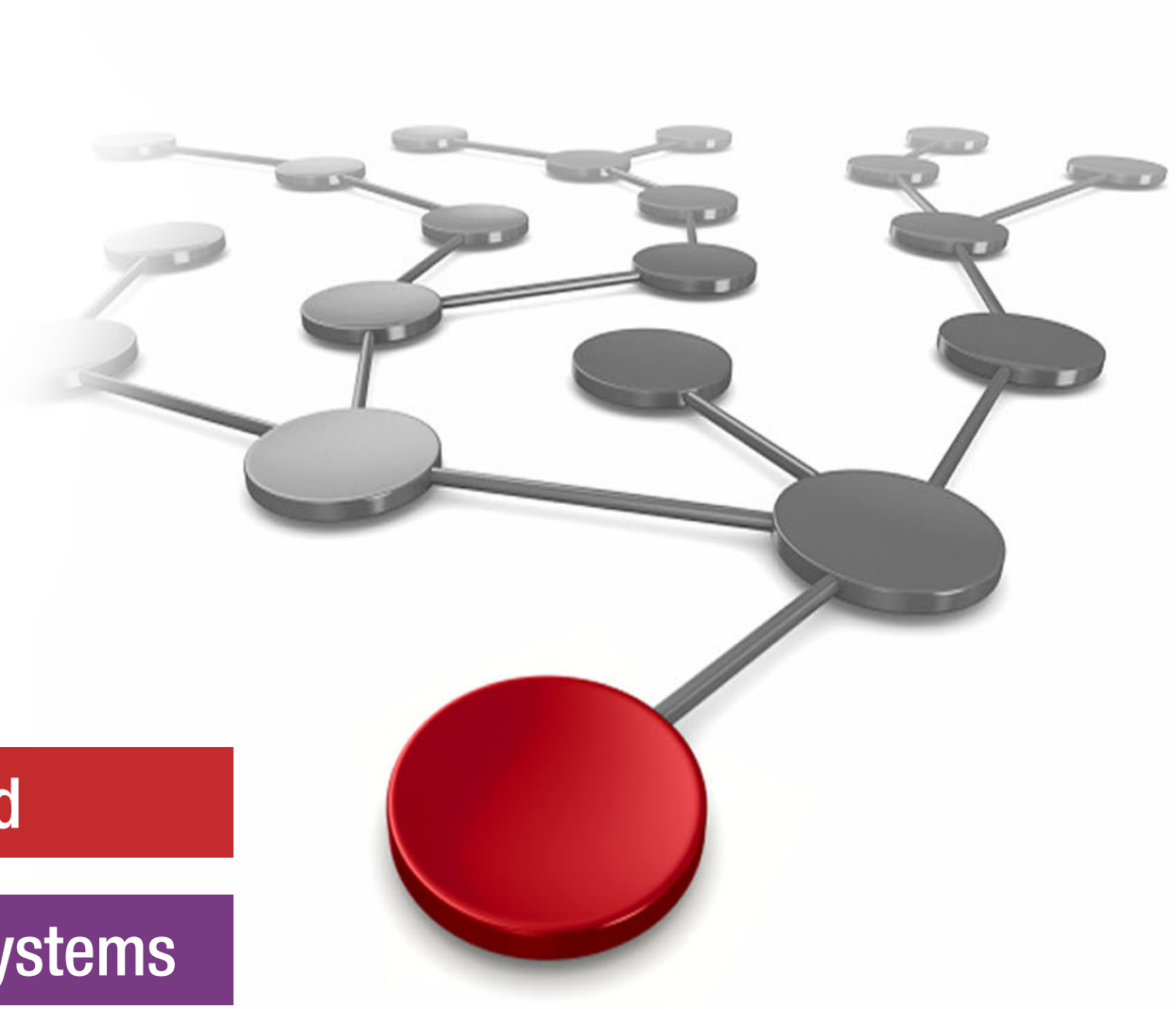
Draft Document for Review May 7, 2024 3:11 pm REDP-5728-00

Redbooks
ibm.com/redbooks

IBM Power S1012 Technical Overview and Introduction

Tim Simon
Henry Vo
Dean Mussari
Tsvetomir Spasov



Cloud

Power Systems

IBM

Redpaper

IBM Redbooks

IBM Power S1012 Introduction

May 2024

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

First Edition (May 2024)

This edition applies to the Power S1012 (9028-21B)
.

This document was created or updated on May 7, 2024.

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 https://www.ibm.com/legal/copytrade.shtml

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

AIX®
IBM®
IBM Z®
Maximo®
POWER®
Power8®
Power9®
PowerHA®
PowerVM®
Redbooks®
Redbooks (logo) ®
System z®
SystemMirror®

The following terms are trademarks of other companies:

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

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

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

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

Preface

This IBM Redpaper publication is a comprehensive guide that covers the IBM Power S1012 (9028-21B) which is an entry level server in the IBM Power10 Scale Out Server line. Like the other Power10 Scale Out Servers, the Power S1012 uses IBM Power10 processor-based technology and supports IBM AIX®, IBM i, and Linux (including Red Hat OpenShift).

The Power S1012 is especially suited for use in small IBM i environments, small database environments and as an Edge computing device running various AI workloads. While the Power S1012 is designed as an entry level system, it still retains the benefits provided by the IBM Power10 chip for reliability, security and enhanced performance.

One unique feature of the Power S1012 is the support of industry standard DIMMs in its memory configuration. In addition the system supports up to four internal NVMe drives for internal storage and four PCIe Gen5 PCIe slots to provide high performance connections to your users and other data sources.

This publication is intended for the following professionals who want to acquire a better understanding of IBM Power server products:

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

This paper is designed to complement the other IBM Redbook publications describing the Power10 server line. This information will also be integrated into the existing *IBM Power S1014, S1022s, S1022, and S1024 Technical Overview and Introduction*, REDP-5675.

Authors

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

Tim Simon is an IBM® Redbooks® Project Leader in Tulsa, Oklahoma, USA. He has over 40 years of experience with IBM, primarily in a technical sales role working with customers to help them create IBM solutions to solve their business problems. He holds a BS degree in Math from Towson University in Maryland. He has worked with many IBM products and has extensive experience creating customer solutions using IBM Power, IBM Storage, and IBM System z® throughout his career.

Henry Vo is an IBM Redbooks Project Leader with 10 years experience in IBM. He has technical expertise in business problem solving, risk/root-cause analysis, and writing technical plans for business. He has held multiple roles at IBM including Project management, ST/FT/ETE Test, Back End Developer, DOL agent for NY. He is a certified IBM zOS Mainframe Practitioner including IBM Z® System programming, Agile, and Telecommunication Development Jumpstart. Henry holds a Master of MIS (Management Information System) from the University of Texas in Dallas.

Dean Mussari is an IBM Power Brand Technical Specialist in the National Market in the USA. He recently came to IBM bringing 35 years of experience working with IBM servers and storage solutions in large retail environments. His main area of expertise is Power Servers with a focus on IBM i. He holds a masters degree in computer science from Loyola University of Chicago.

Tsvetomir Spasov is a Power Servers Hardware Product Engineer in Sofia, Bulgaria. He has 8 years of experience with IBM in a RTS, SME and PE role. His main area of expertise is HMC, FSP, eBMC, POWERLC and GTMS. He holds a masters degree in Electrical Engineering from Technical University of Sofia.

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

Nicole Schwartz Nett, Power Systems Architect and Platform Engineer
IBM Austin

Brandon Pederson, IBM i and Power Product Marketing Manager
IBM Austin

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:

  https://www.linkedin.com/groups/2130806

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

  https://www.redbooks.ibm.com/subscribe

- Stay current on recent Redbooks publications with RSS Feeds:

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

1

IBM Power10 Servers

IBM Power systems are purpose-built for today's demanding and dynamic business environments, and these new systems are optimized to run essential workloads – such as databases and core business applications – as well as maximize the efficiency of containerized applications. An ecosystem of solutions with Red Hat OpenShift also enables IBM to collaborate with clients connecting critical workloads to new, cloud-native services designed to maximize the value of their existing infrastructure investments.

This chapter summarizes the offerings in IBM's Power10 line of servers and introduces the newest member of the Scale Out server family – the IBM S1012 server. The following topics are presented:

- "Introduction to the IBM Power10 server family"
- "Power10 server advantages for your enterprise"
- "Power10 Scale Out systems – now including Power S1012"
- "Introducing the IBM Power10 S1012"

1.1 Introduction to the IBM Power10 server family

IBM Power10 based servers bring unique value to Mission Critical Computing. The IBM Power10 family brings high-end server capabilities in a range of servers purpose-built for requirements ranging from entry-level workloads to extremely scalable high performance servers. These systems not only provide high performance, in a highly available package, but they also support critical security features such as transparent memory encryption and advanced processor/system isolation.

IBM has designed different servers in the Power10 family to meet the needs of different business segments. The highest performing and most scalable server is the IBM Power E1080 which supports up to 240 Power10 cores and 64 TB of memory to support even the most demanding mission critical workloads. The IBM Power E1050 is a mid size four socket Power10 based server supporting up to 96 cores and 16 TB of memory – perfect for data workloads with high memory requirements. The IBM Power10 Scale Out servers are a range of rack mount or tower servers which provide enterprise-class reliability, security at a size to support smaller workloads and meet your scale out computing requirements.

Figure 1-1 illustrates the Power10 server ecosystem which provides support for your AIX, Linux or IBM i workloads for high end, midrange, and scale out environments.

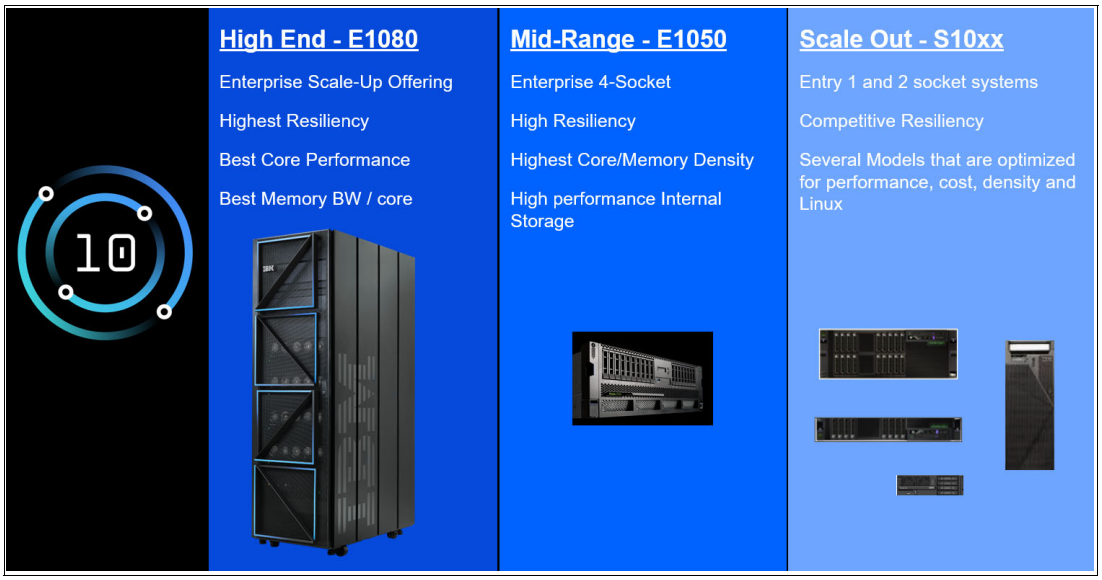


Figure 1-1 Power10 System

The Power10 processor-based server family uses the capabilities of the latest Power10 processor technology – which are defined in section 2.2, “IBM Power10 processor technology” on page 15 – to deliver unprecedented security, reliability, and manageability for your cloud and cognitive workloads. IBM Power10 servers deliver industry-leading price and performance for your mission-critical workloads. The IBM Power10 line of servers provides:

- Improved Total Cost of Ownership
  - Scale up large, mission critical landscapes with fewer systems and less networking.
- Industry leading scale & performance
  - Benchmarks – including a record SAP SD benchmark – show a 2X increase in per core performance compared to x86 alternatives.[1]

[1] https://www.sap.com/dmc/exp/2018-benchmark-directory/#/sd?sort=Benchmark%20Users&sortDesc=true

- Increased availability
  - Greater than 99.999% availability making the platform perfect for mission critical workloads.
- Enhanced Security
  - Encryption enabled for 100% of data at rest and in motion. This is done using Quantum safe standards.
  - Secure and Trusted Boot with TPM module.
- Sustainability
  - The 8-socket E1080 is 2x more energy efficient than a 16-socket x86.
  - Saves 40-80% in energy costs by upgrading to Power10 while reducing carbon footprint.
- On chip AI acceleration
  - Built-in inference engine brings AI closer to your data to reduce cost and complexity.

