Therefore, when an MTBF of 10 years, for example, is quoted, it is not expected that on average each system will run 10 years between failures. Rather, it is more reasonable to expect that on average in a year that 1 server out of 10 will fail. If a population of 10 servers always had exactly one failure a year, a statement of 99.999% availability across that population of servers would mean that the one server that failed would be down about 53 minutes when it failed.

In theory, five 9s of availability can be achieved by having a system design that fails frequently, multiple times a year, but whose failures are limited to small periods of time. Conversely, five 9 s of availability might mean a server design with a large MTBF, but where a server takes a fairly long time to recover from the rare outage.

Figure 4-3 shows that five 9s of availability can be achieved with systems that fail frequently for minuscule amounts of time, or infrequently with much larger downtime per failure.
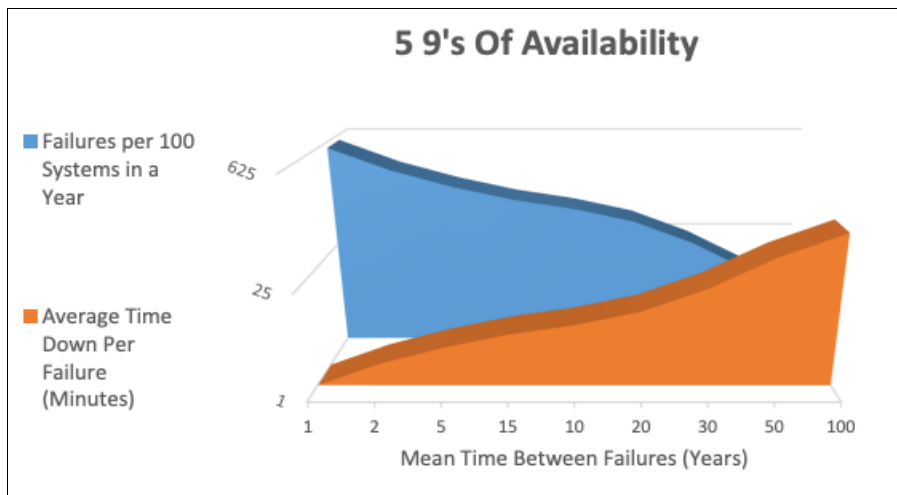
Figure 4-3   Five 9s of availability

Figure 4-3 on page 136 is misleading in the sense that servers with low reliability are likely to have many components that, when they fail, take down the system and keep the system down until repair. Conversely, servers that are designed for great reliability often also are designed so that the systems, or at least portions of the system, can be recovered without having to keep a system down until it is repaired. So, systems with low MTBF would have longer repair times, and a system with five 9s of availability would be synonymous with a high level of reliability.

5

Enterprise solutions

In this chapter, we describe the major solutions that can help enterprises achieve their business goals and the reasons on why IBM Power10 processor-based mid-range servers provide a significant contribution to that end.

5.1 PowerVM virtualization

The PowerVM platform is the family of technologies, capabilities, and offerings that delivers industry-leading virtualization for enterprises. It is the umbrella branding term for IBM Power processor-based server virtualization:

► IBM Power Hypervisor
► Logical partitioning
► IBM Micro-Partitioning®
► Virtual I/O Server (VIOS)
► Live Partition Mobility (LPM)

PowerVM is a combination of hardware and software enablement.

> **Note**: PowerVM Enterprise Edition License Entitlement is included with each Power10 processor-based mid-range server. PowerVM Enterprise Edition is available as a hardware feature (#EPVV); supports up to 20 partitions per core, VIOS, multiple shared processor pools (MSPPs); and also offers LPM.

5.1.1 IBM Power Hypervisor

IBM Power processor-based servers are combined with PowerVM technology and offer the following key capabilities that can help to consolidate and simplify IT environments:

► Improve server usage and share I/O resources to reduce the total cost of ownership (TCO) and better use IT assets.
► Improve business responsiveness and operational speed by dynamically reallocating resources to applications as needed to better match changing business needs or handle unexpected changes in demand.
► Simplify IT infrastructure management by making workloads independent of hardware resources so that business-driven policies can be used to deliver resources that are based on time, cost, and service-level requirements.

Combined with features in the Power10 processor-based mid-range servers, the Power Hypervisor delivers functions that enable other system technologies, including logical partitioning technology, virtualized processors, IEEE virtual local area network (VLAN)-compatible virtual switches, virtual Small Computer Serial Interface (SCSI) adapters, virtual Fibre Channel (FC) adapters, and virtual consoles.

The Power Hypervisor is a basic component of the system's firmware and offers the following functions:

► Provides an abstraction between the physical hardware resources and the LPARs that use them.
► Enforces partition integrity by providing a security layer between LPARs.
► Controls the dispatch of virtual processors to physical processors.
► Saves and restores all processor state information during a logical processor context switch.
► Controls hardware I/O interrupt management facilities for LPARs.

► Provides VLAN channels between LPARs that help reduce the need for physical Ethernet adapters for inter-partition communication.

► Monitors the enterprise Baseboard Management Controller (eBMC) and performs a reset or reload if needed, notifying the operating system (OS) if the problem is not corrected.

The Power Hypervisor is always active, regardless of the system configuration or whether it is connected to the managed console. It requires memory to support the resource assignment of the LPARs on the server. The amount of memory that is required by the Power Hypervisor firmware varies according to several factors:

► Memory usage for hardware page tables (HPTs)
► Memory usage to support I/O devices
► Memory usage for virtualization

Memory usage for hardware page tables

Each partition on the system includes its own HPT that contributes to hypervisor memory usage. The HPT is used by the OS to translate from effective addresses to physical real addresses in the hardware. This translation from effective to real addresses allows multiple OSs to run simultaneously in their own logical address space. Whenever a virtual processor for a partition is dispatched on a physical processor, the hypervisor indicates to the hardware the location of the partition HPT that can be used when translating addresses.

The amount of memory for the HPT is based on the maximum memory size of the partition and the HPT ratio. The default HPT ratio is 1/128th (for AIX, VIOS, and Linux partitions) of the maximum memory size of the partition. AIX, VIOS, and Linux use larger page sizes (16 and 64 KB) instead of using 4 KB pages. The use of larger page sizes reduces the overall number of pages that must be tracked; therefore, the overall size of the HPT can be reduced. For example, the HPT is 2 GB for an AIX partition with a maximum memory size of 256 GB.

When defining a partition, the maximum memory size that is specified is based on the amount of memory that can be dynamically added to the dynamic logical partition (DLPAR) without changing the configuration and restarting the partition.

In addition to setting the maximum memory size, the HPT ratio can be configured. The `hpt_ratio` parameter for the `chsyscfg` Hardware Management Console (HMC) command can be issued to define the HPT ratio that is used for a partition profile. The valid values are 1:32, 1:64, 1:128, 1:256, or 1:512.

Specifying a smaller absolute ratio (1/512 is the smallest value) decreases the overall memory that is assigned to the HPT. Testing is required when changing the HPT ratio because a smaller HPT might incur more CPU consumption because the OS might need to reload the entries in the HPT more frequently. Most customers choose to use the IBM provided default values for the HPT ratios.

Memory usage for I/O devices

In support of I/O operations, the hypervisor maintains structures that are called the *translation control entities* (TCEs), which provide an information path between I/O devices and partitions. The TCEs provide the address of the I/O buffer, indications of read versus write requests, and other I/O-related attributes. Many TCEs are used per I/O device, so multiple requests can be active simultaneously to the same physical device. To provide better affinity, the TCEs are spread across multiple processor chips or drawers to improve performance while accessing the TCEs.

For physical I/O devices, the base amount of space for the TCEs is defined by the hypervisor that is based on the number of I/O devices that are supported. A system that supports high-speed adapters also can be configured to allocate more memory to improve I/O performance. Linux is the only OS that uses these extra TCEs so that the memory can be freed for use by partitions if the system uses only AIX.

Memory usage for virtualization features

Virtualization requires more memory to be allocated by the Power Hypervisor for hardware statesave areas and various virtualization technologies. For example, on Power10 processor-based systems, each processor core supports up to eight simultaneous multithreading (SMT) threads of execution, and each thread contains over 80 different registers.

The Power Hypervisor must set aside save areas for the register contents for the maximum number of virtual processors that are configured. The greater the number of physical hardware devices, the greater the number of virtual devices, the greater the amount of virtualization, and the more hypervisor memory is required. For efficient memory consumption, wanted and maximum values for various attributes (processors, memory, and virtual adapters) must be based on business needs, and not set to values that are significantly higher than actual requirements.

