Redbooks
ibm.com/redbooks

SAP HANA on IBM Power Systems Architectural Summary

Corie Neri

Tim Simon

Henry Vo

John F. Aristizabal

Neelabha Banerjee

Youssef Largou

Analytics

Power Systems

IBM

Redpaper

IBM Redbooks

SAP HANA on IBM Power Systems Architectural Summary

June 2024

REDP-5569-01

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

Second Edition (June 2024)

This edition applies to:

- SAP HANA 2.0 SPS06
- SAP S/4HANA 2022 Release
- Red Hat Enterprise Linux 8.4
- SUSE Linux Enterprise Server for SAP Applications 15 SP4

Contents

Notices

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

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

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

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

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

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

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

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

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

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

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

COPYRIGHT LICENSE:

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

Trademarks

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

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

AIX®
Cognos®
Db2®
FlashCopy®
IBM®
IBM Cloud®
IBM Consulting®
IBM FlashSystem®
IBM Spectrum®
IBM Z®
Micro-Partitioning®
Power®
POWER®
Power Architecture®
Power8®
Power9®
PowerHA®
PowerVM®
Redbooks®
Redbooks (logo) ®
System z®
Tivoli®
WebSphere®

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

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

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

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

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

2015 SUSE LLC. All rights reserved. SUSE and the SUSE logo are registered trademarks of SUSE LLC in the United States and other countries.

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

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

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

Preface

This IBM® Redpaper publication delivers SAP HANA architectural concepts for successful implementation on IBM Power® Systems servers. This update is designed to introduce the Power10 product line and how it enhances support for SAP HANA. The update includes information that SAP HANA now supports Red Hat Enterprise Linux as a supported operating system.

This publication addresses topics for sellers, IT architects, IT specialists, and anyone who wants to understand how to take advantage of running SAP HANA workloads on Power Systems servers. Also, this guide provides documentation to transfer how-to skills to the technical teams, and it provides solution guidance to the sales team. This publication complements documentation that is available at IBM Documentation, and it aligns with educational materials that are provided by IBM Systems.

Authors

Corie Neri is an IBM Redbooks® Project Leader in Phoenix, Arizona, US. He has over 25 years of experience with IBM primarily designing and developing microcode, and providing the documentation to implement the procedures to upgrade IBM VTS and Peer-to-Peer configurations, which evolved into becoming the current Hydra (TS7700) environment. He also provided 24 x 7 technical support and became a subject matter expert (SME), where he informed the Top Guns on the year-to-year changes in the annual Top Gun Seminar. He also attended the onsite installation for the first-time install of the Miscellaneous Equipment Specification (MES) procedure that he designed, so he was available for immediate support and any modifications to optimize the process to minimize the impact to the customer. Before becoming a Redbooks Project Leader for StorM and Open Power, he was a Data Scientist for 6 years. As a Data Scientist, he analyzed TS7700 and TS4500 field records that were possible impacts. He helped determine whether the cause of the issue was already solved in a newer version of code that must be installed or if the issue was a new problem, which needed attention to be raised to development to avoid it from reoccurring. He also helped provide new rules for the automated process for either adding a new rule to add it to the possible impact list or to skip a phrase from being added to the possible impact list to optimize the team's manual review process. He holds a BS degree in Computer Science from Arizona State University.

Tim Simon is a Redbooks Project Leader in Tulsa, Oklahoma, US. 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 shares his technical expertise in business problem solving, risk/root-cause analysis, and writing technical plans for business. His previous roles at IBM include Project management, ST/FT/ETE Test, Back-End Developer, DOL agent for NY. He is certified for IBM zOS Mainframe Practitioner, IBM Z® System programming, Agile, and Telecommunication Development Jumpstart. Henry holds a Master of MIS (Management Information System) from the University of Texas at Dallas.

John F. Aristizabal is an IT Specialist Consultant and the founder of BC SUPPORT in Colombia. With 23 years of experience in the technology industry, he has spent the last 16 years as a certified SAP NetWeaver and SAP HANA consultant. Throughout his career, John has successfully worked on various customer projects, tackling challenges such as installations, upgrades, and database migrations to SAP systems on platforms such as IBM Power, HP-UX, and Intel. He has extensive expertise in working with diverse databases, which include IBM Db2®, Db2/400, Oracle, Sybase, SQL Server, and SAP HANA.

Neelabha Bannerjee has 20 years of experience working as a Technical Architect for SAP Landscape. For the last 18 years as part of SAP Practice at IBM Consulting®, he has been delivering and consulting for IBM clients across the globe. He has multiple SAP Certifications. His areas of expertise include SAP on Hybrid Cloud, SAP Implementation (Greenfield), SAP S/4HANA conversion (Brownfield), Technical architecture design, SAP OS/DB Migration, SAP system consolidation, and SAP Upgrade. He holds a Bachelor of Technology degree in Computer Science & Engineering.

Youssef Largou is the founding director of PowerM, a platinum IBM Business Partner in Morocco. He has 22 years of experience in systems, HPC, middleware, and hybrid cloud, including IBM Power, IBM Storage, IBM Spectrum®, IBM WebSphere®, IBM Db2®, IBM Cognos®, IBM WebSphere Portal, IBM MQ, ESB, IBM Cloud® Pak, SAP HANA and Red Hat OpenShift. He has worked within numerous industries with many technologies. Youssef is an IBM Champion 2020, 2021,2022 and 2023, an IBM Redbooks Platinum Author, and has designed many reference architectures. He has been recognized as an IBM Beacon Award Finalist in Storage, Software-Defined Storage, and LinuxONE five times. He holds an engineering degree in Computer Science from the Ecole Nationale Supérieure des Mines de Rabat and Excecutif MBA from EMLyon.