1.2 Power10 server advantages for your enterprise

IBM Power10 based servers are designed to improve the way that you build information technology solutions in your enterprise. Because of the increased scalability and performance, the built in AI capabilities, the built in security enhancements and the reliability innovations built into the system you can gain the following benefits.

Respond faster to business demands

Set a new standard for performance with containerized cloud-native applications. Designed with from one to eight sockets and with four times the number of processor threads compared to x86 alternative, Power10 servers can be sized to efficiently meet your business requirements.

Streamline insights and automation

Run inferencing models directly on the chip to bring AI closer to core enterprise data and applications. Power10 chips have four matrix math accelerator (MMA) engines per core which are used to accelerate AI inference capabilities. With IBM's broad portfolio of AI software, the processor can help infuse AI workloads into typical enterprise applications to glean more impactful insights from your data.

Protect data from core to cloud

Enable advanced protection with workload isolation and platform integrity to the processor as part of a zero-trust approach. Each Power10 chip has eight cryptographic accelerators per core to offload encryption algorithms and allow for 100% data encryption with no performance implications.

IBM Power's multi-layered approach to security gives you full visibility of your hardware and software. With IBM Power10's hardware-accelerated transparent memory encryption, quantum-safe cryptography and fully homomorphic encryption, it protects your data with comprehensive end-to-end security at every layer of the stack – to address both today's and tomorrow's threats.

IBM Power10 security is based on:

- End to End security with full stack encryption.
  - No additional management setup or performance impact for transparent memory encryption.
  - Compared with IBM Power9 hardware-accelerated encryption – 4x crypto engines in every core.
  - Support Quantum-safe cryptography and Fully homomorphic encryption for defending against threats.

- Confidential Computing: Isolation and Integrity.
  - Enhanced CPU isolation from Service Processors.
  - Limit CPU resources accessible by BMC/FSP.
  - Creates separate domains between service processor and CPU.
  - Performance enhanced side channel avoidance.
  - Protect against data leakage.
  - Enhanced handling of automatic thread isolation from speculation-based attacks.
  - Protect Integrity of Return Stack against Return-Oriented-Programming attacks.
  - New in-core hardware architecture with low hardware footprint and standards-based cryptography.
- Secure and Trusted boot for host and guest LPARs.

Figure 1-2 shows the end to end encryption strategy of the IBM Power10 servers.

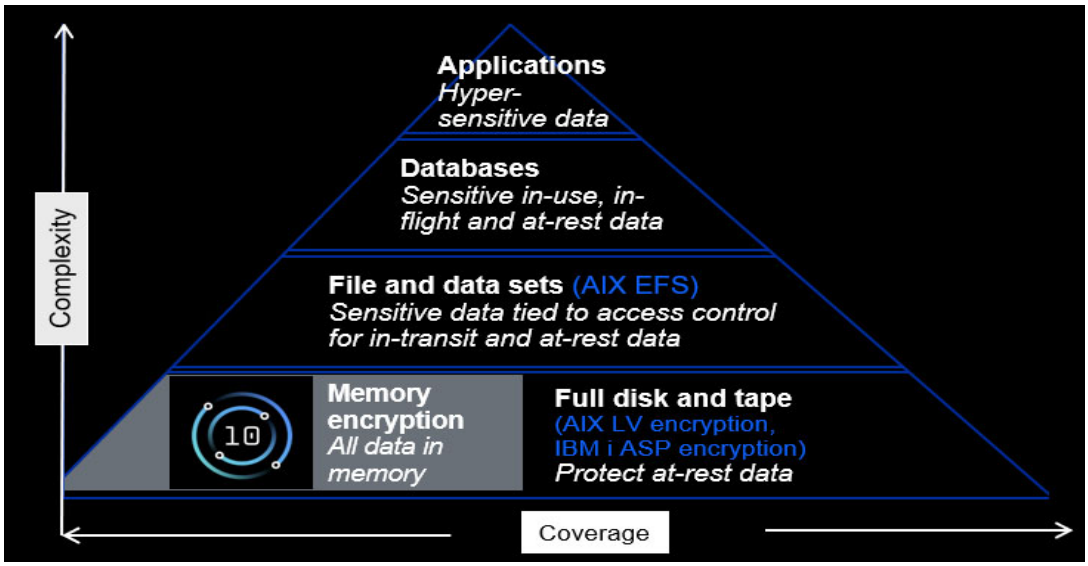


Figure 1-2 End to End encryption

Ensure uptime and reliability.

Built with redundancies, substantial retry capabilities, and self-healing capabilities, Power10 servers provide class-leading infrastructure reliability. Power10 based servers provide better than 99.999% uptime. In addition, there is a rich variety of features and programs that support business continuity and disaster recovery solutions designed to meet your enterprise requirements.

Reduce energy usage and reduce carbon footprint

Upgrading to Power10 can reduce energy consumption, reduce datacenter floor space and increase performance and capacity. For example you can save 28 MWh per year with E1080 when compared to the E880C – enough to power nine houses in Germany for a year.

Note: E880C and E1080 based on Energy Estimator and published rPerf. All others based on published maximum system rPerf performance and maximum energy consumption

German household consumption is 3,149kWh per year as of 2020 based on https://www.iamexpat.de/housing/real-estate-news/energy-consumption-germany-falling-prices-continue-rise

Figure 1-3 shows an example of a customer who migrated their existing workload to Power9 servers and then upgraded later to IBM Power10 servers. In the process they reduced the number of servers that they had to support, reduced the number of software licenses required as well as reducing their energy consumption and carbon footprint.

Figure 1-3 Carbon footprint impact of Power10 servers

1.3 Power10 Scale Out systems – now including Power S1012

To meet the requirements of smaller workloads, IBM delivers a range of entry level and Scale Out servers at price points that allow you to meet your budget requirements. Figure 1-4 shows how IBM Power10 servers can be utilized from the Edge to the Cloud.

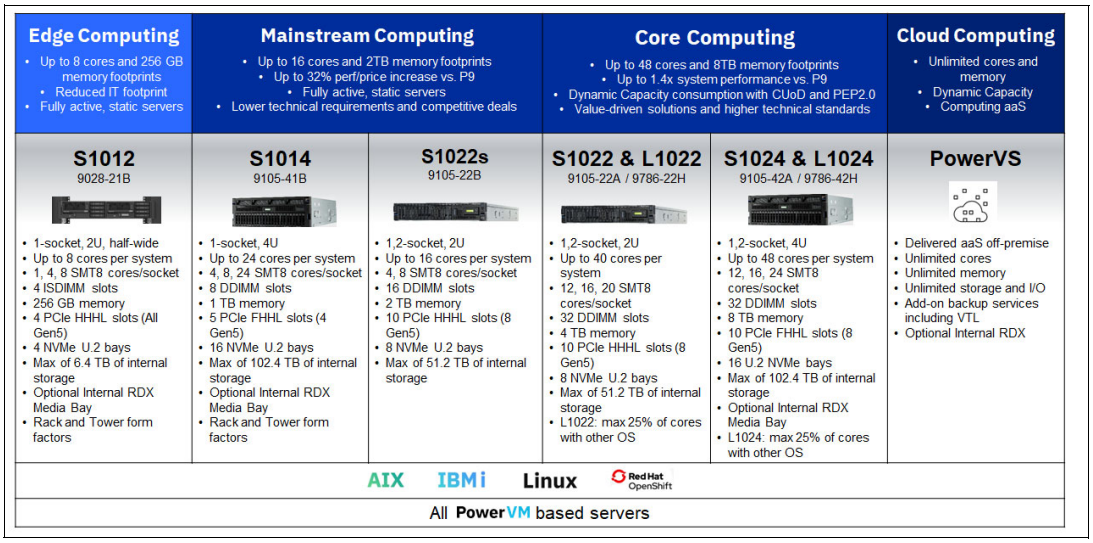


Figure 1-4 Power10 Scale Out options

1.3.1 Entry Systems and Enhanced Performance and Scale Systems

The Power10 Scale Out server offerings can be divided into two groups; Entry Systems and Enhanced Performance and Scale systems. These offerings allow you to customize your Power10 solution to meet your specific requirements. This section details these offerings.

Entry Systems

The entry systems are one or two socket servers which provide processor options starting as small as a single core and ranging up to 16 cores per system. Memory configurations range from 32 GB to 2 TB depending on the model and the memory options chosen. Each of these options offer differing levels of internal NVMe storage and offer from 4 to 10 PCIe Gen4 and Gen5 slots.

In the entry systems all cores are fully active with static activations. The IBM PowerVM hypervisor is fully supported across each of the options. In general these systems provide a 32% performance/price increase over the IBM Power9 based options. They generally have lower technical requirements and are designed for easier installation. Often there are very competitive deals available.

Figure 1-5 gives the highlights of each of the three entry systems (now including the S1012) in the Power10 Scale Out lineup. We provide more detail on the new Power S1012 model in section 1.4, “Introducing the IBM Power10 S1012” on page 7. We then provide additional technical details in Chapter 2, “Power S1012 Technical Details” on page 13 and Chapter 3, “Power S1012 Use Cases” on page 33.

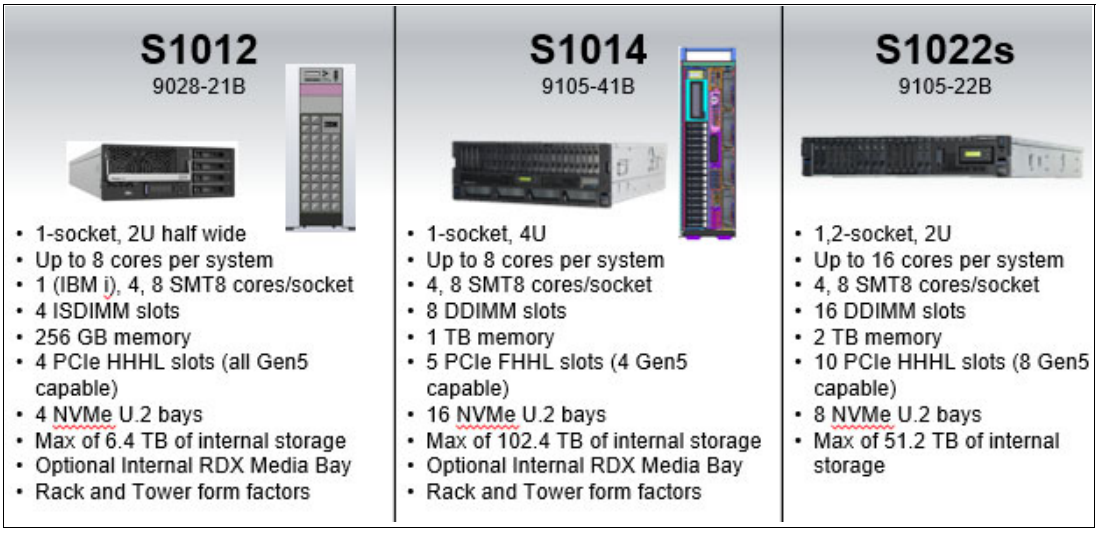


Figure 1-5 Entry Systems

Enhanced Performance and Scale

The enhanced performance and scale group includes four systems. Two of these systems – the S1022 and S1024 – are designed as general purpose systems supporting the full range of operating systems supported by Power10. The other two systems – the L1022 and L1024 – are optimized for Linux workloads. In the L10xx models, while all of the available cores can be used to support Linux partition, the number of cores running non-Linux operating systems are limited.

This group provides up to 50% performance/price increase and 1.4x more system performance compared to their Power9® based alternatives. These models provide up to 48 cores and 8 TB memory footprints. They also offer Dynamic Capacity consumption with Capacity Upgrade on Demand (CUoD) and Private Enterprise Pools 2.0 (PEP2.0) which allows the client to provide a

private cloud environment with dynamic capacity sharing across an enterprise including the use of pay as you go credits within a processor group. There are also several value driven solutions available for different industry workloads.

Figure 1-6 shows the members of these groups and summarizes their capabilities. For more details on these servers please reference *IBM Power S1014, S1022s, S1022, and S1024 Technical Overview and Introduction*, REDP-5675.