Predicting memory that is used by the Power Hypervisor

The IBM System Planning Tool (SPT) is a resource that can be used to estimate the amount of hypervisor memory that is required for a specific server configuration. After the SPT executable file is downloaded and installed, you can define a configuration by selecting the correct hardware platform and the installed processors and memory, and defining partitions and partition attributes. SPT can estimate the amount of memory that is assigned to the hypervisor, which helps you when you change a configuration or deploy new servers.

The Power Hypervisor provides the following types of virtual I/O adapters:

► Virtual SCSI

The Power Hypervisor provides a virtual SCSI mechanism for the virtualization of storage devices. The storage virtualization is accomplished by using two paired adapters: a virtual SCSI server adapter and a virtual SCSI customer adapter.

► Virtual Ethernet

The Power Hypervisor provides a virtual Ethernet switch function that allows partitions fast and secure communication on the same server without any need for physical interconnection or connectivity outside of the server if a Layer 2 bridge to a physical Ethernet adapter is set in one VIOS partition, also known as *Shared Ethernet Adapter* (SEA).

► Virtual FC

A virtual FC adapter is a virtual adapter that provides customer LPARs with an FC connection to a storage area network (SAN) through the VIOS partition. The VIOS partition provides the connection between the virtual FC adapters on the VIOS partition and the physical FC adapters on the managed system.

► Virtual (tty) console

Each partition must have access to a system console. Tasks such as OS installation, network setup, and various problem analysis activities require a dedicated system console. The Power Hypervisor provides the virtual console by using a virtual tty or serial adapter and a set of hypervisor calls to operate on them. Virtual tty does not require the purchase of any other features or software, such as the PowerVM Edition features.

Logical partitions

Logical partitions (LPARs) and virtualization increase the usage of system resources and add a level of configuration possibilities.

Logical partitioning is the ability to make a server run as though it were two or more independent servers. When you logically partition a server, you divide the resources on the server into subsets, which are called *LPARs*. You can install software on an LPAR, and the LPAR runs as an independent logical server with the resources that you allocated to the LPAR.

LPAR also is referred to in some documentation as a *virtual machine* (VM), which makes it look like what other hypervisors offer. However, LPARs provide a higher level of security and isolation and other features that are described in this chapter.

Processors, memory, and I/O devices can be assigned to LPARs. AIX, IBM i, Linux, and VIOS can run on LPARs. VIOS provides virtual I/O resources to other LPARs with general-purpose OSs.

> **Note:** The Power E 1050 server does not support IBM i.

LPARs share a few system attributes, such as the system serial number, system model, and processor FCs. All other system attributes can vary from one LPAR to another.

Micro-Partitioning

When you use the Micro-Partitioning technology, you can allocate fractions of processors to an LPAR. An LPAR that uses fractions of processors is also known as a *shared processor partition* or *micropartition*. Micropartitions run over a set of processors that is called a *shared processor pool* (SPP), and virtual processors are used to enable the OS manage the fractions of processing power that are assigned to the LPAR.

From an OS perspective, a virtual processor cannot be distinguished from a physical processor unless the OS is enhanced to determine the difference. Physical processors are abstracted into virtual processors that are available to partitions.

On a Power10 processor-based server, a partition can be defined with a processor capacity as small as 0.05processing units. This number represents 0.05 of a physical core. Each physical core can be shared by up to 20 shared processor partitions, and the partition's entitlement can be incremented fractionally by as little as 0.05 of the processor. The shared processor partitions are dispatched and time-sliced on the physical processors under the control of the Power Hypervisor. The shared processor partitions are created and managed by the HMC.

> **Note:** Although the Power10 processor-based mid-range server supports up to 20 shared processor partitions, the real limit depends on application workload demands in use on the server.

Processing mode

When you create an LPAR, you can assign entire processors for dedicated use, or you can assign partial processing units from an SPP. This setting defines the processing mode of the LPAR.

Dedicated mode

In dedicated mode, physical processors are assigned as a whole to partitions. The SMT feature in the Power10 processor core allows the core to run instructions from two, four, or eight independent software threads simultaneously.

Shared dedicated mode

On Power10 processor-based servers, you can configure dedicated partitions to become processor donors for idle processors that they own, which allows for the donation of spare CPU cycles from dedicated processor partitions to an SPP. The dedicated partition maintains absolute priority for dedicated CPU cycles. Enabling this feature can help increase system usage without compromising the computing power for critical workloads in a dedicated processor mode LPAR.

Shared mode

In shared mode, LPARs use virtual processors to access fractions of physical processors. Shared partitions can define any number of virtual processors (the maximum number is 20 times the number of processing units that are assigned to the partition). The Power Hypervisor dispatches virtual processors to physical processors according to the partition's processing units entitlement. One processing unit represents one physical processor's processing capacity. All partitions receive a total CPU time equal to their processing unit's entitlement. The logical processors are defined on top of virtual processors. Therefore, even with a virtual processor, the concept of a logical processor exists, and the number of logical processors depends on whether SMT is turned on or off.

5.1.2 Multiple shared processor pools

MSPPs are supported on Power10 processor-based servers. This capability allows a system administrator to create a set of micropartitions with the purpose of controlling the processor capacity that can be used from the physical SPP.

Micropartitions are created and then identified as members of the default processor pool or a user-defined SPP. The virtual processors that exist within the set of micropartitions are monitored by the Power Hypervisor. Processor capacity is managed according to user-defined attributes.

If the IBM Power server is under heavy load, each micropartition within an SPP is assured of its processor entitlement, plus any capacity that might be allocated from the reserved pool capacity if the micropartition is uncapped.

If specific micropartitions in an SPP do not use their processing capacity entitlement, the unused capacity is ceded, and other uncapped micropartitions within the same SPP can use the extra capacity according to their uncapped weighting. In this way, the entitled pool capacity of an SPP is distributed to the set of micropartitions within that SPP.

All IBM Power servers that support the MSPP capability have a minimum of one (the default) SPP and up to a maximum of 64 SPPs.

This capability helps customers reduce total cost of ownership (TCO) when the cost of software or database licenses depends on the number of assigned processor cores.

5.1.3 Virtual I/O Server

The VIOS is part of PowerVM. It is the specific appliance that allows the sharing of physical resources among LPARs to allow more efficient usage (for example, consolidation). In this case, the VIOS owns the physical I/O resources (SCSI, FC, network adapters, or optical devices) and allows customer partitions to share access to them, which minimizes and optimizes the number of physical adapters in the system.

The VIOS eliminates the requirement that every partition owns a dedicated network adapter, disk adapter, and disk drive. The VIOS supports OpenSSH for secure remote logins. It also provides a firewall for limiting access by ports, network services, and IP addresses.

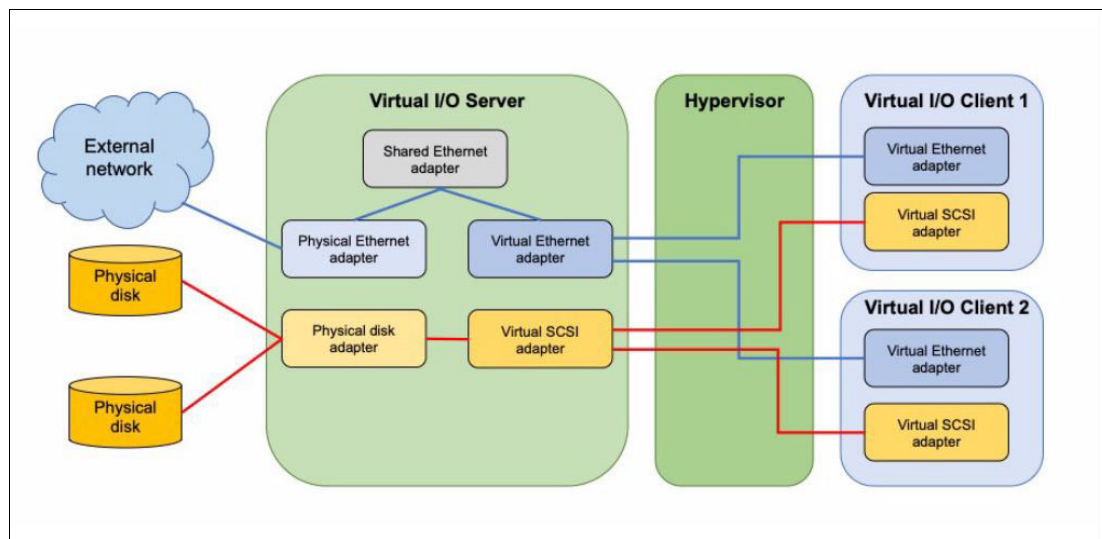Figure 5-1 shows an overview of a VIOS configuration.