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

Pete Heyrman
IBM PowerVM® Hypervisor, IBM Rochester, US

Dr. Edmund Häfele
Senior IT Architect, IBM Germany

Wade Wallace
IBM Redbooks®, Austin Center

Katharina Probst, Walter Orb, Tanja Scheller
IBM Germany

Thanks to the authors of the previous edition:

- **Dino Quintero**
- **Damon Bull**
- **Vinicius Cosmo Cardoso**
- **Cleiton Freire**
- **Eric Kass**

Now you can become a published author, too!

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

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

ibm.com/redbooks/residencies.html

Comments welcome

Your comments are important to us!

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

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

  ibm.com/redbooks

- Send your comments in an email to:

  redbooks@us.ibm.com

- Mail your comments to:

  IBM Corporation, IBM Redbooks
  Dept. HYTD Mail Station P099
  2455 South Road
  Poughkeepsie, NY 12601-5400

Stay connected to IBM Redbooks

- Find us on Facebook:

  http://www.facebook.com/IBMRedbooks

- Follow us on Twitter:

  http://twitter.com/ibmredbooks

- Look for us on LinkedIn:

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

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

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

- Stay current on recent Redbooks publications with RSS Feeds:

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

1

Solution approach for SAP HANA on IBM Power Systems

This chapter provides IBM POWER® architectural details and illustrates the advantages of deploying SAP HANA on IBM Power servers.

This chapter covers the following topics:

1.1 SAP HANA Overview

SAP High Performance Analytic Appliance (SAP HANA) was created as the first In-Memory database for SAP and offered early customers in-memory computing, real-time data processing and analytics capabilities with features such as columnar storage, parallel processing, and distributed computing architecture for greater performance and scalability.

The evolution of SAP HANA has transcended its initial identity as a stand-alone database. It has matured into a seamlessly integrated platform that is renowned for its multifaceted capabilities, including robust database management, sophisticated application development, and advanced analytics functions.

Table 1-1 lists common SAP HANA use cases.

Table 1-1 Common SAP HANA Use case

Use case	Description
Real-time analytics	Enables real-time data analytics so that organizations can make decisions based on up-to-the-moment information. This is especially useful in business environments that require quick responses, such as fraud detection, risk analysis, inventory management, and personalized customer experiences.
Planning and forecasting	SAP HANA is used for enterprise planning and demand forecasting. It enables organizations to perform predictive analysis and what-if scenarios, which helps them optimize resource planning, supply chain management, and strategic decision-making.
Advanced analytics	SAP HANA provides a platform for running advanced analytics, such as predictive analysis, time series analysis, and text analysis. This allows organizations to uncover patterns, identify trends, perform customer segmentation, as well as conduct sentiment analysis to gain valuable data-driven insights.
Real-time business applications	SAP HANA serves as a platform for the development and execution of real-time business applications. This includes applications for financial management, human resources, sales and marketing, and logistics. SAP HANA accelerates the performance of these applications and enables real-time data integration for agile decision-making.
Internet of Things (IoT)	With the ability to process large volumes of real-time data, SAP HANA is used in IoT applications for data analysis and management of sensor-generated data and connected devices. This enables real-time monitoring, asset management, production optimization, and other IoT-related applications.
Supply Chain Optimization	SAP HANA can optimize supply chain management by processing vast amounts of data quickly, enhancing demand forecasting, inventory management, and logistics.
Customer segmentation	Organizations can use SAP HANA's real-time analytics to perform customer segmentation, enabling targeted marketing campaigns and personalized customer experiences.
Fraud Detection and Risk Management	SAP HANA's in-memory processing facilitates rapid analysis of transactional data, enhancing fraud detection and risk management capabilities for financial institutions and other industries.

Note: For this book, SAP HANA is mainly a database for applications that are provided by SAP and other vendors.

Since 2011, SAP HANA has gradually become the prominent SAP DB technology for most SAP products. SAP HANA is an in-memory DB that uses column-based storage to speed up data access and also provides compression to reduce the size of memory required. The data processing is faster than a Relational Database Management System (RDBMS), and the data that is stored when you migrate to SAP HANA is reduced in terms of allocated space in memory.

Initially, SAP HANA was available as an Appliance, which was optimized and preinstalled configurations that combined both hardware and SAP HANA software by different manufacturers, including IBM. The manufacturers worked closely with SAP to ensure that these systems met the requirements and performance standards, compatibility, and scalability defined by SAP. One of the disadvantages of appliances was the inability to grow granularly because growth required double the resources, which doubled the costs.

In 2015, SAP released SAP HANA Tailored Datacenter Integration (TDI). With TDI, customers can use SAP certified infrastructure that the customer had available to install SAP HANA. TDI provides greater flexibility in terms of hardware and allows customers to reduce costs when they expand the capacity of their SAP HANA system. IBM Power Systems provides an advantage to customers because it can granularly allocate resources such as CPU, memory, and disk and can dynamically scale.

As of this writing, SAP HANA offers support for Red Hat Enterprise Linux (RHEL) and SUSE® Linux Enterprise Server only. Both of these operating systems are supported on the IBM Power System platform.

> **Note:** For more information about supported operating systems, see SAP Note 2235581. Frequent validations are recommended.

1.2 IBM Power Server and IBM Power Virtual Server Overview

IBM Power Systems and IBM Power Virtual Server provide organizations with a powerful and flexible platform for running mission-critical workloads, whether on-premises or in the cloud. These offerings combine high performance, reliability, and scalability to support diverse computing needs and enable seamless digital transformation.