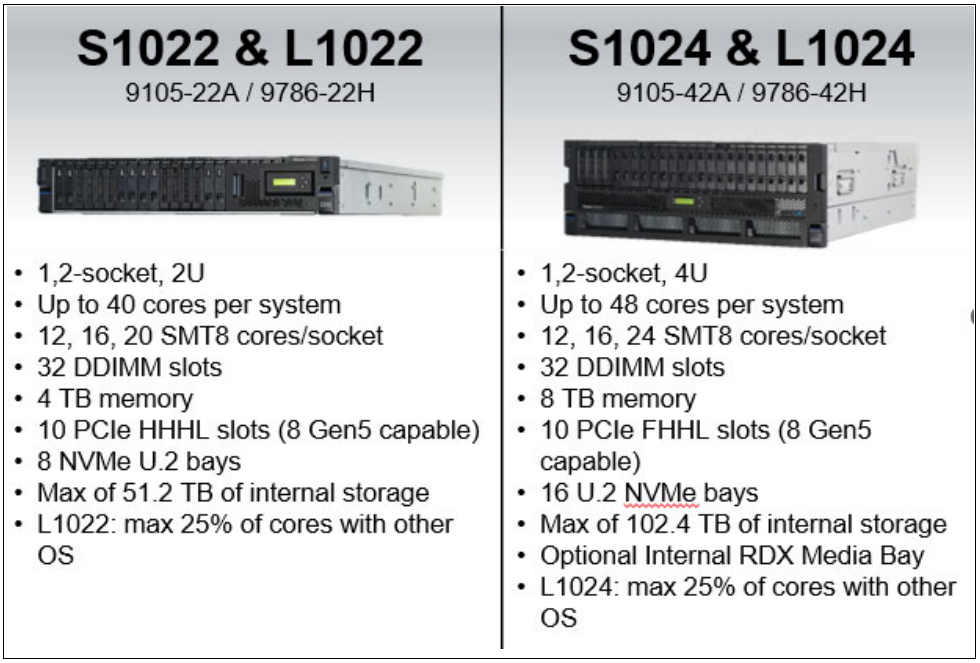


Figure 1-6 Enhanced performance & scale systems

1.4 Introducing the IBM Power10 S1012

The IBM S1012 is a 1-Socket Power10 based system which is offered in a rack (2U 1/2 rack format) or a desk side tower version. The system is designed as an entry point for customers and is particularly suited for low end IBM i servers, small database servers or for Edge computing. It is unique in the Power10 server environment in that it supports industry standard DDR4 DIMMs for its memory configuration.

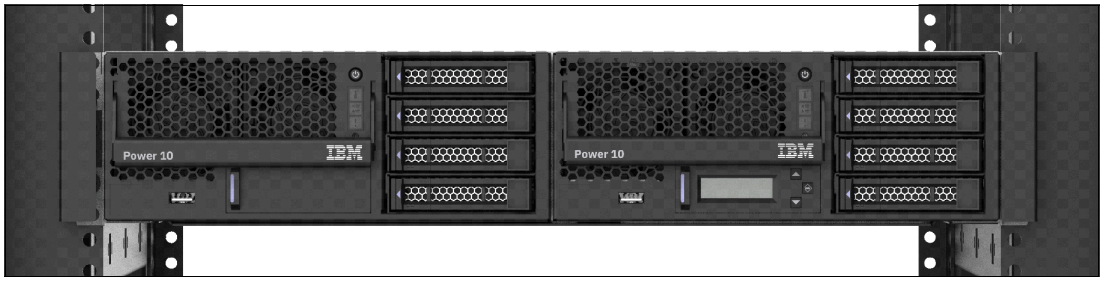


Figure 1-7 Two System S1012 systems rack mounted side-by-side in a single 2u space

1.4.1 Power S1012 Highlights

The following list highlights the features of the Power S1012:

- Rack and Tower form factors
  - Rack version is a 2U half-wide form factor
- Single Power10 eSCM processor with one of the following options
  - 1 SMT8 core (IBM i only)
  - 4 SMT8 cores
  - 8 SMT8 cores (rack only)
- Industry Standard DIMM
  - Four slots for DDR4 DIMMs
  - System memory capacity up 256 GB max
  - Main memory encryption for added security
- 4 PCIe HHHL Gen5 slots
  - 2 PCIe G5 x8 / G4 x16
  - 2 PCIe G5 x8
  - No support for external PCIe drawer
- Built in NVMe enclosure
  - Support up to 4 NVMe U.2 Flash drives
  - 15 mm U.2 format (supports 7 mm in a 15 mm carrier)
  - 800 GB and 1.6 TB options available for NVMe drives
  - Provides up to 6.4 TB of storage
  - No support for external I/O drawers
- Optional RDX
- Secure and Trusted Boot with TPM module
- Titanium power supplies to meet EU Efficiency Directives
  - 2x 800W industry standard
  - 100-240VAC C14 inlet
  - Built in Advanced Thermal and Power Management
- Enterprise BMC managed
  - HMC optional
- Support for IBM i, AIX, and Linux (including OpenShift)

1.4.2 Rack configurations

The rack mount version of the Power S1012 is a 2u half-wide form factor. For installation in a rack, IBM provides a full rack width “cage” to hold either one or two systems in a 2U rack slot. There are three possible configurations for the Power S1012 in the rack:

1. Two side by side Power S1012 systems in the cage as shown in Figure 1-7 on page 7.
2. Single Power S1012 system with a blank cover for the other half of the cage.
3. Single Power S1012 system with an RDX enclosure installed in the other half.

1.4.3 Benefits of the Power S1012

The Power S1012 is designed as an entry system, but it still benefits from the enterprise features that are available with all Power10 servers. The following enterprise and usability features are supported.

Flexible power supply and rack or desk side configurations

The Power S1012 supports both 110 volt and 220 volt input voltages. This means that it can easily integrate into your environment whether you have a machine room environment or not. In your machine room, you can integrate the Power S1012 directly into existing racks and plug it into existing power distribution units (PDUs). The system can be shipped in a rack to your location or you can integrate it into a new or existing rack that you provide. In non machine room environments, it can be plugged directly into a wall outlet – even a 110 volt outlet in the US or other countries that run on 110 volts.

Fully functional AI acceleration

AI acceleration utilizing the MMA engines built right into the core is fully functional in the Power S1012. This functionality supports popular open source library instructions, allowing low latency AI inferencing at the location where the data is being collected, allowing for data gravity.

Enhanced cryptography support

Power10 includes 4x more crypto engines in every core compared to Power9. This accelerates encryption performance across the stack. For example, the performance of the widely used AES encryption algorithm is improved by 2.5x over POWER9. With these innovations, along with new in-core defenses against programming attacks and support for Post Quantum Encryption and Fully Homomorphic Encryption your data is better protected.

Support for memory encryption

With data residing in an increasingly distributed environment, you cannot set a perimeter to it anymore. This reinforces the need for layered security across IT stack. Power10 family of servers introduce a new layer of defense with transparent memory encryption. Transparent memory encryption is designed to simplify encryption and support end-to-end security, without impacting performance, by leveraging hardware features for a seamless user experience. Since this capability is enabled at the silicon level, there is no additional management setup or performance impact.

Processor chip and memory subsystem

The Power S1012 uses the same eSCM module supported in the other Power10 Scale Out servers with the main differences being the number of cores active (one, four or eight) and the number of OMI memory channels.

The use of OMI for memory connections provides a high performance serial interface which is common across all of the Power10 servers. The OMI interface is connected to memory controllers which are soldered onto the system planar in the Power S1012. These buffered memory controllers translate the data packets from the OMI connection to support the industry standard DIMM slots that allows the use of Industry standard DIMMs.

Ease of use

The Power S1012 enclosure is designed to not require extra brackets, EMC covers and shipping brackets are added or removed during installation. The Power S1012 LEDs are designed to support the maintenance and operation of the system and support a number of concurrent maintenance actions. Concurrent maintenance is the ability to fix a system while it is operational – for example installing, removing or replacing an internal NVMe drive or replacing a power supply.

The Power S1012 can be managed with or without the use of an HMC. Although the HMC provides some significant management advantages, for some environments the ability to run without an HMC is a preferred option.

1.4.4 Limitations

There are some functions that are available on other Power10 Scale Out servers that are not supported on the Power S1012.

- Only static activations are supported.

  Like all of the members of the entry level group in the Power10 Scale Out server family, the Power S1012 only supports static activations. All processors and all of the memory are active from the factory, no support is provided for capacity on demand or Power Enterprise Pools.

- No external PCIe expansion drawers or external storage drawers are supported.

  Based on the use cases for this server, there is no plan to have the capability of expanding the number of PCIe slots or the number of NVMe storage devices.

- PCIe slots do not support concurrent maintenance.

  Adding or removing a PCIe card from a PCIe slot in the Power S1012 requires that the server be removed from service and powered down.

- Advanced virtualization functions not enabled.

  Other offerings within the Power10 server family offer a range of advanced virtualization functions in conjunction with the virtual i/o server capability provided by IBM PowerVM®. The following functions are not enabled in the Power S1012:

  – Live Partition Mobility (LPM) is not supported.

    LPM provides the ability to move a running Power LPAR to another system without interruption. This function is not supported for the Power S1012 at general availability.

  – Network virtualization functions required for LPM are not supported.

    There are some network virtualization capabilities that are designed to ensure an LPAR is capable of being moved using LPM. Since LPM is not provided, these functions are not required and so are not supported on the Power S1012 at general availability:

    - Virtual Network Interface Card (vNIC)
    - Hybrid Network Virtualization (HNV)

  – Some advanced memory virtualization functions are not enabled.

    There are some specific advanced memory virtualization functions that have not been enabled on the Power S1012. These functions are used in systems with larger amounts of installed memory and were not enabled as they were not part of the primary use cases of the Power S1012. The following memory functions are not supported:

    - Active memory mirroring (AMM)
    - Virtual Persistent Memory (vPMEM)

1.4.5 Operating System Levels supported at GA

The following is a list of the supported operating system levels that will be supported at the general availability of the Power S1012:

- VIOS
  - VIOS 4.1.0.20
  - VIOS 3.1.4.40 (P9 mode)
- AIX

  For AIX LPARs with any I/O defined, the following versions are supported:

  - AIX Version 7.3 with the 7300-02 Technology Level and Service Pack 7300-02-02-2420, or later (running in P10 mode)
  - AIX Version 7.3 with the 7300-01 Technology Level and Service Pack 7300-01-04-2420, or later (planned availability - July 26, 2024)
  - AIX Version 7.2 with the 7200-05 Technology Level and Service Pack 7200-05-08-2420, or later (running in P9 mode)

  If installing the AIX operating system in an LPAR with only Virtual I/O defined, the following versions are supported:

  - AIX Version 7.3 with the 7300-00 Technology Level and service pack 7300-00-02-2220, or later
  - AIX Version 7.2 with the 7200-05 Technology Level and service pack 7200-05-04-2220, or later
  - AIX Version 7.1 with the 7100-05 Technology Level and Service Pack 7100-05-10-2220, or later (running in P8 mode)

Important: Support for AIX Version 7.1 requires that you have a valid service extension contract in place.

- IBM i
  - IBM i 7.4 TR10 or later
  - IBM i 7.5 TR4 or later
- Linux
  - RHEL 9.2 for PowerLE or later
- OpenShift
  - OpenShift Container Platform 4.15 or later

1.4.6 IBM i Software Tiers

The Power S1012 is eligible for the P05 software tier when ordered with either the single core or four core CPU options. When using the eight core option, the system is in the P10 software tier.

The IBM i P05 and P10 software tiers are based not only on the number of cores, but they also have other system limits to qualify.

- An entry POWER10 P05 server can have a maximum of four (4) cores active with IBM i OS, 64 GB of RAM memory and 3.2 TB of NVMe (SSD) storage. This resource range is generally adequate for most entry small-to-medium (SMB) IBM i users.
- An Enterprise POWER10 P10 server can have up to eight (8) cores active with IBM i OS, 1 TB of RAM memory and 102 TB storage in the CEC.

1.4.7 IBM i Subscription Term Licensing

The IBM i licensing strategy for the Power S1012 is to offer subscription term licenses only. This is in line with the general strategy for all IBM i P05 and P10 offerings to be subscription term licensed only. Note that in the future the P20 and P30 Tiers will also be migrated to subscription term licensing as well.