Figure 5-1   Architectural view of the VIOS

It is a best practice to run dual VIOSs per physical server.

Shared Ethernet Adapter

A SEA can be used to connect a physical Ethernet network to a virtual Ethernet network. The SEA provides this access by connecting the Power Hypervisor VLANs to the VLANs on the external switches. Because the SEA processes packets at Layer 2, the original MAC address and VLAN tags of the packet are visible to other systems on the physical network. IEEE 802.1 VLAN tagging is supported.

By using the SEA, several customer partitions can share one physical adapter. You also can connect internal and external VLANs by using a physical adapter. The SEA service can be hosted only in the VIOS (not in a general-purpose AIX or Linux partition) and acts as a Layer 2 network bridge to securely transport network traffic between virtual Ethernet networks (internal) and one or more (Etherchannel) physical network adapters (external). These virtual Ethernet network adapters are defined by the Power Hypervisor on the VIOS.

Virtual SCSI

Virtual SCSI is used to view a virtualized implementation of the SCSI protocol. Virtual SCSI is based on a client/server relationship. The VIOS LPAR owns the physical I/O resources and acts as a server or in SCSI terms a target device. The client LPARs access the virtual SCSI backing storage devices that are provided by the VIOS as clients.

The virtual I/O adapters (a virtual SCSI server adapter and a virtual SCSI client adapter) are configured by using an HMC. The virtual SCSI server (target) adapter is responsible for running any SCSI commands that it receives, and it is owned by the VIOS partition. The virtual SCSI client adapter allows a client partition to access physical SCSI and SAN-attached devices and LUNs that are mapped to be used by the client partitions. The provisioning of virtual disk resources is provided by the VIOS.

N_Port ID Virtualization

N_Port ID Virtualization (NPIV) is a technology that allows multiple LPARs to access one or more external physical storage devices through the same physical FC adapter. This adapter is attached to a VIOS partition that acts only as a pass-through that manages the data transfer through the Power Hypervisor.

Each partition features one or more virtual FC adapters, each with their own pair of unique worldwide port names. This configuration enables you to connect each partition to independent physical storage on a SAN. Unlike virtual SCSI, only the client partitions see the disk.

For more information and requirements for NPIV, see *IBM PowerVM Virtualization Managing and Monitoring*, SG24-7590.

5.1.4 Live Partition Mobility

With LPM, you can move a running LPAR from one system to another one without disruption. Inactive partition mobility allows you to move a powered-off LPAR from one system to another one.

LPM provides systems management flexibility and improves system availability by avoiding the following situations:

► Planned outages for hardware upgrade or firmware maintenance.

► Unplanned downtime. With preventive failure management, if a server indicates a potential failure, you can move its LPARs to another server before the failure occurs.

For more information and requirements for LPM, see *IBM PowerVM Live Partition Mobility*, SG24-7460.

HMC 10.1.1020.0 and VIOS 3.1.3.21 or later provide the following enhancements to the LPM feature:

► Automatically choose fastest network for LPM memory transfer.
► Allow LPM when a virtual optical device is assigned to a partition.

5.1.5 Active Memory Mirroring

Active Memory Mirroring (AMM) for Hypervisor is available as an option (#EM8G) to enhance resilience by mirroring critical memory that is used by the PowerVM hypervisor so that it can continue operating in a memory failure.

A portion of available memory can be proactively partitioned such that a duplicate set can be used on non-correctable memory errors. This partition can be implemented at the granularity of DIMMs or logical memory blocks.

5.1.6 Remote Restart

Remote Restart is a high availability (HA) option for partitions. If an error occurs that causes a server outage, a partition that is configured for Remote Restart can be restarted on a different physical server. At times, it might take longer to start the server, in which case the Remote Restart function can be used for faster reprovisioning of the partition. Typically, this task can be done faster than restarting the server that stopped and then restarting the partitions. The Remote Restart function relies on technology that is like LPM, where a partition is configured with storage on a SAN that is shared (accessible) by the server that hosts the partition.

HMC 10R1 provides an enhancement to the Remote Restart feature that enables remote restart when a virtual optical device is assigned to a partition.

5.1.7 IBM Power processor modes

Although they are not virtualization features, the IBM Power processor modes are described here because they affect various virtualization features.

On IBM Power servers, partitions can be configured to run in several modes, including the following modes:

► Power8

This native mode for Power8 processors implements version 2.07 of the IBM Power instruction set architecture (ISA). For more information, see Processor compatibility mode definitions.

► Power9

This native mode for Power9 processors implements version 3.0 of the IBM Power ISA. For more information, see Processor compatibility mode definitions.

► Power10

This native mode for Power10 processors implements version 3.1 of the IBM Power ISA. For more information, see Processor compatibility mode definitions.

Figure 5-2 shows the available processor modes on a Power10 processor-based mid-range server.
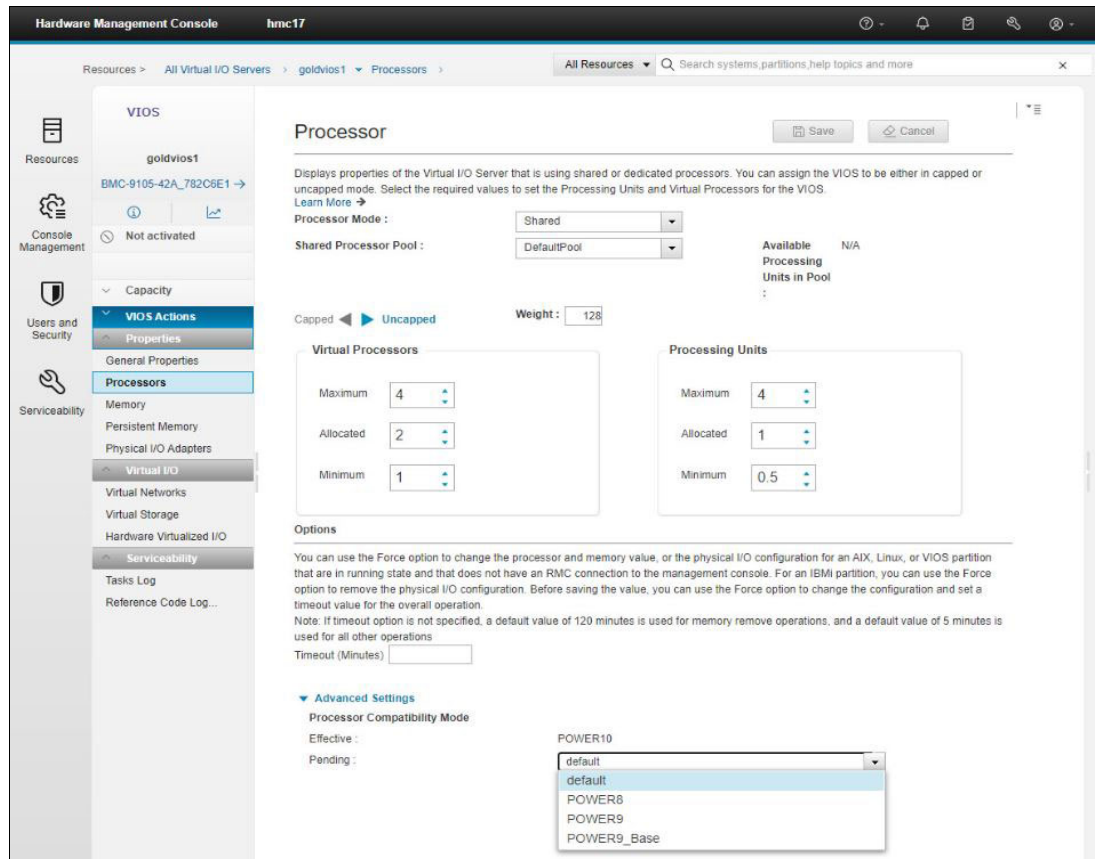
Figure 5-2   Processor modes

Processor compatibility mode is important when LPM migration is planned between different generation of servers. An LPAR that might be migrated to a machine that is managed by a processor from another generation must be activated in a specific compatibility mode.