1.2.1 IBM Power Systems

IBM Power Systems is a family of high-performance and scalable servers designed to deliver superior processing power, reliability, and flexibility for various enterprise workloads. These systems are based on IBM POWER Architecture®, which has a strong history of serving mission-critical applications and data-intensive workloads. IBM Power Systems are known for their ability to handle complex and demanding computing tasks across different industries. IBM Power Systems includes some key features:

- IBM Power Systems are equipped with POWER processors, which are known for their high performance, parallel processing capabilities, and ability to handle large data workloads efficiently.
- These systems offer a high-memory capacity and support for fast I/O connectivity, enabling smooth processing of data-intensive applications.
- IBM Power Systems are designed for high availability and business continuity, making them suitable for mission-critical applications that require continuous uptime.

- These servers offer scalability so that organizations can add resources as needed to accommodate growing workloads without compromising performance.
- IBM Power Systems provide robust virtualization capabilities so that organizations can create and manage multiple virtual environments, which increases resource usage and reduces hardware costs.
- IBM Power Systems support various operating systems, including IBM AIX®, IBM i, and Linux, which provides the flexibility to run diverse workloads on the same hardware.
- IBM Power Systems are certified for running SAP HANA, which helps businesses that are seeking optimal performance for their SAP applications.

1.2.2 IBM Power Virtual Server

IBM Power Virtual Server is a cloud-based offering that extends the capabilities of IBM Power Systems into the cloud environment. Organizations can use it to deploy virtual IBM Power Systems instances on IBM Cloud, providing a secure and scalable solution for running Power workloads in the cloud. The IBM Power Virtual Server includes the following key features:

- IBM Power Virtual Server enables organizations to flexibly deploy virtual Power Systems instances based on their specific workload requirements. It provides options for customizing virtual machine sizes and configurations.
- With Power Virtual Server, organizations can quickly provision virtual Power Systems instances, eliminating the need for lengthy hardware procurement and setup processes.
- The cloud-based nature of Power Virtual Server allows seamless scaling of resources up or down based on demand, which helps to ensure optimal performance and cost efficiency.
- IBM Cloud ensures that robust security measures are available to protect data and workloads that are hosted on Power Virtual Server instances. IBM Cloud adheres to industry standards and compliance requirements.
- Power Virtual Server integrates with other IBM Cloud services, enabling organizations to create hybrid cloud environments and use additional cloud-based capabilities.
- By adopting a cloud-based approach, organizations can achieve cost savings by paying only for the resources they consume, avoiding upfront hardware investments, and reducing maintenance costs.

By using IBM Power, customers can deploy SAP software in containers by using Red Hat OpenShift. This provides clients with the option to use Red Hat Ansible to create SAP containers, so they can automate processes and eliminate manual errors. With this added flexibility, clients can more easily deploy multiple SAP images faster.

1.2.3 SAP and IBM partnership

The partnership between IBM and SAP is a strategic collaboration that brings together the strengths of both companies to deliver innovative solutions and value to customers worldwide. With a long-standing relationship spanning decades, IBM and SAP work together to optimize SAP's software applications on IBM's cutting-edge hardware and cloud platforms.

The IBM-SAP partnership includes several key features:

- Optimized Solutions

  IBM and SAP collaborate to optimize SAP's enterprise software, including SAP HANA, S/4HANA, and other SAP applications, on IBM's high-performance hardware and cloud infrastructure. This ensures that customers can use the full potential of SAP's solutions for data processing, analytics, and business transformation.

- Cloud Integration

  IBM's cloud offerings, such as IBM Cloud and IBM Cloud Private, provide a robust and secure platform for hosting SAP applications. This collaboration enables organizations to run SAP workloads on IBM Cloud with the assurance of industry-leading security, scalability, and performance.

- Technology Innovation

  The partnership fosters joint research and development efforts to drive technological innovation. IBM and SAP work together to explore emerging technologies, such as artificial intelligence, machine learning, blockchain, and Internet of Things (IoT), and how they can be integrated into SAP's solutions to deliver even greater business value.

- IBM Consulting

  IBM's extensive IBM consulting capabilities complement the SAP software portfolio to provide customers with end-to-end consulting, implementation, and support services. Together, they offer seamless migration strategies, industry-specific solutions, and best practices to help organizations realize the full potential of their SAP investments.

- Industry Expertise

  IBM and SAP's combined expertise spans across various industries, including banking, healthcare, retail, manufacturing, and more. This industry-focused approach allows them to tailor solutions that address the unique challenges and requirements of each sector, delivering measurable business outcomes.

- Client Success Stories

  IBM and SAP collaborated on numerous successful projects to help customers achieve digital transformation, streamline operations, and improve business agility. Their shared client success stories demonstrate the value of their joint solutions across different business scenarios. The IBM-SAP partnership continues to evolve, driven by a shared commitment to innovation, customer success, and advancing digital transformation for businesses worldwide. Together, they are poised to address the evolving needs of the digital era and help organizations thrive in an increasingly competitive and data-driven landscape.

- Knowledge Sharing

  The partnership allows for extensive knowledge sharing and sharing of ideas between the two companies. This knowledge exchange provides for continuous improvement and innovation in their joint solutions, which helps to keep them at the forefront of technology advancements.

- Strategic Alignment

  The long-term partnership demonstrates a strong strategic alignment between IBM and SAP. Both companies share a common vision of delivering value-driven solutions to customers and driving digital transformation across industries.