Subscription licensing is a trend across the industry. As customers continue to move to a consumption model to provide additional flexibility in their Information Technology infrastructure, subscription licenses allow better connection between their consumption and their costs.

Subscription Term value to clients

Moving to subscription licensing provides the following benefits to our clients.

- Lower up-front costs; consistent payments
- Increased Flexibility (run / move to where you want)
- Simplicity (eliminate complex license/SWMA transfers)
- Improved technology currency

Acquisition options

Clients who acquire the Power S1012 server will have the following IBM i Licensing options:

- Acquire or renew IBM i Subscription Term licenses
- Convert IBM i non-expiring licenses from another “donor” machine to IBM i Subscription Term license at a lower-priced subscription term option. Note the following requirements:
  - Active SWMA on the donor is required.
  - Transfer of IBM i non-expiring licenses is not supported.
  - Conversion pricing applies for 3-year, 4-year, 5-year Subscription Terms.

1.4.8 Warranty coverage

Maintenance service for the Power S1012 is provided under warranty after purchase. IBM will attempt to resolve your problem over the telephone or electronically through an IBM website. There are also remote support capabilities for direct problem reporting, remote problem determination, and resolution with IBM. Following problem determination, if IBM determines on-site service is required, scheduling of service will depend upon the warranty conditions and availability of parts.

The default warranty for the Power S1012 is:

Warranty Period: 3 years
Service Level: IBM CRU & On-Site, 9x5 Next Business Day

The following warranty service upgrades are available:

- IBM On-Site Repair, 9x5 Same Day and
- IBM On-Site Repair 24x7 Same Day

Service levels are response-time objectives and are not guaranteed. The specified level of warranty service may not be available in all worldwide locations. Contact your local IBM representative or your reseller for country-specific and location-specific information.

2

Power S1012 Technical Details

This chapter provides a more in-depth view of the Power S1012 server. In order to understand the full abilities of the Power S1012, we spend a little time reviewing the architecture and capabilities of the IBM Power10 chip and the benefits it brings to your business in the area of performance, scalability, and security. We then describe the different components of the Power S1012 and show the different configuration options available for you to customize the server to your business needs.

The following topics are discussed:

- “Power S1012 Overview”
- “IBM Power10 processor technology”
- “IBM Power S1012 architecture and components”
- “Physical planning”

2.1 Power S1012 Overview

The Power10 processor-based Scale Out servers integrate the capabilities of the latest Power10 processor technology to deliver unprecedented security, reliability, and manageability for your cloud and cognitive infrastructure and delivers industry-leading price and performance for your mission-critical workloads. These systems are ready for hybrid cloud deployment, with Enterprise grade virtualization capabilities built into the system firmware with the PowerVM hypervisor.

The new Power10 S1012 server is designed as an entry level Power10 Scale Out server targeted for clients with smaller CPU and memory requirements that are provided by other Power10 Scale Out servers. This server is ideal for smaller IBM i clients and for use in small database servers. It is also an excellent choice for use in edge computing – for example, clusters that are running Red Hat OpenShift. The Power S1012 is available in either a rack mount or a standalone tower form factor. The rack mount is shown in Figure 2-1.

Figure 2-1 Rack mount version of a single Power S1012 System (sled)

The tower version is shown in Figure 2-2.

Figure 2-2 Tower version of the Power S1012

The Power S1012 is a single socket server supporting up to eight SMT8 cores. The system supports industry standard RDIMMs providing four slots and supporting up to 256 GB of memory. Memory encryption is supported for added security.

The system provides four PCIe Gen5 slots which support half-height, half-length card formats for connectivity to networks and external storage. There is no on-board network card, network connectivity is provided by an appropriate network interface controller (NIC) installed in one of the PCIe slots.

Internal storage is provided by four NVMe U.2 Flash Bays supporting up to 6.4 TB of internal NVMe storage. There is no SAS backplane supported on the Power S1012. RDX and other USB devices are supported.

Power for the Power S1012 is provided by dual Titanium power supplies which meet EU Efficiency Directives and support voltages from 100 VAC to 240 VAC with a C14 inlet.

System management is provided by a new Enterprise BMC card, use of the hardware management controller (HMC) is optional.

The S1012 is supported by IBM i, AIX, Linux and Red Hat OpenShift Container Platform.

2.2 IBM Power10 processor technology

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

- Data plane bandwidth

  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

  New core micro-architecture, flexibility, larger caches, and reduced latencies.
- End-to-end security

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

  Up to threefold energy-efficiency improvement in comparison to Power9 processor technology.
- Artificial intelligence (AI) infused core

  A 10-20x matrix math performance improvement per socket compared to the 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 S1012 server.

The IBM Power10 Processor session material as presented at the 32nd HOT CHIPS conference is available at this web page.

[1] https://hotchips.org/

2.2.1 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 nonuniform cache access (NUCA) L3 cache.

The processor supports the following three distinct functional interfaces that all can run with a signaling rate of up to 32 Gigatransfers per second (GTps):

- Open memory interface

  The Power10 processor has eight memory controller unit (MCU) channels that support one open memory interface (OMI) port with two OMI links each[2]. One OMI link aggregates eight lanes that are running at 32GTps and connects to one memory buffer-based differential DIMM (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 1TBps.

- SMP fabric interconnect (PowerAXON)

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

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

  Because of the versatile characteristic of the technology, it is also referred to as *PowerAXON* interface (Power A-bus/X-bus/OpenCAPI/Networking[3]). The OpenCAPI and the memory clustering and memory interception use cases can be pursued in the future and as of this writing are not used by available technology products.

- PCIe Version 5.0 interface

  To support external I/O connectivity and access to internal storage devices, the Power10 processor provides differential Peripheral Component Interconnect Express version 5.0 interface bus (PCIe Gen 5) with a total of 32 lanes.

The IBM Power10 processor has several design characteristics that differentiate it from other processor technologies and lead to it industry leading performance. Some of these are listed in Table 2-1 on page 17.

[2] The OMI links are also referred to as OMI subchannels.
[3] A-buses (between CEC drawers) and X-bussed (within CEC drawers) provide SMP fabric ports.

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

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

a. Complimentary metal-oxide-semiconductor (CMOS)
b. Power instruction set architecture (Power ISA)

The Power10 processor core inherits the modular architecture of the Power9 processor core, but with a redesigned and enhanced micro-architecture that significantly 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 acceleration engines can deliver significant performance gains for machine learning, particularly for AI inferencing workloads.

2.2.2 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 single-threaded (ST) mode. The ability to support multiple concurrent threads within the processor allows Power10 based servers to scale to support even the most demanding workloads,

If more than one hardware thread is active, the processor runs in SMT mode. In addition to the ST mode, the Power10 processor core supports the following 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 significantly enhances the throughput and scalability of multi-threaded applications and optimizes the compute density for single-threaded 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.

The Power10 processor-based scale-out servers use 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 *8-way simultaneous multithreading (SMT8) mode*. A Power10 core with SMT8 capability is named Power10 SMT8 core. The Power10 core also supports modes with four active threads (SMT4), two active threads (SMT2), and one single active thread (ST).

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

Figure 2-3 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.

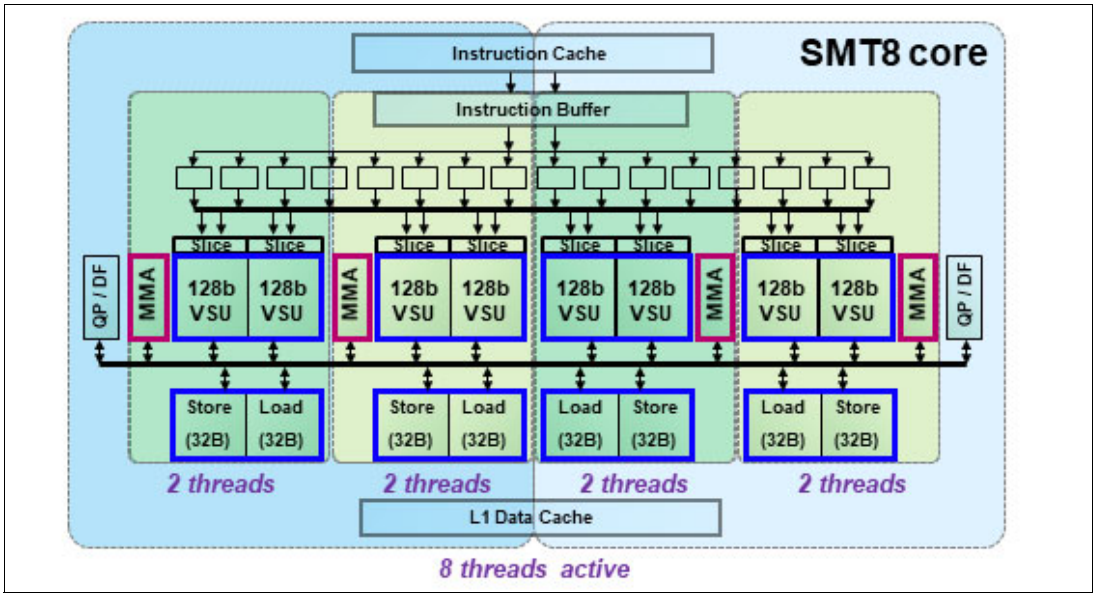


Figure 2-3 Power10 SMT8 core

Each of the two execution resource domains supports 1 - 4 threads and includes four vector scalar units (VSU) of 128-bit width, two matrix math accelerator (MMA) units, 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 also can be used as a combined execution resource, which is then named *super-slice*. When operating in SMT8 mode, eight SMT threads are subdivided in pairs that collectively run on two adjacent slices, as indicated by colored backgrounds in different shades of green in Figure 2-3.

In SMT4 or lower thread modes, one to two threads each share a four-slice resource domain. Figure 2-3 also shows other essential resources that are shared among the SMT threads, such as instruction cache, instruction buffer, and 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 a significant improvement of the performance per watt signature compared with the previous Power9 core implementation.

The overall energy efficiency per socket is better by a factor of approximately 2.6, which demonstrates the advancement in processor design that is manifested by the Power10 processor.

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 square parentheses.

Enhanced computation resources

The following computational resource enhancements are available:

- Eight vector scalar unit (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)/SHA) +400% operations].
- Four units for matrix math accelerator (MMA) acceleration each capable of producing a 512-bit result per cycle (new) [+400% Single and Double precision FLOPS plus support for reduced precision AI acceleration].
- Two units for quad-precision floating-point and decimal floating-point operations instruction types.

Larger working sets

The following major changes were implemented in working set sizes:

- L1 instruction cache: 2 x 48 KB 6-way (96 KB total); +50%
- L2 cache: 2 MB 8-way; +400%
- L2 translation lookaside buffer (TLB): 2 x 4K entries (8K 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)
- Translation lookaside buffer (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 on a cache miss only

In addition to significant 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

2.2.3 Power10 compatibility modes

The Power10 core implements the Processor Compatibility Register (PCR) as described in the Power instruction set architecture (ISA) version 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 version 3.0) or Power8 (Power ISA version 2.07) processors. The support for processor compatibility modes also enables older operating systems versions of AIX, IBM i, Linux, or Virtual I/O server environments to run on Power10 processor-based systems.

The Power10 processor-based scale-out servers support the Power8, Power9 Base, Power9, and Power10 compatibility modes.

2.2.4 IBM Power10 processor packaging

The Power10 processor is delivered in a variety of package options across the IBM Power10 server line. For the high end enterprise server – Power10 E1080 – the processor is packaged as a single core module (SCM). Up to 16 SCMs can be interconnected to provide a 16 socket high performance enterprise server. For the mid-range and Scale Out servers, the processors are packaged as dual core modules (DCM). The Scale Out systems utilize either one or two DCMs and up to four DCMs can be integrated in the mid-range enterprise E1050 servers as shown in Figure 2-4.