Note: Migrating an LPAR from a Power7 processor-based server to a Power10 processor-based mid-range server by using LPM is not supported; however, the following steps can be completed to accomplish this task:

1. Migrate LPAR from a Power7 processor-based server to a Power8 or Power9 processor-based server by using LPM.

2. Migrate the LPAR from the Power8 or Power9 processor-based server to a Power10 processor-based mid-range server.

The OS running on the Power7 processor-based server must be supported on the Power10 processor-based mid-range server or must be upgraded to a supported level before completing these steps.

5.1.8 Single-root I/O virtualization

Single-root I/O virtualization (SR-IOV) is an extension to the Peripheral Component Interconnect Express (PCIe) specification that allows multiple OSs to simultaneously share a PCIe adapter with little or no runtime involvement from a hypervisor or other virtualization intermediary.

SR-IOV is PCI standard architecture that enables PCIe adapters to become self-virtualizing. It enables adapter consolidation through sharing, much like logical partitioning enables server consolidation. With an adapter capable of SR-IOV, you can assign virtual *slices* of a single physical adapter to multiple partitions through logical ports, which is done without a VIOS.

5.1.9 More information about virtualization features

The following IBM Redbooks publications provide more information about the virtualization features:

► *IBM PowerVM Best Practices*, SG24-8062
► *IBM PowerVM Virtualization Introduction and Configuration*, SG24-7940
► *IBM PowerVM Virtualization Managing and Monitoring*, SG24-7590
► *IBM Power Systems SR-IOV: Technical Overview and Introduction*, REDP-5065

5.2 IBM PowerVC overview

IBM Power Virtualization Center (PowerVC) is an advanced virtualization and cloud management offering (that is built on OpenStack) that provides simplified virtualization management and cloud deployments for IBM AIX and Linux VMs, including Red Hat OpenShift (CoreOS) running on IBM Power. PowerVC is designed to improve administrator productivity and simplify the cloud management of VMs on IBM Power servers.

By using PowerVC, you can perform the following tasks:

► Create VMs and resize the VMs CPU and memory.

► Attach disk volumes or other networks to those VMs.

► Import VMs and volumes so that they can be managed by IBM PowerVC.

► Deploy new VMs with storage and network from an image in a few minutes.

► Monitor the usage of resources in your environment.

► Take snapshots of a VM or clone a VM.

► Migrate VMs while they are running (live migration between physical servers).

► Automated Remote Restart VMs in a server failure.

► Automatically balance cloud workloads by using the Dynamic Resource Optimizer (DRO).

► Use advanced storage technologies, such as VDisk mirroring, IBM HyperSwap, and IBM Global Mirror.

► Put a server into maintenance mode with automatic distribution of LPARs to other servers and back by using LPM.

► Create a private cloud with different projects/tenants that are independent from each other but use the same resources.

► Create a self-service portal with an approval workflow.

► Meter resource usage as a basis for cost allocation.

- ► Improve resource usage to reduce capital expenses and power consumption.
- ► Increase agility and execution to respond quickly to changing business requirements.
- ► Increase IT productivity and responsiveness.
- ► Simplify IBM Power virtualization management.
- ► Accelerate repeatable, error-free virtualization deployments.

IBM PowerVC can manage AIX and Linux based VMs that are running under PowerVM virtualization that are connected to an HMC or that use NovaLink. This release supports the scale-out and the enterprise IBM Power servers that are built on IBM Power8, IBM Power9, and Power10 servers.

> **Note:** The Power E1050 server is supported by PowerVC 2.0.3 or later. If an additional fix is needed, see IBM Fix Central.

5.2.1 IBM PowerVC functions and advantages

When more than 70% of IT budgets are spent on operations and maintenance, IT customers legitimately expect vendors to focus their new development efforts to reduce IT costs and foster innovation within IT departments.

IBM PowerVC gives IBM Power customers the following advantages:

- ► It is deeply integrated with IBM Power.
- ► It provides virtualization management tools.
- ► It eases the integration of servers that are managed by PowerVM in automated IT environments, such as clouds.
- ► It is a building block of IBM Infrastructure as a Service (IaaS), which is based on IBM Power.
- ► PowerVC integrated with other cloud management tools, such as Ansible, Terraform, or Red Hat OpenShift, can be integrated into orchestration tools, such as the IBM Cloud Automation Manager, VMware vRealize, or SAP Landscape Management (LaMa).
- ► PowerVC also provides an easy exchange of VM images between private and public clouds by using the integration of IBM Cloud Object Storage into PowerVC.

IBM PowerVC is an addition to the PowerVM set of enterprise virtualization technologies that provide virtualization management. It is based on open standards and integrates server management with storage and network management.

Because IBM PowerVC is based on the OpenStack initiative, IBM Power can be managed by tools that are compatible with OpenStack standards. When a system is controlled by IBM PowerVC, it can be managed in one of three ways:

- ► By a system administrator by using the IBM PowerVC GUI
- ► By a system administrator that uses scripts that contain the IBM PowerVC Representational State Transfer (REST) application programming interfaces (APIs)
- ► By higher-level tools that call IBM PowerVC by using standard OpenStack API

The following PowerVC offerings are positioned within the available solutions for IBM Power cloud:

- ► IBM PowerVC: Advanced Virtualization Management
- ► IBM PowerVC for Private Cloud: Basic Cloud

- ► IBM Cloud Automation Manager: Advanced Cloud
- ► VMware vRealize: Advanced Cloud

PowerVC provides a systems management product that enterprise customers require to manage effectively the advanced features that are offered by IBM premium hardware. It reduces resource use and manages workloads for performance and availability.

For more information about PowerVC, see *IBM PowerVC Version 2.0 Introduction and Configuration*, SG24-8477.

5.3 Digital transformation and IT modernization

Digital transformation is the process of using digital technologies to create business processes or modify existing ones.

It also means changing the culture and experiences to meet the changing needs of the business and the market. This reimaging of business in the digital age is digital transformation.

For IBM, digital transformation takes a customer-centric and digital-centric approach to all aspects of a business, from business models to customer experiences, processes, and operations. It uses artificial intelligence (AI), automation, hybrid cloud, and other digital technologies to leverage data and drive intelligent workflows, faster and smarter decision-making, and a real-time response to market disruptions. Ultimately, it changes customer expectations and creates business opportunities.

To date, one of the main technologies that is recognized to enable the digital transformation and the path to application modernization is Red Hat OpenShift. There are advantages to associate Red Hat OpenShift with the IBM Power processor-based platform.

5.3.1 Application and services modernization

Modernizing applications and services (both legacy and cloud-native) on containers that are deployed through a platform that is designed with Kubernetes orchestration at its core is arguably the most forward-looking strategy any business will benefit from. Unsurprisingly, it is also at the heart of the IBM and Red Hat strategy for the hybrid multicloud reality of tomorrow's digital landscape.

Red Hat OpenShift is a container orchestration platform that is based on Kubernetes that helps develop containerized applications with open-source technology. It facilitates management and deployments in hybrid and multicloud environments by using full-stack automated operations.

Containers are key elements of the IT transformation and journey toward modernization. Containers are software executable units in which the application code is packaged, along with its libraries and dependencies, in common ways so that it can be run anywhere, both on desktop and on any type of server or on the cloud. To do this task, containers take advantage of a form of OS virtualization in which OS functions are used effectively both to isolate processes and to control the amount of CPU, memory, and disk that those processes have access to.

Containers are small, fast, and portable because unlike a VM, they do not need to include a guest OS in every instance, and can instead simply leverage the functions and resources of the host OS.

Containers first appeared decades ago with releases such as FreeBSD jails and AIX Workload Partitions (WPARs), but most modern developers remember 2013 as the beginning of the modern container era with the introduction of Docker.

One way to better understand a container is to understand how it differs from a traditional VM. In traditional virtualization (both on-premises and in the cloud), a hypervisor is used to virtualize the physical hardware. Each VM contains a guest OS, a virtual copy of the hardware that the OS requires to run, with an application and its associated libraries and dependencies.

Instead of virtualizing the underlying hardware, containers virtualize the OS (usually Linux) so that each individual container includes only the application and its libraries and dependencies. The absence of the guest OS is the reason why containers are so light, fast, and portable.

With IBM Power, you enable the possibility of having a high container ratio per core with multiple CPU threads and co-localized cloud-native applications and applications that are related to AIX, which use API connections to business-critical data for higher bandwidth and lower latency than other technologies. Only with IBM Power do you have a flexible and efficient usage of resources, manage peaks, and support both traditional and modern workloads with the eventual support of Capacity on Demand (CoD) or sharing processor pools.