IBM Power Systems showcase technical leadership with numerous production databases hosted and advanced features like processor pool sharing and virtual persistent memory:

- The IBM platform facilitates digital transformation and reduces application migration and modernization costs.
- It offers a serverless-like system with flexible consumption options for various workloads from core to edge applications.

- Clients benefit from twice the compute and memory density fitting seamlessly into their data centers.
- IBM Power Systems provide industry-leading uptime, with a maximum of 20 times less downtime when compared to x86 server vendors.
- The platform enables innovation and collaboration without interruptions, ensuring seamless hybrid cloud integration for enhanced resources.
- Data security is paramount, with platform integrity from the processor to the cloud, following a zero-trust architecture.
- Data fabrics replace silos so that data is more accessible for enterprise analytics and AI-driven decision-making.
- IBM's built-in inference engine brings AI closer to the data, which reduces cost and complexity for clients.

Figure 1-1 shows the collaboration timeline between SAP HANA and IBM Power9® and Power10.

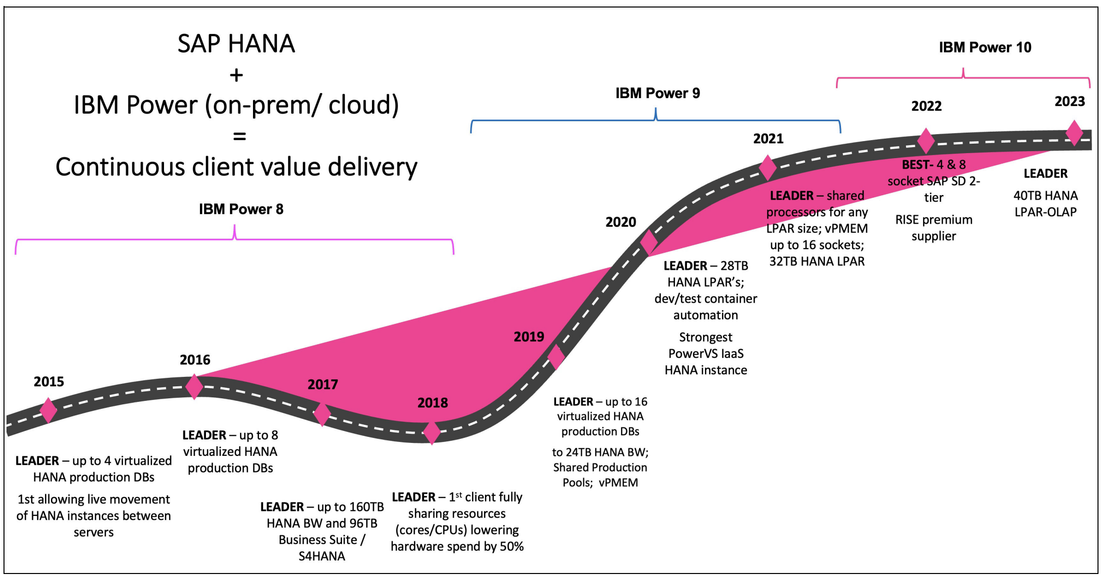


Figure 1-1 SAP HANA and Power Systems collaboration timeline

1.3 SAP HANA architecture

SAP HANA is an in-memory database and platform that is designed to handle high-performance data processing and analytics. Its architecture is designed to take advantage of in-memory computing, which enables fast data access and real-time processing.

1.3.1 Main components of the SAP HANA System

In the context of a functioning SAP HANA system, there are several interconnected processes, or services, working together. An SAP HANA system is composed of different components. The following list describes some of the main components:

- Name server

  This component runs exclusively on the system DB, and it is the directory of SAP HANA. Information about all running components can be found here, including distributed components in multi-host servers.

- Compile server

  Its main function is compiling the stored procedures and programs. It does not store information. There is no persistency. It runs on the system database and serves to all system databases.

- Preprocessor server

  It is an auxiliary server that is used by the index server to analyze text data and extract information on what the text search capabilities are based. It runs in the system DB and serves all system databases.

- Index server

  The most important component of SAP HANA. It contains the data stores and the engines to process them. The index server runs on every system database.

- XS Engine (Classic Server or Advanced Runtime)

  Allows you to write and run applications based on SAP HANA. Some of the applications or management tools run there. By default, the Classic version is installed. However, this version is deprecated, and SAP recommends migrating to the XS Advanced Runtime version.

Figure 1-2 shows an architectural diagram representation of these main components.