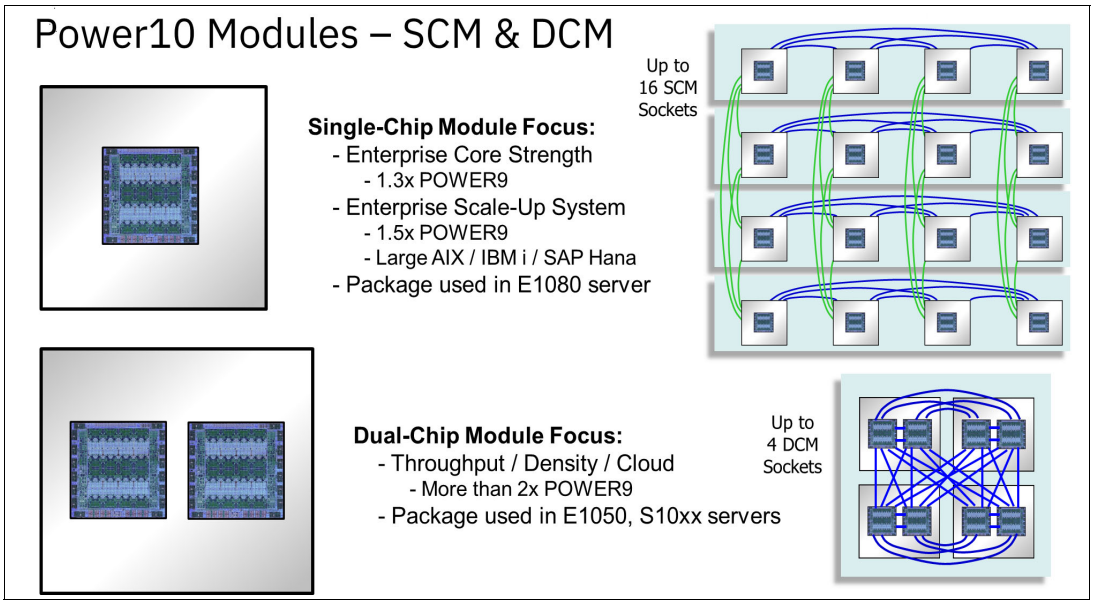


Figure 2-4 Single Core and Dual Core Modules in the Power10 server

2.2.5 Processor packaging in Scale Out servers

For the Power10 processor based scale-out servers, the Power10 processor is packaged as either a DCM or as an eSCM:

- The DCM contains two directly coupled Power10 processor chips plus more logic that is needed to facilitate power supply and external connectivity to the module.

  The Power S1022 and the Power S1024 servers use DCM modules, as does the Power S1014 when equipped with the 24-core processor option.

- The eSCM is a special derivative of the DCM where all active compute cores run on the first chip and the second chip contributes only extra PCIe connectivity, essentially a switch:

  The Power S1012, Power S1014 (4-core and 8-core options) and the Power S1022s servers are based on eSCM technology. The benefit to this is that these eSCM based servers qualify as single socket servers for the purposes of licensing. This can significantly reduce the software licensing costs for many software packages.

Figure 2-5 illustrates the differences between the DCM and eSCM modules.

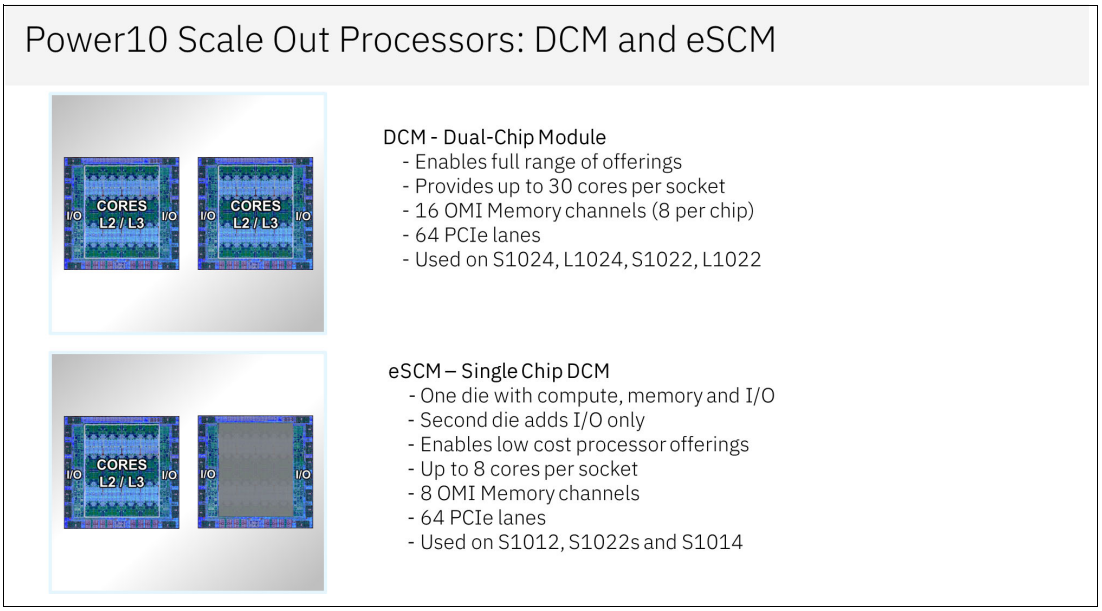


Figure 2-5 Scale Out module options DCM and eSCM

2.3 IBM Power S1012 architecture and components

The Power S1012 is designed to provide an entry level configuration option within the Power10 server family. The processor is delivered as an eSCM with the processor and memory components enabled on Chip-0 and the PCIe connectivity split between Chip-0 and Chip-1.

The Power S1012 includes a system board with the CPU and four Industry Standard DIMM slots, four PCIe Gen5 slots, four NVMe M2 slots and dual power supplies. The system is managed by an on-board eBMC card.

The system is available in either a rack-mount version or a tower model. The rack mount format is a 2U half-wide format. Figure 2-6 on page 22 shows the rack enclosure components.

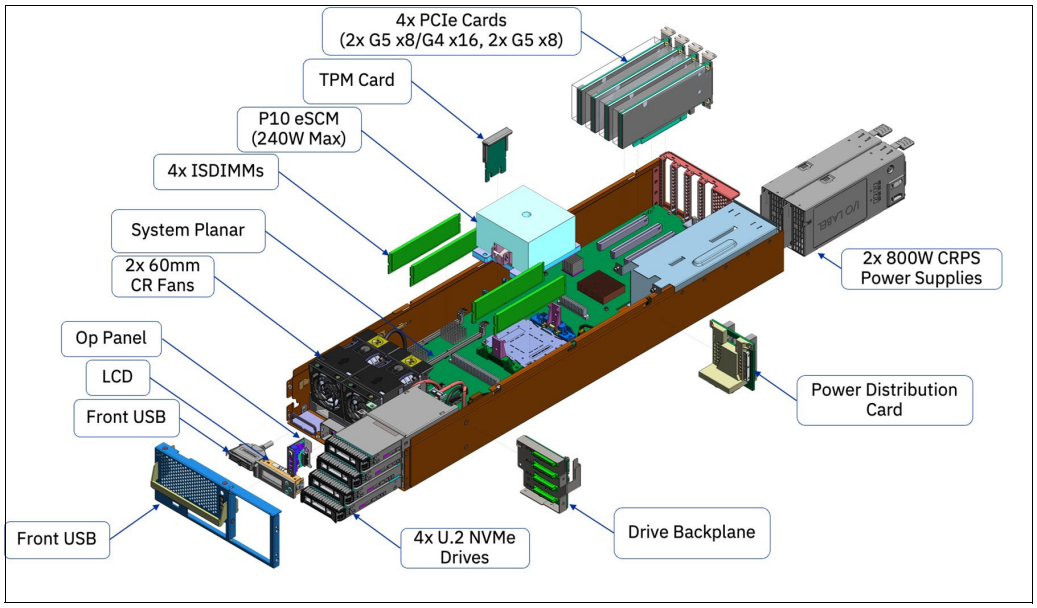


Figure 2-6 Rack mount enclosure components

Figure 2-7 shows the components of the tower version of the Power S1012.

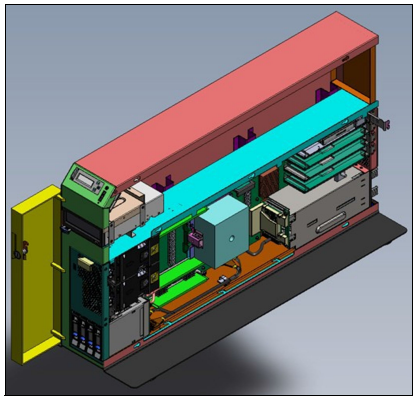


Figure 2-7 Desk side (tower) components

2.3.1 IBM Power S1012 system board

The Power S1012 has a new form factor – half rack server enclosure – compared to the original Power10 Scale Out servers and hence required a new system board. The system board contains:

- Processor - eSCM
- Voltage Regulators
- Memory controllers
- Industry Standard DIMM slots
- Service Processor (eBMC) components
- Gen4 Switch
- 4 PCIe Slots (all support x8 Gen 5)

Figure 2-8 is a high level view of the new system board.

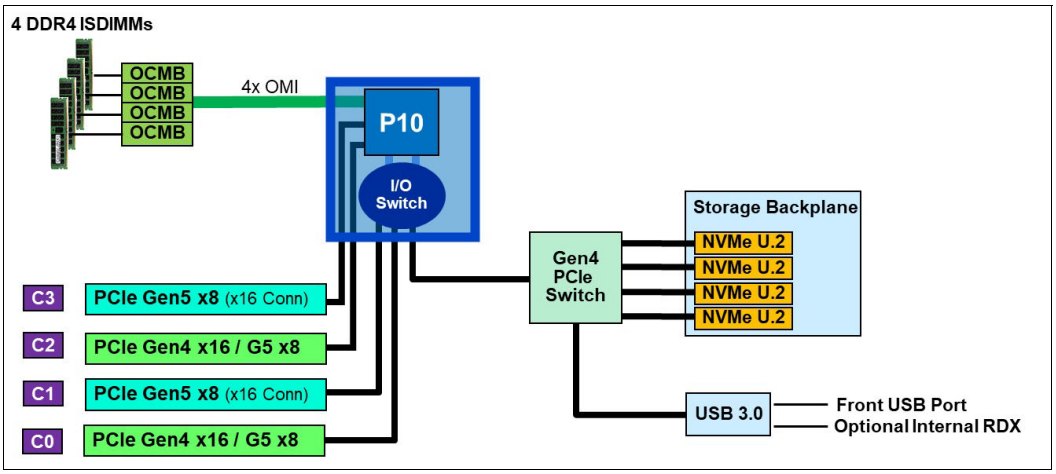


Figure 2-8 System board components

2.3.2 Processor module options for Power S1012

By design, the Power S1012 is a 1-socket server which supports the installation of one eSCM module. (The difference between the DSM and the eSCM processor modules is shown in section 2.2.5, “Processor packaging in Scale Out servers” on page 21). The Power S1012 can be procured with one, four or eight cores active.

The supported processor activation types and use models vary within the Power10 processor based scale-out server model types. The Power S1012 only supports the classical static processor activation model. All functional cores of the configured modules are delivered with processor activation features at initial order. This use model provides static and permanent processor activations.

There are three processor offerings available in the Power S1012 which are detailed in Table 2-2.

Table 2-2 Processor options available

Feature Code	Processor Entitlement[a]	Processor Cores	Typical Frequency Range	IBM i Software Tier
EPGZ	EPFY	8	3.00 to 3.90 GHz	P10
EPG7	EPFV	4	3.00 to 3.90 GHz	P05
EPG3[b,c]	EPFW	1	3.00 to 3.90 GHz	P05

a. One processor entitlement must be ordered for each processor core.
b. Support for IBM i in native mode only.
c. The one core processor option is only available in the tower configuration.

Tip: Although the one core processor option is not available in the rack configuration, it is possible to have a single core server in a rack-mounted form factor by purchasing the four core processor (EPG7) with quantity three of feature code 2319. FC 2319 allows you to permanently deactivate a single core in a server - usually used to reduce software licensing requirements.

The processors support Enhanced Workload Optimized Frequency to support optimum performance. The maximum frequency is achievable without reducing cores. Processor

frequencies are dynamic and are set to Max Performance Mode by default, The default can be changed using the system management tools. For more details on power and performance management reference this document. Additional information is also provided in Chapter 2 of *IBM Power S1014, S1022s, S1022, and S1024 Technical Overview and Introduction*, REDP-5675.

2.3.3 Memory subsystem

The Power10 processor was architected with an enhanced memory connection - the Open Memory Interface (OMI). The Open Memory Interface enables extremely low latency and high bandwidth RAM using serial connections. Using serial memory communications to off-chip controllers reduces the number of signaling lanes to and from the chip, increases the bandwidth and makes the processor agnostic towards what technology is in the memory, making the system flexible and future proofed.