The Red Hat Ansible product complements the solution with best-in-class automation.

5.3.2 System automation with Ansible

Enterprises can spend much precious administrative time performing repetitive tasks and using manual processes. Tasks such as updating, patching, compliance checks, provisioning new VMs or LPARs, and ensuring that the correct security updates are in place, are taking time away from more valuable business activities.

The ability to automate by using Ansible returns valuable time to the system administrators.

Red Hat Ansible Automation Platform for IBM Power is fully enabled, so enterprises can automate several tasks within AIX and Linux that include deploying applications. Ansible also can be combined with HMC, PowerVC, and Power Virtual Server to provision infrastructure anywhere you need, including cloud solutions from other IBM Business Partners or third-party providers based on IBM Power processor-based servers.

A first task after the initial installation or setup of a new LPAR is to ensure that the correct patches are installed. Also, extra software (whether it is open-source software, independent software vendor (ISV) software, or perhaps the business' own enterprise software) must be installed. Ansible features a set of capabilities to roll out new software, which makes it popular in Continuous Delivery/Continuous Integration (CD/CI) environments. Orchestration and integration of automation with security products represent other ways in which Ansible can be applied within the data center.

Despite the wide adoption of AIX in many different business sectors by different types of customers, Ansible can help introduce IBM Power processor-based technology to customers who believe that AIX skills are a rare commodity that is difficult to find in the marketplace but want to take advantage of all the features of the hardware platform. The Ansible experience is identical across IBM Power or x86 processor-based technology, and the same steps can be repeated in IBM Cloud.

AIX skilled customers also can benefit from the extreme automation solutions that are provided by Ansible.

The IBM Power processor-based architecture features unique advantages over commodity server platforms such as x86 because the engineering teams that are working on the processor, system boards, virtualization. and management appliances collaborate closely to ensure an integrated stack that works seamlessly. This approach is in stark contrast to the multivendor x86 processor-based technology approach, in which the processor, server, management, and virtualization must be purchased from different (and sometimes competing) vendors.

The IBM Power stack engineering teams work closely to deliver an enterprise server platform, which results in an IT architecture with industry-leading performance, scalability, and security (see Figure 5-3).
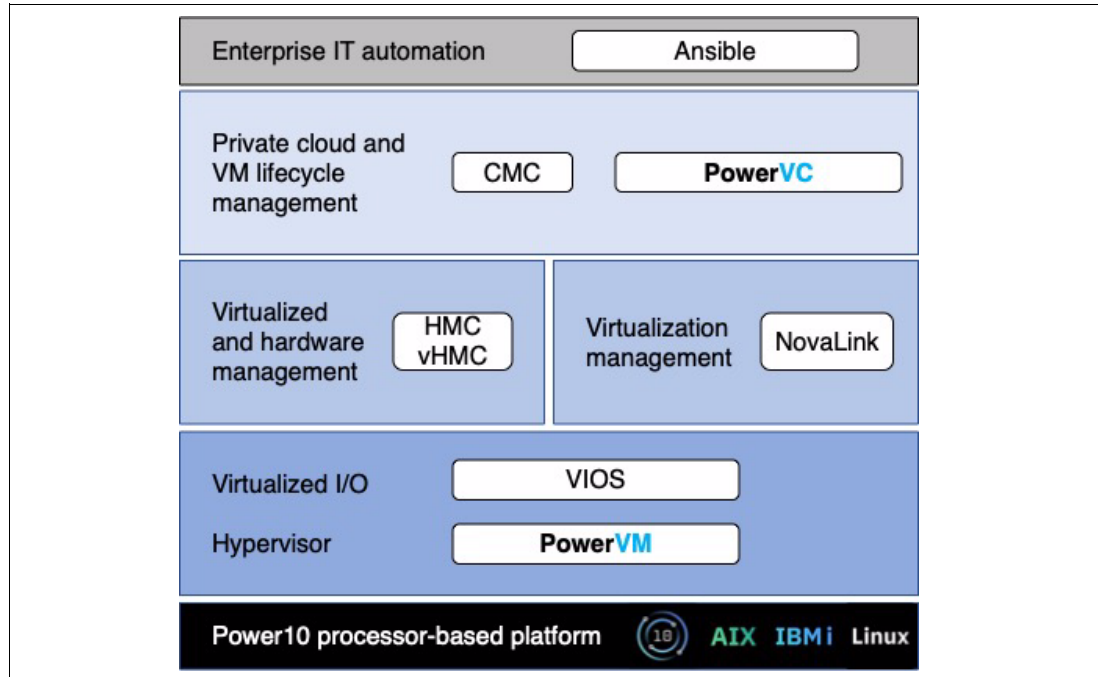
Figure 5-3   IBM Power stack

Every layer in the IBM Power stack is optimized to make the Power10 processor-based technology the platform of choice for mission-critical enterprise workloads. This stack includes the Ansible Automation Platform, which is described next.

Ansible Automation Platform

Ansible Automation Platform integrates with IBM Power processor-based technology, which is included in the Certified Integrations section of the Red Hat Ansible website.

The various Ansible collections for IBM Power processor-based technology, which (at the time of writing) were downloaded more than 25,000 times by customers, are now included in the Red Hat Ansible Automation Platform. As a result, these modules are covered by the Red Hat 24x7 enterprise support team, which collaborates with the respective IBM Power processor-based technology development teams.

IBM Power in the Ansible ecosystem

A series of Ansible collections is available for the IBM Power processor-based platform. A *collection* is a group of modules, playbooks, and roles. By embracing Ansible in the IBM Power community, the AIX and IBM i communities have the right set of modules that are available. Some examples are development of tools to manage AIX logical volumes, which put module interfaces over the installation key command, or managing the AIX init tab entries.

> **Note:** The Power E 1050 server does not support IBM i.

Our OS teams develop modules that are sent to the Ansible open-source community (named Ansible Galaxy). Every developer can post any object that can be a candidate for a collection in the open Ansible Galaxy community and possibly certified to be supported by IBM with a subscription to Red Hat Ansible Automation Platform (see Figure 5-4).
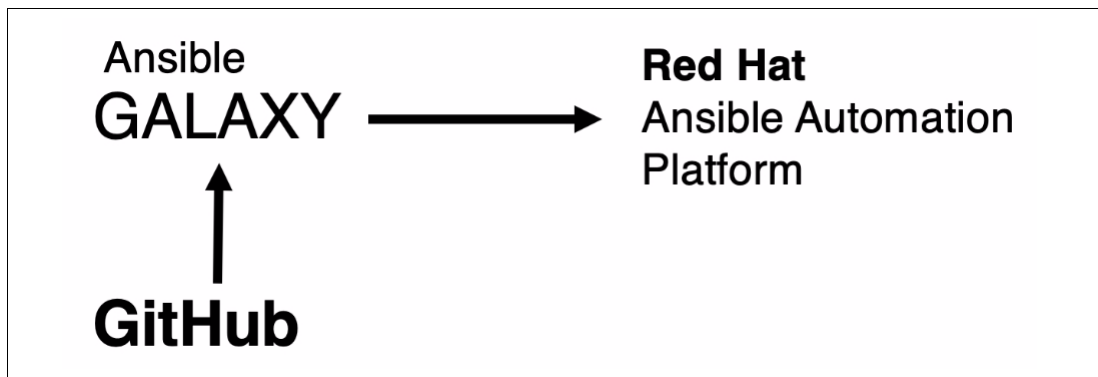
Figure 5-4   IBM Power content in the Ansible ecosystem

Ansible modules for AIX

The IBM Power AIX collection, as part of the broader offering of Ansible Content for IBM Power, is available from Ansible Galaxy and has community support.

The collection includes modules and sample playbooks that help to automate tasks. You can find it at this web page.

Ansible modules for HMC

IBM Power HMC collection provides modules that can be used to manage configurations and deployments of HMC systems. The collection content helps to include workloads on IBM Power processor-based platforms as part of an enterprise automation strategy through the Ansible ecosystem.

For more information about this collection, see this web page.

Ansible modules for VIOS

The IBM Power VIOS collection provides modules that can be used to manage configurations and deployments of IBM Power VIOS systems. The collection content helps to include workloads on IBM Power processor-based platforms as part of an enterprise automation strategy through the Ansible system.

For more information, see this web page.