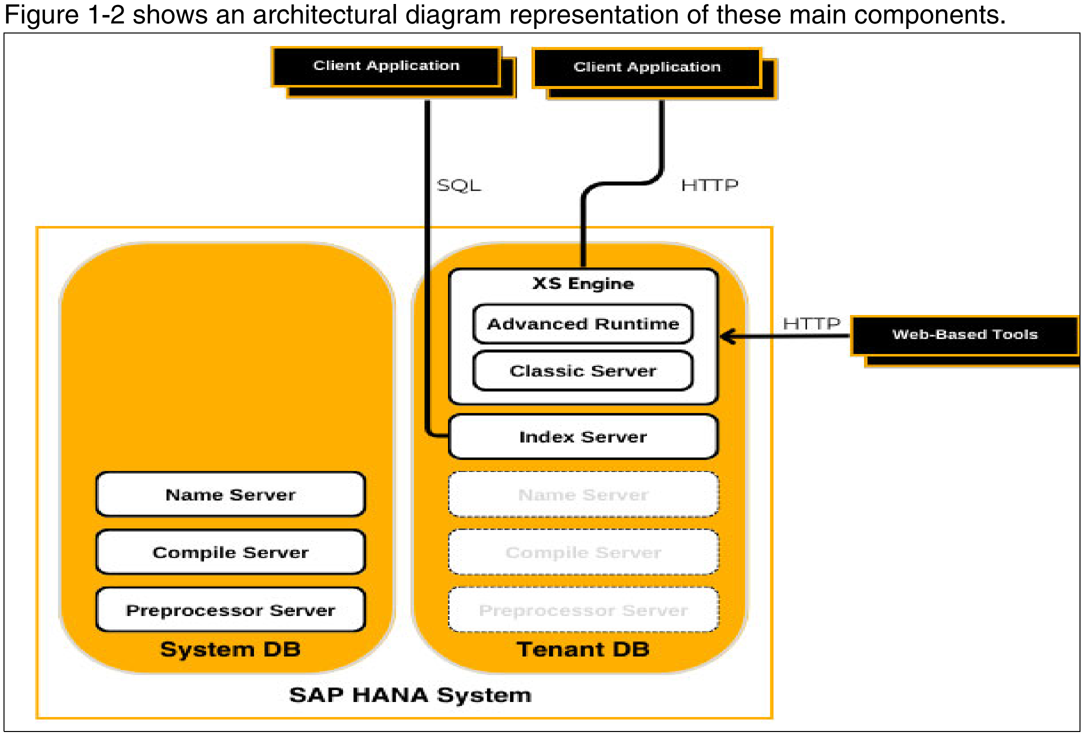


Figure 1-2 SAP HANA System: main components

It is important to understand that although SAP HANA is an In-Memory database, it can store the information on disk to ensure that the data is not volatile. This persistent mechanism is also responsible for ensuring the data integrity during an uncontrolled shutdown or a controlled shutdown of the SAP HANA database.

1.3.2 Multi-Database Container (MDC)

The initial iteration of SAP HANA emerged as version 1.0 SPS00 Revision 000.00. At the time of writing this book, the most recent iteration of SAP HANA stands as version 2.0 SPS07 Revision 0.71.00. The architecture of SAP HANA has undergone significant transformation from its inception to the present day.

SAP HANA version 1.0 SPS09 Revision 90 introduced the Multi-Database Container (MDC) concept that is integrated into the SAP HANA system. This feature facilitates the use of multiple databases, referred to as tenants, by various applications within a singular SAP HANA environment. This framework ensures the distinct separation of database users, database catalogs, repositories, backups, traces, and other elements for each individual tenant.

Before MDC was available, Multiple Components One Database (MCOD) and Multi-Component Operating System (MCOS) setups provided some flexibility and optimized hardware footprint. However, they are no longer recommended. For SAP HANA 2.0 SPS 01 and later, the MDC mode is mandatory.

Figure 1-3 shows the different options to deploy SAP HANA System MDC:

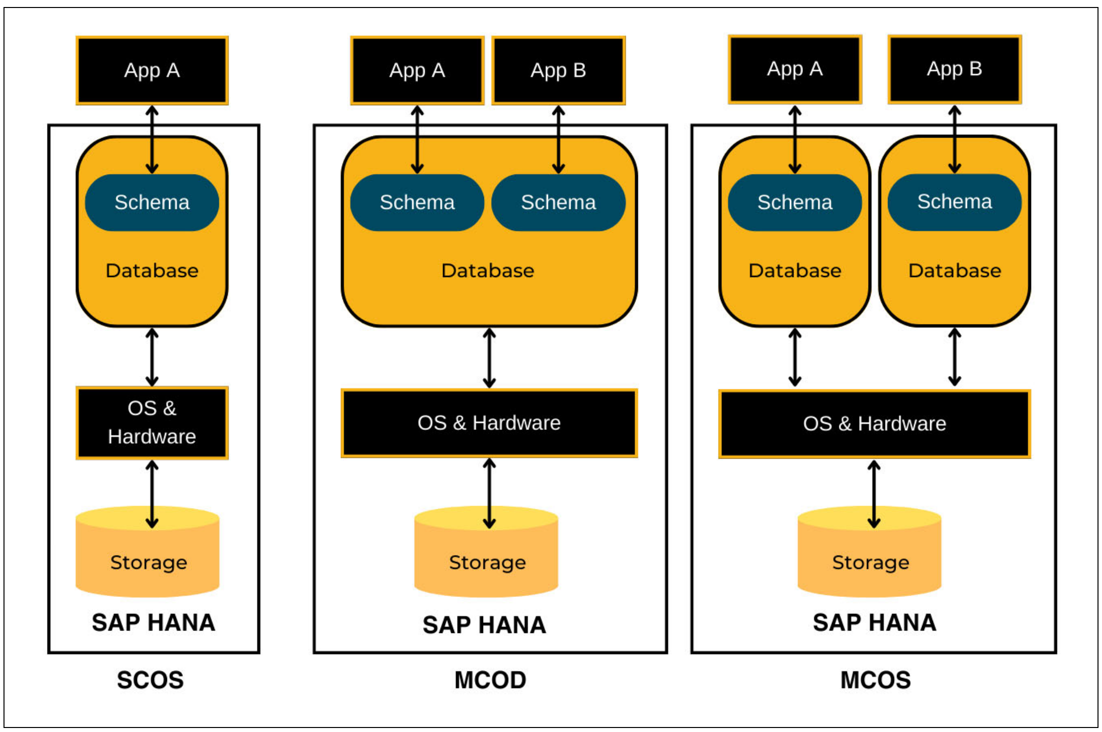


Figure 1-3 Options to deploy SAP HANA System without Multi-Database Container

Figure 1-4 shows the different deployment options for an SAP HANA System with MDC:

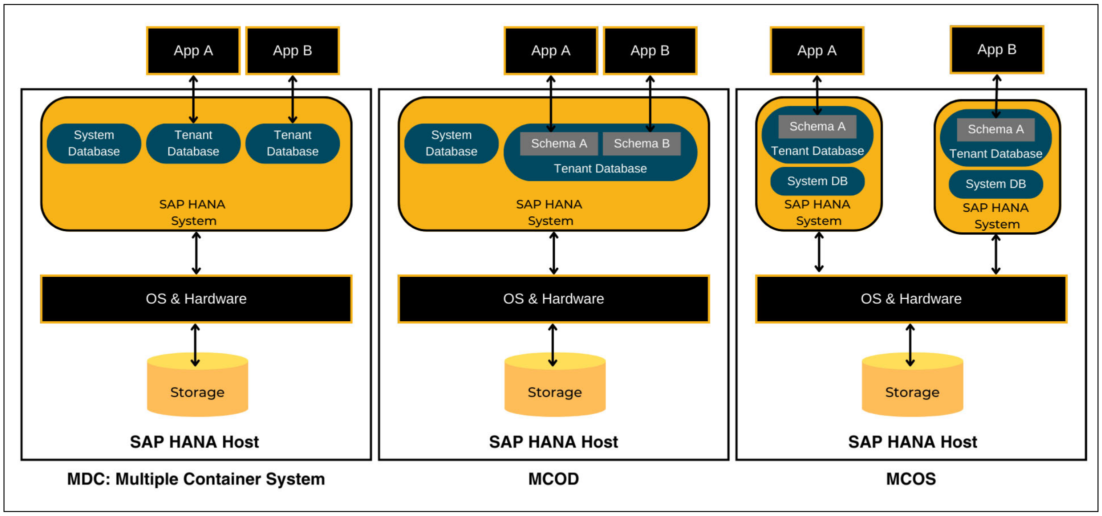


Figure 1-4 Options to deploy SAP HANA System with Multi Database Container (MDC)

It is important to mention that an SAP HANA system can be installed on a single host or across different hosts (multiple-host, distributed system, or scale-out system), providing greater scalability and availability to the installation.

When an SAP HANA system is installed, a system database is automatically created, and there is only one per system. Its main function is to maintain general information about the entire system and all existing tenant databases. Some important administration tasks for all tenant databases in the system are run from this system database. Examples of such tasks include performing backup and restore operations, creating and deleting new system databases, and configuring global parameters.

Similarly, a system database is created by default when installing an SAP HANA system, and this tenant database acquires the same ID as the SID of the installed database.

With the default implementation, it is possible to have up to 20 tenants in a single SAP HANA system. However, if enough resources are available, that limit can also be increased by increasing the number of ports to support more tenants in the configuration.

The implementation of SAP HANA on Power Systems provides the maximum benefits of multitenancy because several systems can be consolidated in a single server. Even with increasing computing resources demand, the performance is kept in check with features such as Power Systems Capacity on Demand (CoD), which scales up by activating suspended processors or memory within a server when needed.

The following list summarizes the benefits of using multitenancy with Power Systems:

- Integrated system administration and monitoring
- Optimized systems resources usage
- Flexible landscape management
- Lower total cost of ownership (TCO)

Because you can combine several SAP HANA instances in one DB system, the overall management of operating system (OS) setup, administration, monitoring, and backup is simplified. Instead of performing the same activity in multiple systems, it is done only once.

Another benefit is the reuse of resources in one server. The reuse of resources typically requires less computational resources because you have only a single OS for several DB instances that are shared among two or more tenants.

1.3.3 SAP HANA sizing considerations and advantages on Power Systems

For any IT system, capacity planning and the sizing of its components is a key activity in various phases of SAP projects. Because of large memory footprint requirements, sizing is more critical when in-memory DBs such as SAP HANA are part of the solution.

SAP defines the sizing process as the set of activities that is required to determine the correct hardware configuration capacity for a workload, which includes CPU, memory, storage size, I/O, and network bandwidth. The process is a combined effort of customers, SAP and hardware vendors, which include Power Systems.

An SAP System comprises software components that provide functionalities within an SAP solution. It is installed and configured as a unit with components possibly spread across multiple servers. An SAP System includes a database, one or more application server instances (ABAP or Java), SAP Central Services, and optional components. Each system is identified by a unique SAP system ID.

An SAP landscape is a logical group of systems based on the same SAP product and release but serving different purposes. SAP applications fall into two categories, their performance depending on the CPU and RAM resources of the hosting server:

- Application server component
- SAP HANA database

To effectively support a production system of SAP S/4HANA, you must also set up additional systems such as development, quality assurance, sandbox, training, and possibly more development or QA systems. Each of these systems must be considered during the sizing process, and the number and size of SAP application servers for each application must be considered.

Figure 1-5 shows an overview of a typical SAP system landscape of a client.