The Power S1012 is unique in that it supports industry standard DIMMs. Using industry standard DIMMs allows the Power S1012 to reduce the cost of memory in the system and hence provide a lower entry point for Power10 servers. To enable the use of these industry standard DIMMs the controller which provides the interface between the Power10 OMI interface and the DDR memory slots is soldered onto the system planar. There are four DIMM slots available on the system board.

Note: While Active Memory Mirroring is available in other Power10 Scale Out servers, it is not supported by the Power S1012. However, Active Memory Expansion – available on AIX and VIOS LPARs – uses compression to logically expand the memory on an LPAR and is supported on the Power S1012.

DIMMs supported

The Power S1012 provides a high bandwidth buffered memory architecture supporting up to 102 GB/s peak memory bandwidth per socket. The there are four DIMM slots on the system board, each of which supports a DDR4 memory DIMM. The maximum memory is 256 GB per server. Table 2-3 details the memory feature codes available on the system.

Table 2-3 Memory feature codes supported on the Power S1012

Feature Code	DIMM Size	DRAM Speed	Memory Bandwidth (Per Socket)
EMBN	2x 16 GB 2U DIMM	3200 MHz	102 GB/s
EMBW	2x 32 GB 2U DIMM	3200 MHz	102 GB/s
EMBY	2x 64 GB 2U DIMM	3200 MHz	102 GB/s
EMBP	Active Memory Expansion[a]	n/a	n/a

a. Add on feature

DIMM plug rules

DIMMs must be installed in pairs. Each feature code shown in Table 2-3 delivers two DIMMs. You can populate either two slots or four slots. Table 2-4 shows the memory DIMM placement rules.

Table 2-4 DIMM placement diagram

DIMM slot	P0-C6	P0-C7	P0-C9	P0-C8
First Pair	x		x	
Second Pair	x	x	x	x

Important: All DIMMs within a single system must be the same type.

Figure 2-9 shows the DIMM slot locations described in Table 2-4 on page 24.

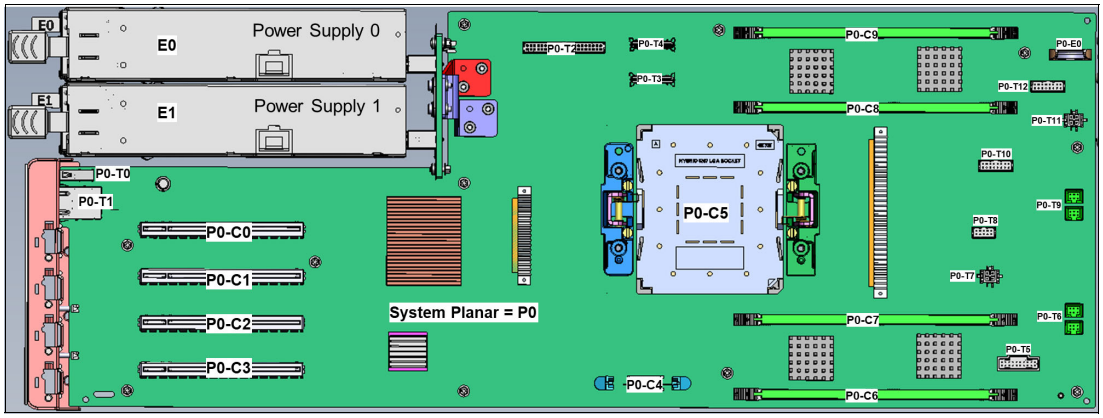


Figure 2-9 System board with location codes

2.3.4 Storage

Internal storage for the Power S1012 is provided by up to four NVMe drives installed in the chassis. There is no SAS backplane supported on the Power S1012 and there is no support for any SAS or NVME storage drawer.

There is the option to attach devices via the USB ports provided on the system board including support for the RDX device docking station. External SAN storage is supported, if desired, through the use of the appropriate SAN cards – either Fiber Channel or Ethernet connected.

Internal storage options

There is one PCIe Gen4 switch on the system planar to support the connection of the NVMe drives and the USB controller to the processors. Up to four NVMe drives are supported on the NVMe riser card installed in the system. The system also supports an internal RDX drive attached via the USB controller. In the rack mount configuration there is an external RDX enclosure which can be installed in the space adjacent to the system in the rack.

NVMe Drives

The Power S1012 has four NVMe drive slots. The drive communications and power are provided through cables to the system board. High speed Oculink cables are used for the PCIe and control signals.

The NVMe slots in the riser card are 15mm slots which support both 15 mm NVMe U2 drives and 7 mm NVMe U2 drives using a 15 mm carrier. Each NVMe drive can be independently assigned to an LPARs for use as a boot disk or as data storage. The NVMe U2 drives support concurrent add, remove, and replace.

The following NVMe drives are supported in the Power S1012:

- 0.8TB 4K U.2 15mm 18W PCIe Gen4 Enterprise Class
- 1.6TB 4K U.2 15mm 18W PCIe Gen4 Enterprise Class

The NVMe drives are installed in the front of the system as shown in Figure 2-10.

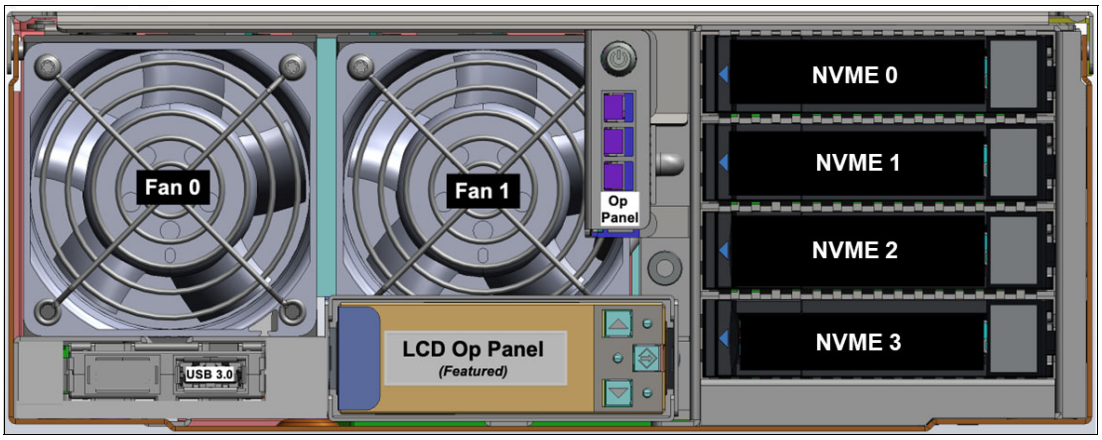


Figure 2-10 NVMe drive placement

2.3.5 RDX drive

The Power S1012 supports the connection of an RDX disk enclosure for additional internal storage and backups. The RDX can be installed internally in the tower/desk side format and can be installed in a rack enclosure for the rack mounted format.

The RDX is attached via the USB port on the system and supports the following media types:

- 320GB Disk Cartridge (EU08)
- 500GB Disk Cartridge (1107)
- 1TB Disk Cartridge (EU01)
- 1.5TB Disk Cartridge (EU15)
- 2TB Disk Cartridge (EU2T)

2.3.6 Host USB 3.0 ports

There are 2 USB 3.0 ports included in the system:

- 1x USB 3.0 port in the front
- 1x USB 3.0 port internal only for RDX attach

The host USB controller can be deconfigured by the customer via the Enterprise BMC System Management web interface, disabling all the host USB ports. Individual ports cannot be disabled.

2.3.7 I/O subsystem

The Power S1012 provides four PCIe slots to provide connectivity to devices and networking. There is not a LAN card on the system board, so one of the slots must be used for a network card. There are no external I/O drawers supported.

PCIe slots

There are four PCIe slots provided in the Power S1012. One slot is required for a LAN card for connectivity to the system. The other three are available for other PCIe requirements. All of the slots support Gen5 cards. As PCIe is backward compatible, these slots also support

earlier generation cards as well. Table 2-5 lists the characteristics for each PCIe slot in the system.

Important: The PCIe slots are not hot pluggable in the Power S1012.

Table 2-5 PCIe Slot Properties Overview

IO Slot Location Code	Source	PCIe Spec and Lanes	Max Data Rate	Card Size	Power Capability
P0-C0	Proc Chip 1 -E1	G5 x8 / G4 x16	32 GT/s	HHHL	75W
P0-C1	Proc Chip 1 -E0	G5 x8	32 GT/s	HHHL	75W
P0-C2	Proc Chip 1 -E1	G5 x8 / G4 x16	32 GT/s	HHHL	75W
P0-C3	Proc Chip 1 -E0	G5 x8	32 GT/s	HHHL	75W

Supported PCIe adapters

The PCIe adapters shown in Table 2-6 will be supported in the Power S1012 at general availability.

Table 2-6 Adapters supported at general availability

Adapter Feature Code	Adapter Description
EN1B	2-port PCIe Gen3 x8 Fiber Channel 32Gb/s
EN1K	2-port PCIe Gen4 x8 Fiber Channel 32Gb/s
EN2B	2-port PCIe Gen3 x8 Fiber Channel 16Gb/s
EJ2C	PCIe Gen3 x8 SAS Tape HBA 12Gb
EC71	2-port PCIe Gen4 x8 25/10/1Gb EN ConnectX-6
EC73	2-port 25Gb EN ConnectX-6 Crypto
EC2X	4-port PCIe Gen3 x8 10Gb ENET
EC2Y	4-port 1GB ENET
5260	4-port x4 1Gb ENET[a]

a. Migration support only

The Power S1012 may support a future PCIe adapter that runs hot or has strict temperature limits. In order to adequately cool these adapters, the fan floor could be set at a higher RPM when the adapters are installed in the system. Though temperature monitoring is also supported on some of the adapters, a high fan floor may still be desired to prevent oscillations in fan speeds. Higher fan floors can result in noticeable increases in system noise. Adapter support will likely change after the Power S1012 general availability. As adapters are added or removed from support on Power S1012, the most up-to-date list can be found in this IBM Support document.

Firmware slot capabilities

Each adapter in the system is allocated a direct memory address (DMA) space to enable communication from the PCIe adapter and the system. In the Power S1012 PCIe slots are allocated DMA space using the following algorithm:

- All slots are allocated a 2GB default DMA window.
- All I/O adapter slots (except the embedded USB) are allocated Dynamic DMA Window (DDW) capability. The amount of DDW allocated is based on the amount of memory installed in the server. DDW capability is calculated assuming 4K I/O mappings:
  - For systems with less than 64GB of memory, slots are allocated 16GB of DDW capability.
  - For systems with at least 64GB of memory, but less than 128GB of memory, slots are allocated 32GB of DDW capability.
  - For systems with 128GB or more of memory, slots are allocated 64GB of DDW capability.
- Slots can be enabled with Huge Dynamic DMA Window capability (HDDW) using the I/O Adapter Enlarged Capacity setting in ASMI.
  - HDDW enabled slots are allocated enough DDW capability to map all of installed platform memory using 64K I/O mappings.
  - Minimum DMA window size for HDDW enabled slots is 32GB.
  - Slots that are HDDW enabled will be allocated the larger of the calculated DDW capability or HDDW capability.

2.4 Physical planning

The Power S1012 is supported in two form factors: rack mount and tower (desk side). There is support for converting the rack mount to a tower form factor or converting a tower to a rack mount form factor. These two form factors are shown in Figure 2-11.

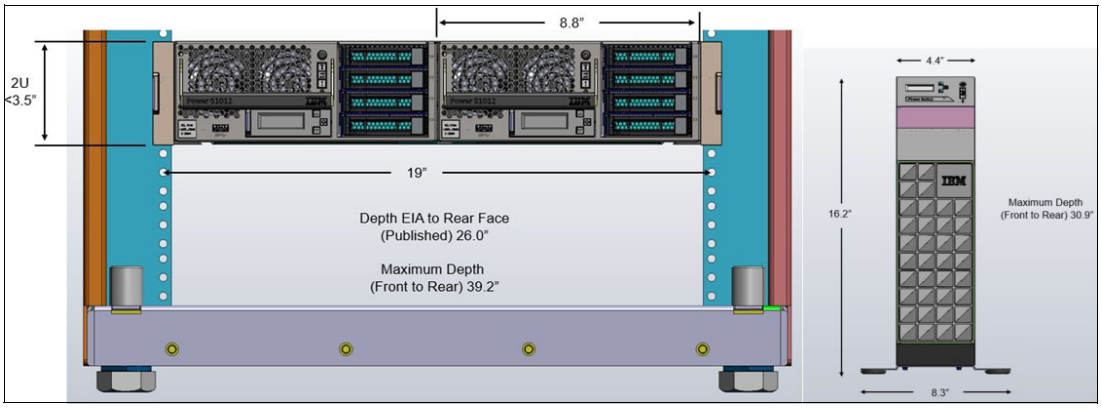


Figure 2-11 Rack and tower form factors

Factory integration for the Power S1012 is supported in the 7965-S42 rack and field integration is supported in previous 7014-T00 and 7014-T42 enterprise racks.