Ansible modules for Oracle on AIX

This repository contains a collection that can be used to install an Oracle single instance database 19c on AIX OS and that creates a test database on an AIX file system and on Oracle ASM. This collection automates the Oracle 19c database installation and creation steps.

For more information, see this web page.

Ansible modules for Oracle RAC on AIX

This collection provides modules that can be used to install Oracle 19c RAC on AIX OS. This collection also automates the Oracle 19c RAC installation and creation steps.

For more information, see this web page.

Ansible modules for SAP on AIX

The Ansible Content for SAP Software on AIX provides roles to automate administrator tasks for SAP installations on AIX, such installing the SAP Host Agent, starting and stopping SAP instances, and upgrading the SAP kernel.

For more information, see this web page.

5.4 Protecting trust from core to cloud

The IT industry has long relied on perimeter security strategies to protect its most valuable assets, such as user data and intellectual property. Using firewalls and other network-based tools to inspect and validate users entering and leaving the network is no longer enough because digital transformation and the shift to a hybrid cloud infrastructure are changing the way industries do business.

Many organizations also are adapting their business models, and many organizations have thousands of people connecting from home computers outside the control of an IT department. Users, data, and resources are scattered worldwide, making it difficult to connect them quickly and securely. Without a traditional local infrastructure for security, employees' homes are more vulnerable to compromise, putting the business at risk.

Many companies are operating with a set of security solutions and tools even without them being fully integrated. As a result, security teams spend more time on manual tasks. They lack the context and information that are needed to effectively reduce their organization's attack surface. Rising data breaches and rising global regulations have made securing networks difficult.

Applications, users, and devices need fast and secure access to data, so much so that an entire industry of security tools and architectures was created to protect them.

Although enforcing a data encryption policy is an effective way to minimize the risk of a data breach that, in turn, minimizes costs, only a few enterprises at the worldwide level have an encryption strategy that is applied consistently across the entire organization, largely because it adds complexity, costs, and negatively affects performance, which means missed service-level agreements (SLAs) to the business.

The rapidly evolving cyberthreat landscape requires focus on cyber-resilience. Persistent and end-to-end security is the only way to reduce exposure to threats.

Prevention is the best protection

Since prevention is the best protection, Power10 processor-based servers provide industry leading isolation, and integrity that help prevent ransomware from being installed. The following features and implementation help customers protect their businesses:

► Host/firmware secure and trusted boot.

► Guest OS secure boot (AIX now, and Linux upcoming).

► Built-in OS runtime integrity: AIX Trusted Execution, and Linux IMA.

► Most secure multi-tenant environment with orders of magnitude lower Common Vulnerabilities and Exposures (CVEs) versus x86 hypervisors.

► Orders of magnitude lower CVEs for AIX.

► Simplified patching with PowerSC.

► Multi-factor authentication (MFA) with PowerSC MFA.

Early detection is critical

Integrated security and compliance management with PowerSC makes it harder to misconfigure and easier to detect anomalies. Offerings such as IBM Security QRadar® enhance inherent security with early anomaly detection.

Fast and efficient recovery

It is easier to deploy resiliency strategies with IBM PowerHA and IBM Storage Safeguarded Copy, and by leveraging IBM Storage and security services for fast detection and automated recovery of affected data.

5.4.1 Power10 processor-based technology integrated security ecosystem

IBM Power processor-based platforms always offer the most secure and reliable servers in their class. Introducing Power10 processor-based technology in 2021, IBM further extended the industry-leading security and reliability of the IBM Power processor-based platform with a focus on protecting applications and data across all the hybrid cloud environments. IBM also introduced significant innovations along the following major dimensions:

► *Advanced data protection* that offers simple to use and efficient capabilities to protect sensitive data through mechanisms such as encryption and MFA.

► *Platform security* ensures that the server is hardened against tampering, continuously protects its integrity, and ensures strong isolation among multi-tenant workloads. Without strong platform security, all other system security measures are at risk.

► *Security innovation for modern threats* helps you stay ahead of new types of cybersecurity threats by using emerging technologies.

► *Integrated security management* addresses the key challenge of ensuring correct configuration of the many security features across the stack, monitoring them, and reacting if unexpected changes are detected.

The Power10 processor-based servers are enhanced to simplify and integrate security management across the stack, which reduces the likelihood of administrator errors.

In the Power10 processor-based scale-out servers, all data is protected by a greatly simplified end-to-end encryption that extends across the hybrid cloud without detectable performance impact and prepares for future cyberthreats.

Power10 processor-core technology features built-in security integration:

► Stay ahead of current and future data threats with better cryptographic performance and support for quantum-safe cryptography and fully homomorphic encryption (FHE).

► Enhance the security of applications with more hardened defense against return-oriented programming (ROP) attacks.

► Simplified single-interface hybrid cloud security management without any required setup.

► Protect your applications and data with the most secure VM isolation in the industry with orders of magnitude lower CVEs than hypervisors that are related to x86 processor-based servers.

Also, workloads on the Power10 processor-based scale-out servers benefit from cryptographic algorithm acceleration, which allows algorithms, such as Advanced Encryption Standard (AES), SHA2, and SHA3 to run significantly faster than Power9 processor-based servers on a per core basis. This performance acceleration allows features, such as AIX Logical Volume Encryption, to be enabled with low performance overhead.

5.4.2 Crypto engines and transparent memory encryption

Power10 processor-based technology is engineered to achieve faster encryption performance compared to IBM Power9 processor-based servers. Power10 processor-based scale-out servers are updated for today's most demanding standards and anticipated future cryptographic standards, such as post-quantum and FHE, and brings new enhancements to container security.

Transparent memory encryption is designed to simplify encryption and support end-to-end security without affecting performance by using hardware features for a seamless user experience. The protection that is introduced in all layers of an infrastructure is shown in Figure 5-5.
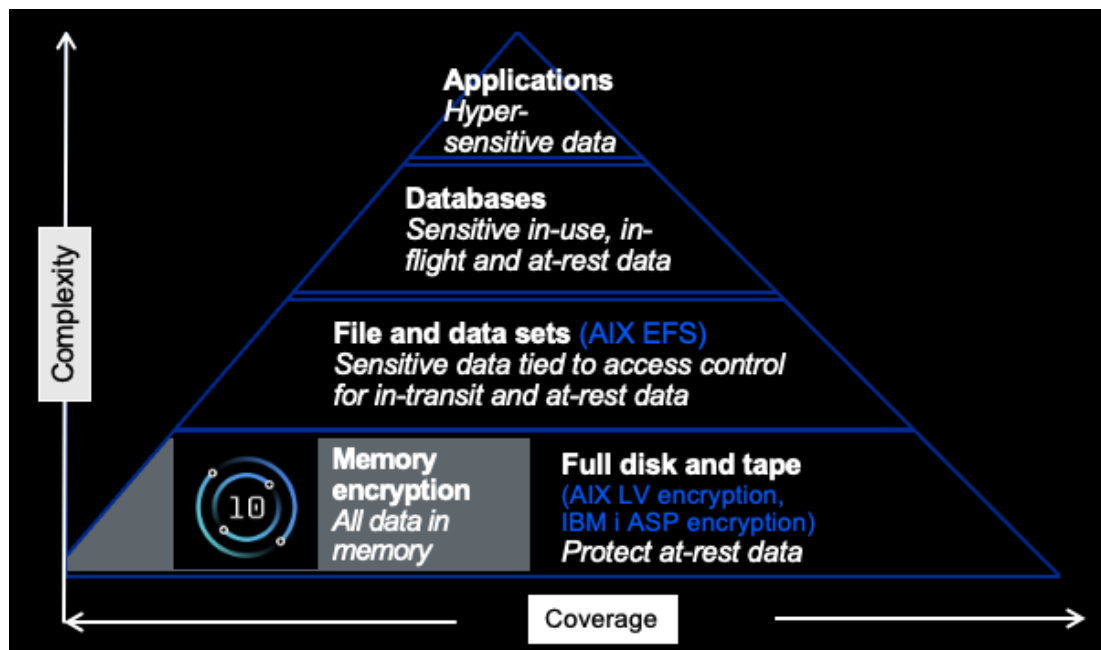

Figure 5-5   Protecting data in memory with transparent memory encryption

5.4.3 Quantum-safe cryptography support

To be prepared for the quantum era, the Power10 processor-based server is built to efficiently support future cryptography, such as quantum-safe cryptography and FHE. The software libraries for these solutions are optimized for the Power10 processor-based chip ISA, and they are or will be available in the respective open source communities.