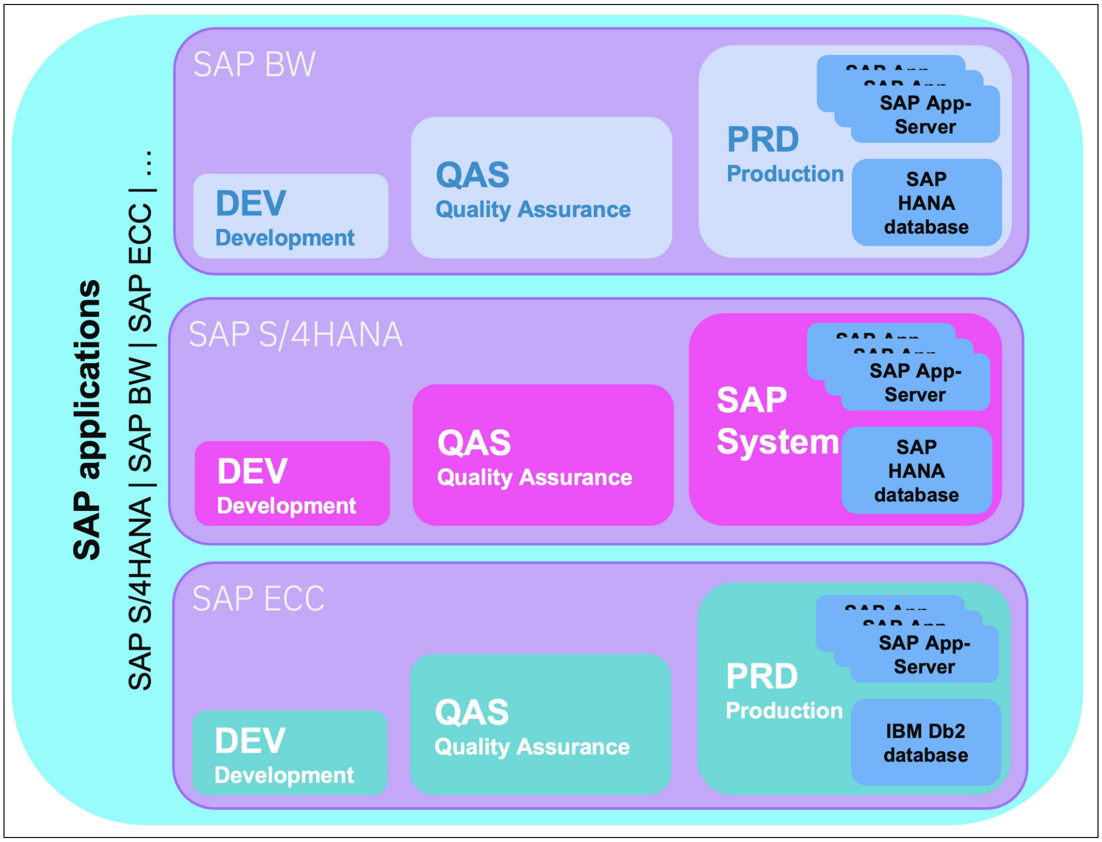


Figure 1-5 Overview of SAP HANA landscape

When gathering client data for all the landscapes, inquire about the following details:

- Description/Type. Specifies the application name or details, using SIDs as identifiers, to record the SAP application, such as SAP S/4HANA, SAP BW/4HANA, SAP ERP, SAP Solution Manager, etc.
- Memory. Represents the required server memory (RAM) for the SAP HANA database or application server. For SAP HANA, this information can be obtained from the sizing report or the existing HANA database on the Power/x86 system.
- SAP Application Performance Standard (SAPS). Refers to the CPU requirements for the database and SAP application servers. for more information, see SAP Application Performance Standard (SAPS).
- Workload Class. Indicates the type of workload.
- Workload Type. Identifies whether the SAP HANA system is a production or nonproduction system. This information is crucial for determining the percentage of shared processor pool entitlement that can be applied.
- CPU Allocation. Specifies whether dedicated cores or shared processor pool should be used.
- Server Placement. Might involve High Availability (HA) and Disaster Recovery (DR) requirements to safeguard the systems in case of failure.

The SAP HANA database can be used for Online Transactional Processing (OLTP) or Online Analytic Processing (OLAP) applications, with examples being S/4HANA, mainly for OLTP, and BW/4HANA, mainly for OLAP.

Application server components, running SAP application logic, exhibit medium CPU usage on average, with periods of high peak CPU usage. Memory (RAM) requirements tend to be low. Application servers are supported on either AIX or Linux operating systems.

For OLTP workloads, the SAP HANA database shows low average CPU usage but peaks during busy periods due to extensive internal database processing required to convert in-memory data into a format storable on persistent storage.

SAP HANA, an in-memory database, is exclusively supported on the Linux operating system. For OLAP workloads, the SAP HANA database demonstrates low average CPU usage but peaks with very high CPU demands. OLAP applications are primarily read-only, with data changes occurring when new data is loaded into the database at specific intervals. Similar to OLTP, OLAP workloads have high memory requirements, and SAP HANA is supported only on the Linux operating system.

Table 1-2 shows the SAP HANA and SAP Application servers and legacy databases workload class.

Table 1-2 SAP Workload class

	SAP Application Servers and Legacy Databases	SAP HANA Database OLAP (SAP BW)	SAP HANA Database OLTP (SAP S/4HANA)
CPU Peak	High	Very High	High
CPU Average	Medium	Low	Low
Memory	Low	Very High	Very High
Target OS	Windows, IBM i, AIX, Linux	Linux	Linux

To discuss sizing in an SAP context, the SAP Application Performance Standard (SAPS)[1] is used, measuring the performance of a system configuration in an SAP environment. SAPS is derived from the Sales and Distribution (SD) benchmark, defining 100 SAPS as 2,000 fully processed business order line items per hour.

Note: SAP takes responsibility for the sizing methodology and determining the required number of SAPS for a specific client's SAP application workload.