Table 2-7 on page 29 provides details on the physical dimensions of the Power S1012 in its different form factors – a single system, a single system in a rack enclosure and a tower system.

Table 2-7 Physical dimensions of rack mount and tower form factors

	Single system	Single system (in Rack)	Tower
Max weight (lbs)	25	45	54 (est)
Height (in)	3.5	3.5	16.2
Width, overall (in)[a]	8.8	19	4.4 (8.3 with tip plate)
Depth (in)[b]	26	30.2	30.9
Rack mountable	Yes	Yes	No
Ferrous materials	Yes	Yes	Yes

a. Width, overall is measured to the outside edges of drawer bezels and tower pedestal. Width, within rack is measured for main chassis, which fits between rack EIA flanges.
b. Depth for drawer: From rack EIA outside surface to back PCIe tailstock surface of the system. Depth for tower: From front edge of the bezel to back edge of the pedestal.

2.4.1 Power subsystem

The Power S1012 system is designed to operate within a power range supported by 1x 800W power supplies when fully configured. All systems will ship with 2x 800W power supplies for redundancy.

2.4.2 Power Supplies

The power supplies supported in the Power S1012 are 800 Watt AC supplies with Card edge connector. The 800 W power supply is a Common Redundant Power Supply (CRPS) Titanium power supply which supports input voltages of 100-127VAC or 200VAC-240VAC. The power supplies are concurrently maintainable.

The Power S1012 system has redundant power supplies in a 1+1 configuration (1 required, 1 redundant). Both supplies are required and redundant supplies are always installed at the factory. The systems will continue to run with the loss of the redundant power supplies.

The power supplies connect directly to the power distribution board – using a card edge connection – which is bolted down to the system planar. Each power supply unit (PSU) has an interlock mechanism that prevents the PSU from being removed from the chassis while its line cord is connected. This insures that input power is removed from the PSU prior to it being removed from the chassis. Communication between the integrated controller in the power supply and the eBMC occurs via an I2C bus. Diagnostics are implemented by eBMC POWR code. Diagnostic LEDs are located within the power supply chassis, visible from the rear. The power supply uses a bi-color LED; Amber and Green. Table 2-8 shows the LED states for each power supply operating state.

Table 2-8 Power supply LED indicators

Power Supply Condition	LED State
Output ON and OK.	Green
No AC power to any power supply.	OFF
AC present but only 12VSB on. The Power supply is off or is in *Smart On* state.	Blink Green (1 Hz)

Power Supply Condition	LED State
The AC cord is unplugged or AC power is lost while the second power supply is operational (with AC power connected).	Amber
Power supply warning events such as: ► high temperature ► high power ► high 12V current ► slow fan The power supply continues to operate	Blink Amber (1 Hz)
Power supply critical event causing main output power to shutdown.	Amber
Power supply firmware updating	Blink Green (2 Hz)

2.4.3 Thermal Design and Control

The Power S1012 is designed and tested to meet the ASHRAE A3 specifications. In some cases, the system configuration or performance may be reduced to stay within the thermal envelope and protect the system from thermal damage. Power10 modules run with Workload Optimized Frequency (WOF) enabled by default.

New on POWER10 processors, each processor module VPD contains a #v record with the static temperatures for the core, IO, and MMA regions. The core temperature is used for ramping fans, throttling, and error call-outs. The IO temperature is used for ramping fans.

Air Movement Devices

Two counter-rotating 60x76mm fans and the power supply fans are used to cool the system. The fan speeds are controlled by the firmware in conjunction with the eBMC. The power supply fans are set once by firmware at a high enough speed to cool a portion of the system. Thermal control loops are running to protect processors, memory, I/O, disk drives and voltage regulator modules. The loops control fan speed and force power reduction actions where needed. Errors are logged and components are called out when the temperatures cannot be maintained within the thermal limits.

Ambient Temperature Thresholds

There are three ambient temperature thresholds that generate error log events and power recovery actions. These limits are defined as:

Warning Limit The ambient warning temperature is set at the ambient operating limit. The ambient temperature should be reduced if it exceeds this threshold.

A customer notification will occur when the temperature exceeds this threshold. A second notification should occur when the warning condition is reset – the temperature falls below the threshold.

Critical Limit The critical temperature limit is the temperature at which the system should not continue normal operations and should be shutdown. Power control generates an EPOW3 indicating the threshold has been exceeded and if the ambient temperature does not fall back below the critical ambient reset threshold, then the power control system forces

power off after a defined number of minutes without any additional notifications.

A customer notification will occur when the temperature exceeds this threshold. A second notification should occur when the warning condition is reset – the temperature falls below the threshold.

Maximum Limit The maximum ambient temperature is the level at which the system is not allow to operate and must be shutdown immediately. Power control generates an EPOW4 and forces power off in a defined number of seconds.

These limits are summarized in Table 2-9.

Table 2-9 Temperature limit summary

Limit	Temperature Limit	Ambient Temperature Limit
Warning Limit	45° C	42° C
Warning Limit reset	42° C	39° C
Critical Limit EPOW3 – force power of in ~15 minutes	53° C	50° C
Critical Limit reset	48° C	45° C
Maximum Limit EPOW4 – force power off in ~20 seconds	58° C	55° C

3

Power S1012 Use Cases

The Power S1012 can be a valuable asset in your infrastructure for a myriad of uses. As it is based on the same technology as all of the other Power10 servers, it provides a unique mix of enterprise availability, flexibility and performance that is not found in the x86 based servers that you might otherwise choose to fit in this entry level class. In this chapter we describe some specific areas where the Power S1012 can provide unique value for your enterprise Information Technology infrastructure.

This chapter describes the following use cases:

- “IBM i entry system”
- “Edge computing”
- “Entry database solutions”
- “Advanced configuration options”

3.1 IBM i entry system

IBM i – and its predecessors – has a long history of providing benefits to users. At its core, the IBM i platform is designed to adapt to the ever-changing needs of both business and computing. Its defining characteristic, the “integration” represented by the “i” in IBM i, can help you gain more value from advanced technology with fewer resources and higher reliability

3.1.1 How IBM i benefits your business

Many thousands of companies around the world rely on IBM i because they want a more resilient, more secure and more cost-efficient alternative to Microsoft Windows technology-based servers for their most important business data and applications.

Midsized companies in particular have two key requirements: to maximize their IT investments and to exploit these investments as the company’s requirements grow. Unlike Windows technology-based servers, the IBM i operating environment is almost always used to run multiple business applications and databases securely and efficiently on the same server. As a result, clients report that they have fewer servers to manage with IBM i when compared to Windows. This optimization of assets helps companies avoid the costs of acquiring and managing a new server every time the business needs to deploy another application. Ease of deployment, upgrades and management gives IBM i a significant advantage when evaluating the total cost of operation (TCO).

The IBM i business is known for its adaptability. This has allowed IBM to invest in making IBM i a platform that provides a strong foundation for technical and business innovation – both for IBM and for clients. IBM i provides a secure integrated system that has been used by many independent software vendors (ISVs) as a base for their applications which include financial services, health care, retail management, wholesale distribution and insurance (to name a few). Why has IBM i been so successful in this wide variety of industries? We provide some of the reasons here.

Integration

The defining characteristic of IBM i is the integration of software solutions into operating system to enhance its adaptability for industry and business need, giving you more value with fewer resources and a highly reliable system. It is fully integrated meaning the database, middleware, security, runtime and hypervisor are integrated into the stack and licensed as one solution.

Adaptability

The adaptability of IBM i has allowed IBM to invest in making the platform provide a strong foundation for technical and business innovation. Applications can be integrated with new technologies (AI, IOT, data analytics for example) to add value and opportunities for applications.

Lower TCO

IBM i is a fully integrated operating system, meaning the database, middleware, security, runtime and hypervisor are integrated into the stack and licensed as one solution. Clients don't have to pay for support on each individual component, thereby lowering total cost of ownership (TCO) and simplifying the platform running your business applications. Autonomous features require little to no administration, and can be managed remotely from any device with IBM Navigation for i.

Nearly 2X ROI

Forrester Consulting found that clients deploying IBM i on premise or in the cloud realized – on average – a 191% return on investment (ROI) and a payback of just 6 months for their business. These clients also saw savings of $1.06 million in reduced system downtime costs and an increase in productivity of $470,000 over three years[1].

Security

Authority collection capabilities enable you to track and monitor who is using which objects and how, which is critical for determining security policies and rules. These capabilities provide enhanced security for your most business-critical and sensitive applications and data. IBM i is a Object-based architecture, meaning that a program cannot masquerade as something else. This protects the environment from viruses, and security breaches.

Flexible Licensing and Support

One annual payment provides both software licenses and SWMA together in a single subscription price. This provides lower TCO and increases flexibility.

Simplified Systems Management

IBM i has autonomous features that require little to no administration, reducing the need for large system administration teams. Manage IBM i environments remotely from any device with IBM Navigator for i and IBM i Mobile Access.

3.1.2 Power S1012 Benefits for IBM i customers

The Power S1012 system is ideal for the small to midsize shop that needs all of the benefits that IBM i has to offer – in a smaller package. Many IBM i customers have multiple locations (stores, distribution centers, warehouses) where there is not a “data center” and where space is a premium resource. The Power S1012 has the security, reliability and performance that are required for the workloads that are run in those locations, all in a nicely compact footprint (either rack mounted or tower) and at a very competitive price point.

Customer example 1

Our first customer example is a company with a large number of small stores. The requirement in each store is one or two processor cores and less than 128 GB of memory. A single 4 port network card and mirrored NVMe drives complete the configuration. These requirements easily fit within the Power S1012 with a reduced footprint compared to other options. The flexibility of either rack-mount or tower configurations, and the ability to run with standard 110 Volt power outlets add additional benefits and cost savings. Support for IBM i provides the ease of management and integrated solution to which the customer is accustomed.

Customer example 2

Customer is running a small warehouse where physical compute is relegated to a small space. They require the availability that IBM i is known for. CPU workload is consistently in the 20-40% range. If they require HA, the solution fits within 2U space in a Rack.

Customer example 3

Multiple similar IBM i environments in a distributed environment. In this example a chain retailer runs a small IBM i environment in every store. They are identical in software and configuration. CPU usage is consistent and predictable. The Power S1012 includes four NVMe drives – mirrored for availability which provides sufficient storage.

[1] https://www.ibm.com/account/reg/signup?formid=urx-52179

Applications include:

- Inventory management
- Ordering
- Signage
- Hand-held RF gun support
- Point of sale.

The environment includes a single thin client with two network interface ports. Port one is directly connected into the Power S1012 and is used for console connectivity by service personal only. Port two is connected to the store network where it is used as an interactive client for the IBM i system.

At the POS an item is scanned, the price is looked up and verified, inventory is adjusted, and an order is placed to the distribution center for replenishment.

3.2 Edge computing

The explosive growth and increasing computing power of IoT devices has resulted in unprecedented volumes of data and those data volumes continue to grow as 5G networks increase the number of connected mobile devices. In the past, the promise of cloud and AI was to automate and speed up innovation by driving actionable insight from data. But the unprecedented scale and complexity of data that is created by connected devices has outpaced network and infrastructure capabilities.

Sending all device-generated data to a centralized data center or to the cloud causes bandwidth and latency issues. Edge computing offers a more efficient alternative; data is processed and analyzed closer to the point where it's created. Because data does not traverse a network to a cloud or data center to be processed, latency is reduced. Edge computing – and mobile edge computing on 5G networks – enables faster and more comprehensive data analysis which creates the opportunity for deeper insights, faster response times and improved customer experiences.

Edge computing is a distributed computing framework that brings enterprise applications closer to data sources such as IoT devices or local edge servers. This proximity to data at its source can deliver strong business benefits, including faster insights, improved response times and better bandwidth availability. Characteristically is distributed, software defined, and flexible. The value of Edge is the movement of computing resources to the physical location where data is created, transacted or stored, thereby increasing enablement of business processes, decisions, and intelligence outside of the core IT environment. Figure 3-1 on page 37 shows an example of Edge computing.