Quantum-safe cryptography refers to the efforts to identify algorithms that are resistant to attacks by classical and quantum computers in preparation for the time when large-scale quantum computers are built.

Homomorphic encryption refers to encryption techniques that permit systems to perform computations on encrypted data without decrypting the data first. The software libraries for these solutions are optimized for the IBM Power processor-based chip ISA.

5.4.4 IBM PCIe3 Crypto Coprocessor BSC-Gen3 4769

IBM PCIe3 Crypto Coprocessor BSC-Gen3 4769 (#EJ37) provides hardware security protection for the most sensitive secrets. It supports IBM AIX, IBM i without VIOS, or supported distributions of Linux OSs.

> **Note:** The Power E 1050 server does not support IBM i.

IBM PCIe3 Crypto Coprocessor BSC-Gen3 4769 provides a comprehensive set of cryptographic functions, including the common AES, Triple Data Encryption Standard (TDES), Rivest-Shamir-Adleman (RSA), and error correction code (ECC) functions for data confidentiality and data integrity support. In addition, IBM Common Cryptographic Architecture (CCA) features extensive functions for key management and many functions of special interest to the banking and finance industry.

The co-processor holds a security-enabled subsystem module and batteries for backup power. The hardened encapsulated subsystem contains two sets of two 32-bit PowerPC 476FP reduced-instruction-set-computer (RISC) processors running in lockstep with cross-checking to detect soft errors in the hardware.

It also contains a separate service processor that is used to manage:
- ▶ Self-test and firmware updates.
- ▶ RAM, flash memory, and battery-powered memory.
- ▶ Secure time-of-day.
- ▶ Cryptographic quality random number generator.
- ▶ AES, Data Encryption Standard (DES), TDES, Hash-Based Message Authentication Code (HMAC), Cipher-based Message Authentication Code (CMAC), MD5, and multiple Secure Hash Algorithm (SHA) hashing methods.
- ▶ Modular-exponentiation hardware, such as RSA and ECC.
- ▶ Full-duplex direct memory access (DMA) communications.

A security-enabled code-loading arrangement allows control program and application program loading and refreshes after co-processor installation in your server.

IBM offers an embedded subsystem control program and a cryptographic API that implements the CCA Support Program that can be accessed from the internet at no charge to the user.

The 4769 PCIe Cryptographic Coprocessor is designed to deliver the following functions:

► X.509 certificate services support.

► American National Standards Institute (ANSI) X9 TR34-2019 key exchange services that use the public key infrastructure (PKI).

► Elliptic Curve Digital Signature Algorithm (ECDSA) secp256k1.

► CRYSTALS-Dilithium, a quantum-safe algorithm for digital signature generation and verification.

► An RSA algorithm for digital signature generation and verification with keys up to 4096 bits.

► High-throughput SHA, MD5 message digest algorithm, HMAC, CMAC, DES, TDES, and AES-based encryption for data integrity assurance and confidentiality, including AES Key Wrap (AESKW) that conforms to ANSI X9.102.

► Elliptic-curve cryptography for digital signature and key agreement.

► Support for smart card applications and personal identification number (PIN) processing.

► Secure time-of-day.

► Visa Data Secure Platform (DSP) point-to-point encryption (P2PE) with standard Visa format-preserving encryption (FPE) and format-preserving, Feistel-based Format Preserving Encryption (FF1, FF2, and FF2.1). Format Preserving Counter Mode (FPCM) is defined in ANSI x9.24 Part 2.

5.4.5 IBM PowerSC support

Power10 processor-based scale-out servers benefit from the integrated security management capabilities that are offered by:

► IBM PowerSC.

► The IBM Power software portfolio for managing security and compliance on every IBM Power processor-based platform that is running AIX.

► The supported distributions and versions of Linux.

► PowerSC is introducing more features to help customers manage security end-to-end across the stack to stay ahead of various threats. Specifically, PowerSC 2.0 adds support for Endpoint Detection and Response (EDR), host-based intrusion detection, block listing, and Linux.

Security features are beneficial only if they can be easily and accurately managed. Power10 processor-based scale-out servers benefit from the integrated security management capabilities that are offered by IBM PowerSC, and the IBM Power software portfolio for managing security and compliance on IBM Power processor-based platforms (AIX and Linux on IBM Power).

PowerSC is sitting on top of the Power solution stack, and it provides features such as compliance automation to help with various industry standards, real-time file integrity monitoring, reporting to support security audits, patch management, and trusted logging.

By providing all these capabilities through a web-based user interface (UI), PowerSC simplifies management of security and compliance.

PowerSC MFA provides additional assurance that only authorized people access the environments by requiring at least one additional authentication factor to prove you are the person that you say that you are. MFA is included in PowerSC 2.0. Because stolen or guessed passwords are still one of the most common ways for hackers to get into systems, having an MFA solution in place allows you to prevent a high percentage of potential breaches.

PowerSC 2.0 also includes EDR. It provides:

► Intrusion Detection and Prevention (IDP)

► Log Inspection and Analysis

► Anomaly detection, correlation, and incident response

► Response triggers

► Event context and filtering

5.4.6 Secure Boot and Trusted Boot

IBM Power servers provide a highly secure server platform. IBM Power10 processor-based hardware and firmware includes PowerVM features to provide a more secure platform for cloud deployment.

The key PowerVM features include:

► A secure initial program load (IPL) process or the Secure Boot feature allows only correctly signed firmware components to run on the system processors. Each component of the firmware stack, including hostboot, the Power Hypervisor, and partition firmware (PFW), is signed by the platform manufacturer and verified as part of the IPL process.

► A framework to support remote attestation of the system firmware stack through a hardware trusted platform module (TPM).

The terms *Secure Boot* and *Trusted Boot* have specific connotations. The terms are used as distinct yet complementary concepts.

Secure Boot

The Secure Boot feature protects system integrity by using digital signatures to perform a hardware-protected verification of all firmware components. It also distinguishes between the host system trust domain and the Flexible Service Processor (FSP) trust domain by controlling service processor and service interface access to sensitive system memory regions.

Trusted Boot

The Trusted Boot feature creates cryptographically strong and protected platform measurements that prove that particular firmware components have run on the system. You can assess the measurements by using trusted protocols to determine the state of the system and use that information for security decisions.

5.4.7 Enhanced CPU: baseboard management controller isolation

Separating CPU and service processor trust domains in Power10 processor-based scale-out servers improves protection from external attacks.

Power10 processor-based technology introduces innovations to address emerging threats, specifically with additional features and enhancements to defend against application domain vulnerabilities such as ROP attacks (a security exploit technique that is used by attackers to run code on a target system). This capability uses a new in-core hardware architecture that imparts minimal performance overhead (approximately only 1 - 2% for some sample workloads tested).

The baseboard management controller (BMC) chip is connected to the two network interface cards through Network Connectivity Status Indicator (NCSI) (to support the connection to HMCs) and also have a PCIe x1 connection that connects to the backplane. This connection is used by PowerVM for partition management traffic, but it cannot be used for Guest LPAR traffic. A Guest LPAR needs its physical or virtual network interface PCIe card (or cards) for an external connection.

Hardware assist is necessary to avoid tampering with the stack. The IBM Power platform added four instructions (hashst, hashchk, hashstp, and hashchkp) to handle ROP in IBM Power ISA 3.1B.