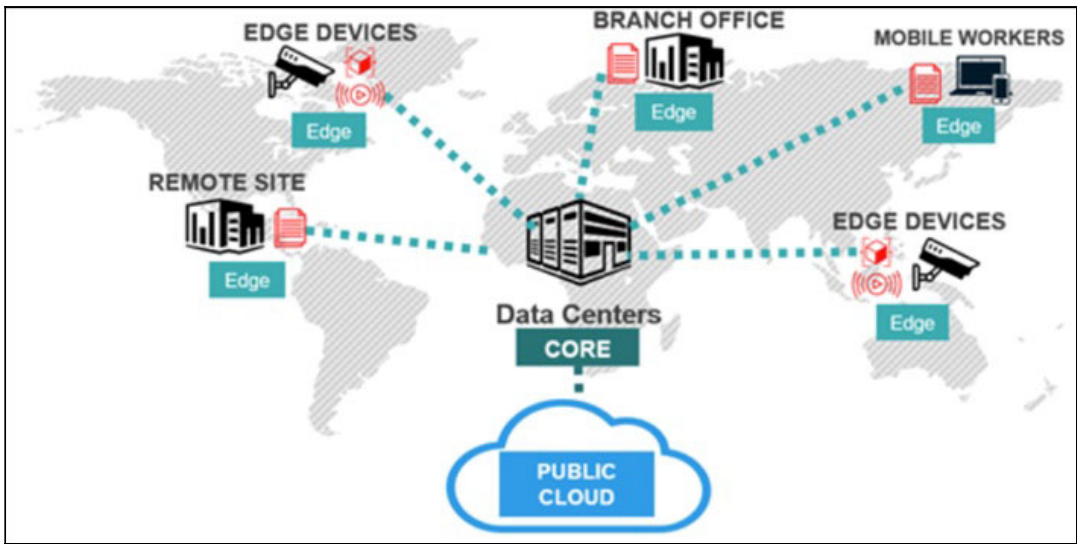


Figure 3-1 An example Edge and Core environment

Power S1012 benefits for Edge Computing with AI

The Power S1012 provides several features that make it an excellent choice for an Edge Computing environment. The Power S1012 provides an entry level compute platform that brings benefits designed into the Power10 chip such as:

- Built in AI acceleration circuitry to run AI workloads without the requirement for expensive GPU cards.
- Built in reliability, availability and serviceability options to provide a highly available AI platform.
- Remote management capability through the integrated eBMC card.

The small size and flexibility of the Power S1012 make it an ideal platform for Edge computing locations with limited space and with minimal power requirements. The ability to be either rack mounted or configured as a standalone tower mode makes the Power S1012 a good candidate for use:

- In remote office, stores, manufacturing locations
- In manufacturing equipment management environments with IoT
- In new locations with new applications designed for Edge computing
- As an alternative to potentially expensive cloud deployments for Edge computing environments.

There are many additional use cases to consider as you look at future Edge computing applications. Power10 customers can benefit from Edge/Remote Computing in:

- Endpoint Security and Surveillance
- In-store Point-of-sales customer experience (scoring)
- Edge/Remote container-based applications
- Health care patient monitoring
- Manufacturing equipment management with IoT (IBM Maximo®)
- Utilities: power grid optimization, reduce energy waste
- Embedded solutions for OEM/ESA contracts. Explore opportunities with Linux-based ISVs focused on ROBO deployments

3.3 Entry database solutions

One challenge for an enterprise is providing the Information Technology resources needed by the end users while staying inside budget parameters. This is especially evident when the applications require a database solution for small databases.

In today's world, keeping data safe and secure can be a challenge. IBM Power10 servers, including the Power S1012 provide many functions that help do that. With built-in encryption features, data can be encrypted within the database without affecting performance. In addition, IBM Power has been shown to be more secure and has significantly fewer security vulnerabilities which often lead to security breaches. The Power S1012 provides an excellent solution for entry level database solutions, providing a platform that is secure, reliable and can provide the performance needed to keep users satisfied – all in a footprint that reduces the space requirements within the infrastructure.

When using a commercial database, it is common that the database vendor will charge based on the size or class of the machine that is hosting the database. This can be especially difficult to manage when the database workload is a small part of your overall application, but the expenses related to the third party database are significant.

One option that might be considered is the use of open source database solutions. While this will likely provide you with savings for the database software, it brings along additional difficulties as changes may need be made in your applications due to different requirements for the database. This could bring additional costs of training and application rework into your application environment which could counteract the savings achieved from choosing the open source solution.

Another, possibly better, option is to take advantage of special offerings from the database supplier that provide an entry level database solution running on a smaller machine at a reduced cost point. For Oracle, this offering is known as the Oracle Database Standard Edition.

3.3.1 Oracle Database Standard Edition 2 on Power S1012

Databases have been a critical component in keeping businesses running for many years. Many enterprises have turned to Oracle to provide the database technology that they require. The challenge is that for small to medium size applications, the cost of the database is a significant part of the total application cost and the number of databases is expanding rapidly as new applications come online.

Oracle has multiple options for running your database. For high end databases, the best option is using Oracle Enterprise Edition which has all of the features to support your enterprise class applications. For smaller databases, Oracle has another option that can save up to 33% of the costs on each DB instance, Oracle Standard Edition 2 (SE2).

The savings opportunity when using Oracle SE2 comes from the fact while Oracle Enterprise Edition is charged per core (based on a core factor for the processor type being used) Oracle SE2 is charged per socket, no matter how many cores are provided per socket. For consolidating a number of smaller databases, Oracle SE2 is a good option.

There are some restrictions involved with running Oracle SE2. The first one is that it is limited to servers with a maximum of two sockets. Oracle considers a single Power10 DCM to be two sockets, the Power10 servers that are eligible to run SE2 are the Power S1012 and the Power S1014. It is also important to understand that another restriction for using the SE2 is that each database is limited to a maximum of sixteen threads. As the Power10 chip is designed

with SMT8, where each core provides eight strong threads, this is even a stronger reason to consider consolidating multiple databases to a single Power10 server.

The Power S1012 brings all of the benefits of Power10 to the Oracle environment. With the built in power hypervisor, consolidation of multiple servers is easy and efficient and does not have the overhead of software virtualization products as seen in the x86 world. Add to that the proven reliability and security of the Power platform over time and this brings even more advantages compared to x86 alternatives.

Power10 adds additional benefits with its built in transparent memory encryption, further adding to the security of enterprise critical databases. If you are looking to add AI capabilities, Power10 provides built in AI inferencing capability as well.

For smaller environments, the Power S1012 with its 8-core processor might be a good fit and could replace older Power8 and Power9 servers currently running SE2. With the 2.8x performance per core advantage of Power10 over x86 options, this might also be a good option for upgrading any x86 based SE2 implementations.

3.4 Advanced configuration options

The form factor of the Power S1012 provides a unique opportunity to create a single rack high availability solution for your smaller workloads. When you install a single Power S1012 server into a rack, there is an available space next to that server where you could install an additional Power S1012 server.

3.4.1 Single rack high availability solution

Even though the Power S1012 has excellent reliability and availability characteristics, no server is 100% available and you should have a plan for any time where your server might be unavailable. Having a secondary server available to run your workloads while the primary server is unavailable – due to either a planned or unplanned outage – can now be done much more economically with the Power S1012.

To enable this solution, two separate AC sources should be available within the rack so that a unique AC power source is provided for each server in this high availability pair. Additionally you should ensure that your users can access the secondary server when the primary server is down using the appropriate network connections. You also need to ensure that the data required for your applications to run is available to the secondary server. Figure 3-2 illustrates this concept.

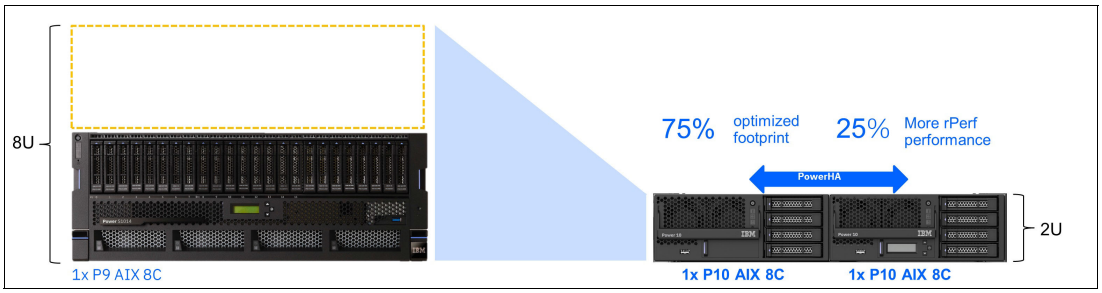


Figure 3-2 High availability solution within 2U rack space

IBM has several HA solutions available that will assist you in managing your server environment and quickly switch between the active and backup server with a minimal outage

time. IBM Virtual Machine Recovery Manager is an easy-to-use tool that allows you to restart your system (no matter what operating system you are running) on a backup server. Another option for AIX and IBM i systems is IBM PowerHA® SystemMirror®, or for Linux systems you can use Pacemaker. For further information on IBM's High Availability offerings see the following:

- *IBM Power Systems High Availability and Disaster Recovery Updates: Planning for a Multicloud Environment*, REDP-5663
- *IBM PowerHA SystemMirror and IBM VM Recovery Manager Solutions Updates*, REDP-5694
- *IBM Virtual Machine Recovery Manager for IBM Power Cookbook*, SG24-8539
- *IBM PowerHA SystemMirror for AIX Cookbook*, SG24-7739

3.4.2 Enhanced IT management capability

The Power Hardware Management Console (HMC) provides many benefits for managing your IBM Power servers, but many enterprises have chosen not to implement them as they require an additional server to host the HMC.

With the unique form factor of the Power S1012, it is possible to install an additional Power S1012 in the adjacent slot to your production server which can be utilized as an HMC in your environment allowing you to take advantage of the additional management and monitoring tools built into the HMC. This is illustrated in Figure 3-3.

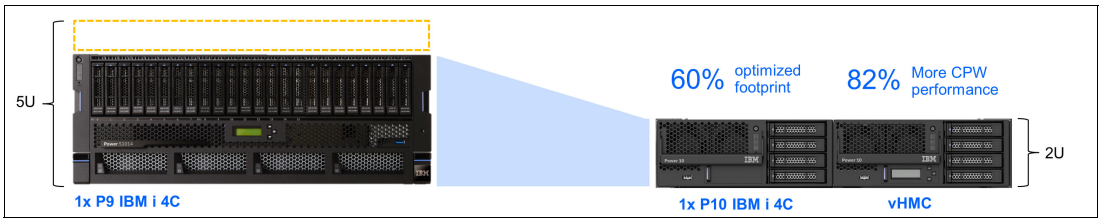


Figure 3-3 Enhanced management capabilities with HMC

Related publications

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

IBM Redbooks

The following IBM Redbooks publications provide additional information about the topic in this document. Note that some publications referenced in this list might be available in softcopy only.

- *IBM Power S1014, S1022s, S1022, and S1024 Technical Overview and Introduction*, REDP-5675
- *IBM Power E1080 Technical Overview and Introduction, REDP-5649*
- *IBM Power E1050: Technical Overview and Introduction*, REDP-5684
- *IBM Power Systems High Availability and Disaster Recovery Updates: Planning for a Multicloud Environment*, REDP-5663
- *IBM PowerHA SystemMirror and IBM VM Recovery Manager Solutions Updates*, REDP-5694
- *IBM Virtual Machine Recovery Manager for IBM Power Cookbook*, SG24-8539
- *IBM PowerHA SystemMirror for AIX Cookbook*, SG24-7739

You can search for, view, download or order these documents and other Redbooks, Redpapers, Web Docs, draft and additional materials, at the following website:

ibm.com/redbooks

Online resources

These websites are also relevant as further information sources:

- What is edge computing?
  https://www.ibm.com/topics/edge-computing?utm_content=SRCWW&p1=Search&p4=43700074715395988&p5=e&gad_source=1&gclid=CjwKCAjw8diwBhAbEiwA7i_sJci52c8JsoQ49vX_KgVhUZvjfL4Bby3w9OCkRyHzSH5rDE_6xrdiVBoCoqYQAvD_BwE&gclsrc=aw.ds

Help from IBM

IBM Support and downloads

ibm.com/support

IBM Global Services

ibm.com/services

REDP-5728-00

ISBN DocISBN

Printed in U.S.A.