Abbreviations and acronyms

4S	4-socket		**DF**	decimal floating-point
ACE	access control entry		**DLPAR**	dynamic logical partition
ACF	Access Control File		**DMA**	direct memory access
AES	Advanced Encryption Standard		**DME**	dense math engine
AESKW	AES Key Wrap		**DPM**	dynamic performance mode
AI	artificial intelligence		**DRAM**	dynamic RAM
AME	Active Memory Expansion		**DRO**	Dynamic Resource Optimizer
AMM	Active Memory Mirroring		**DSP**	Data Secure Platform
ANSI	American National Standards Institute		**EA**	effective address
AOC	Active Optical Cable		**eBMC**	enterprise Baseboard Management Controller
API	application programming interface		**ECC**	error correction code
ARI	analog rack interface		**ECDSA**	Elliptic Curve Digital Signature Algorithm
ASMI	Advanced System Management Interface		**ED/FI**	Error Detection and Fault Isolation
BMC	baseboard management controller		**EDR**	Endpoint Detection and Response
BSC	Blind-Swap Cassette		**EEH**	enhanced error handling
CAC	Common Access Card		**EMEA**	Europe, Middle East, and Asia
CCA	Common Cryptographic Architecture		**ERAT**	effective-to-real address translation
CEC	central electronic complex		**ESA**	Electronic Service Agent
CLI	command-line interface		**eSCM**	entry single-chip module
CMAC	Cipher-based Message Authentication Code		**ESM**	Enclosure Services Manager
CMOS	complimentary metal-oxide-semiconductor		**FC**	Fibre Channel
			FFDC	First Failure Data Capture
CoD	Capacity on Demand		**FHE**	fully homomorphic encryption
CRC	cyclic redundancy check		**FIFO**	first in - first out
CRU	customer-replaceable unit		**FIR**	Fault Isolation Register
CSP	Client Support Portal		**FPCM**	Format Preserving Counter Mode
CSU	customer setup		**FPE**	format-preserving encryption
CUoD	Capacity Upgrade on Demand		**FRU**	field-replaceable unit
CVEs	Common Vulnerabilities and Exposures		**FSP**	Flexible Service Processor
			GA	General Availability
DCM	dual-chip module		**GCC**	GNU Compiler Collection
DDIMM	Differential Dual Inline Memory Module		**GTMS**	Global Total Microcode Support
			GTps	giga-transfers per second
DDR4	Double Data Rate 4		**HA**	high availability or highly available
DDW	Dynamic DMA Window		**HDDW**	Huge Dynamic DMA Window
DES	Data Encryption Standard		**HLFH**	Half Length Full Height
DEXCR	Dynamic Execution Control Register		**HLHH**	Half Length Half Height
			HMAC	Hash-Based Message Authentication Code

HMC	Hardware Management Console	**NX**	nest accelerator
HNV	Hybrid Network Virtualization	**OIDC**	OpenID Connect
HPT	hardware page table	**OMI**	Open Memory Interface
IaaS	infrastructure as a service	**OS**	operating system
IBM	International Business Machines Corporation	**OSHA**	Occupational Safety and Health Administration
IBM CMC	IBM Cloud Management Console	**P2PE**	point-to-point encryption
IBM ESS	IBM Entitled Systems Support	**PCIe**	Peripheral Component Interconnect Express
IBM SSR	IBM System Services Representative	**PCR**	Processor Compatibility Register
IDP	Intrusion Detection and Prevention	**PDU**	power distribution unit
iPDU	intelligent power distribution unit	**PEC**	PCI Express controller
IPL	initial program load	**PEP2**	Power Enterprise Pools 2.0
IPMI	Intelligent Platform Management Interface	**PFW**	partition firmware
		PHB	PCI host bridge
IPS	idle power saver mode	**PIN**	personal identification number
ISA	instruction set architecture	**PIR**	Processor Instruction Retry
ISV	independent software vendor	**PIV**	Personal Identification Verification
KVM	keyboard/video/mouse	**PKI**	public key infrastructure
LA	Latin America	**PMIC**	Power Management Integrated Circuit
LaMa	Landscape Management		
LAN	local area network	**PowerAXON**	Power/A-bus/X-bus/OpenCAPI/Net working
LDAP	lightweight directory access protocol		
		PSU	power supply unit
LE	Little Endian	**QP**	quad-precision floating-point
LMB	local memory bus	**RAS**	reliability, availability, and serviceability
LPAR	logical partition		
LPM	Live Partition Mobility	**RDHX**	Rear Door Heat Exchanger
MCU	memory controller unit	**REST**	Representational State Transfer
MES	Miscellaneous Equipment Specification	**RHEL**	Red Hat Enterprise Linux
		RISC	reduced-instruction-set-computer
MFA	Multi-Factor Authentication	**ROP**	return-oriented programming
MMA	Matrix Math Accelerator	**RSA**	Rivest-Shamir-Adleman
MPM	maximum performance mode	**rsh**	Remote Shell
MSP	Mover Server Partition	**SAN**	storage area network
MSPP	multiple shared processor pools	**SAS**	serial-attached SCSI
MTBF	mean time between failures	**SCM**	single-chip module
MTFR	MPM typical frequency range	**SCSI**	Small Computer Serial Interface
MTM	machine type and model	**SDS**	software-defined storage
NA	North America	**SEA**	Shared Ethernet Adapter
NCSI	Network Connectivity Status Indicator	**SEL**	system event log
		SFP	Service Focal Point
NPIV	N_Port ID Virtualization	**SHA**	Secure Hash Algorithm
NTP	Network Time Protocol	**SIMD**	single instructions multiple data
NUCA	non-uniform cache access	**SIU**	SMP interconnect unit
NVMe	Non-Volatile Memory Express		

SLA	service-level agreement
SMP	symmetric multiprocessing
SMT	simultaneous multithreading
SP	service pack
SPCN	system power control network
SPP	shared processor pool
SPT	System Planning Tool
SR-IOV	single-root I/O virtualization
SRDF	Symmetrix Remote Data Facility
SRR	Simplified Remote Restart
SSD	solid state drive
SSH	Secure Shell
SSIC	IBM System Storage Interoperation Center
SSO	single sign-on
ST	single-threaded
SUE	Special Uncorrectable Error
SUMA	Service Update Management Assistant
SVC	SAN Volume Controller
SWMA	Software Maintenance Agreement
TCE	translation control entity
TCO	total cost of ownership
TIM	thermal interface material
TL	technology level
TLB	translation lookaside buffer
TNC	Trusted Network Connect
TPM	trusted platform module
UAK	Update Access Key
UI	user interface
USB	Universal Serial Bus
vHMC	virtual HMC
VIOS	Virtual I/O Server
VLAN	virtual local area network
VM	virtual machine
VMI	Virtualization Management Interface
vNIC	virtual Network Interface Controller
VSU	vector scalar unit
VSX	Vector Scalar eXtension
WORM	write once, read many
WPAR	Workload Partition
YB	yottabyte

Related publications

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

IBM Redbooks

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

► *IBM PowerAI: Deep Learning Unleashed on IBM Power Systems Servers*, SG24-8409

► *IBM Power E1080 Technical Overview and Introduction*, REDP-5649

► *IBM Power S1014, S1022s, S1022, and S1024 Technical Overview and Introduction*, REDP-5675

► *IBM Power System AC922 Technical Overview and Introduction*, REDP-5494

► *IBM Power System E950: Technical Overview and Introduction*, REDP-5509

► *IBM Power System E980: Technical Overview and Introduction*, REDP-5510

► *IBM Power System L922 Technical Overview and Introduction*, REDP-5496

► *IBM Power System S822LC for High Performance Computing Introduction and Technical Overview*, REDP-5405

► *IBM Power Systems H922 and H924 Technical Overview and Introduction*, REDP-5498

► *IBM Power Systems LC921 and LC922: Technical Overview and Introduction*, REDP-5495

► *IBM Power Systems S922, S914, and S924 Technical Overview and Introduction Featuring PCIe Gen 4 Technology, REDP-5595*

► *IBM PowerVC Version 2.0 Introduction and Configuration*, SG24-8477

► *IBM PowerVM Best Practices*, SG24-8062

► *IBM PowerVM Virtualization Introduction and Configuration*, SG24-7940

► *IBM PowerVM Virtualization Managing and Monitoring*, SG24-7590

► *SAP HANA Data Management and Performance on IBM Power Systems*, REDP-5570

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

167

Online resources

The following websites are also relevant as further information sources:

► IBM Documentation:

https://www.ibm.com/docs/en

► IBM Fix Central:

http://www.ibm.com/support/fixcentral/

► IBM Portal for OpenPower - Power9 Monza Module:

https://www.ibm.com/systems/power/openpower/tgcmDocumentRepository.xhtml?aliasId=POWER9_Monza

► IBM Power:

https://www.ibm.com/it-infrastructure/power

► IBM Power10 documentation:

https://www.ibm.com/docs/en/power10

► IBM Storage:

https://www.ibm.com/it-infrastructure/storage

► IBM Systems Energy Estimator:

https://ibm.biz/energy-estimator

► IBM System Planning Tool:

https://www.ibm.com/support/pages/ibm-system-planning-tool-power-processor-based-systems-0

► OpenCAPI:

http://opencapi.org/technical/use-cases/

► OpenPOWER Foundation:

https://openpowerfoundation.org/

► Power Systems Capacity on Demand:

http://www.ibm.com/systems/power/hardware/cod/

► Support for IBM Systems:

https://www.ibm.com/mysupport/s/?language=en_US

Help from IBM

IBM Support and downloads

ibm.com/support

IBM Global Services

ibm.com/services

Get connected

ibm.com/redbooks