Matching the expected workload calculated in SAPS to the proposed hardware is the responsibility of SAP's hardware vendor partners who also perform hardware sizing. Sizing involves determining hardware requirements such as memory, CPU power, disk space, I/O capacity, and network bandwidth, and converting business requirements into hardware specifications. It is performed early in the project and influenced by both business and technological aspects. Specific optimization techniques of IBM Power Systems, unavailable on x86 Intel servers, are used in the sizing exercise. The sizing exercise can reduce excess CPU capacity requirements and improve usage of the purchased servers, which can result in lower Total Cost of Ownership (TCO).

1.3.4 Sizing process overview

The sizing process for newly implemented SAP applications, which is known as Greenfield Sizing, involves identifying main load drivers from business processes and estimating

[1] https://www.sap.com/about/benchmark/measuring.html

quantities. Clients with limited experience in SAP applications provide input through a questionnaire about users and business volume metrics.

SAP's Quick Sizer tool uses this data to estimate CPU requirements (in SAPS), memory, disk space, and I/O throughput. This is typically done in the early project phases when detailed business processes and data volumes might not be available.

For existing SAP applications or systems, known as Brown field Sizing, different approaches are used depending on the solution's lifecycle stage.

When migrating an existing SAP NetWeaver system to SAP HANA, there are two available sizing reports:

1. The SAP S/4 HANA sizing report determines the required SAP HANA RAM size, disk size, and CPU estimate for migrating the SAP NetWeaver-based ECC system to SAP HANA.
2. The BW sizing report is used for migrating an SAP Business Warehouse system to SAP HANA.

SAP provides tools to assist with sizing:

- Quick Sizer
- Sizing Guidelines

Figure 1-6 shows an overview flow of the sizing process.

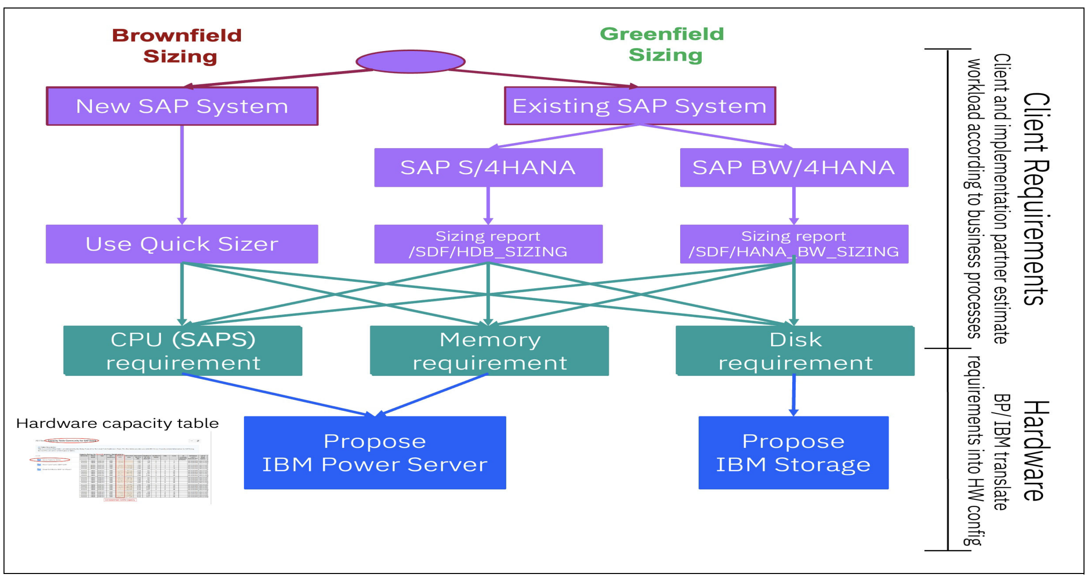


Figure 1-6 Overview of the SAP HANA sizing process

1.3.5 IBM Power CPU Virtualization

Virtualization is an advantage that is offered by IBM Power Systems and is provided by the built-in PowerVM hypervisor, which comes integrated with every IBM Power server. In contrast, x86 customers purchase the VMware vSphere virtualization software separately. IBM Power Systems employ Logical Partitions (LPARs) to partition each virtual machine (VM).

SAP HANA instances and application servers are hosted in dedicated LPARs, with CPU cores and RAM tailored to suit their specific workloads. A benefit of IBM Power Systems is the Shared Processor Pool, a feature not available on x86 servers. On IBM Power Systems each

LPAR can be sized with CPU cores based on average usage, which helps avoid wastage of idle CPU cores during peak workload times. The surplus capacity is pooled together and assigned to LPARs only when additional resources are needed. This results in a more efficient use of CPU cores across all LPARs, as not all LPARs require additional capacity simultaneously. However, memory remains static and must be sized to accommodate the SAP HANA data and working set requirements.

IBM PowerVM virtualization allows for the creation of more LPARs on the server, which leads to a greater reduction in the required number of CPU cores. This contributes to better resource utilization and performance optimization.

A Shared Processor Pool enables each LPAR to be sized based on average usage, which eliminates the wastage of CPU cores that would otherwise remain idle to accommodate peak workload times. The *Entitled CPU capacity* is assured for each LPAR, and if required, they can request additional CPU capacity from the shared pool. Unused cycles from an LPAR are contributed back to the pool, benefiting other LPARs. Because not all LPARs simultaneously require additional capacity, the allocated CPU cores in the common shared pool can be significantly fewer than the total CPU cores needed for all LPARs if such a feature were unavailable (Figure 1-7).

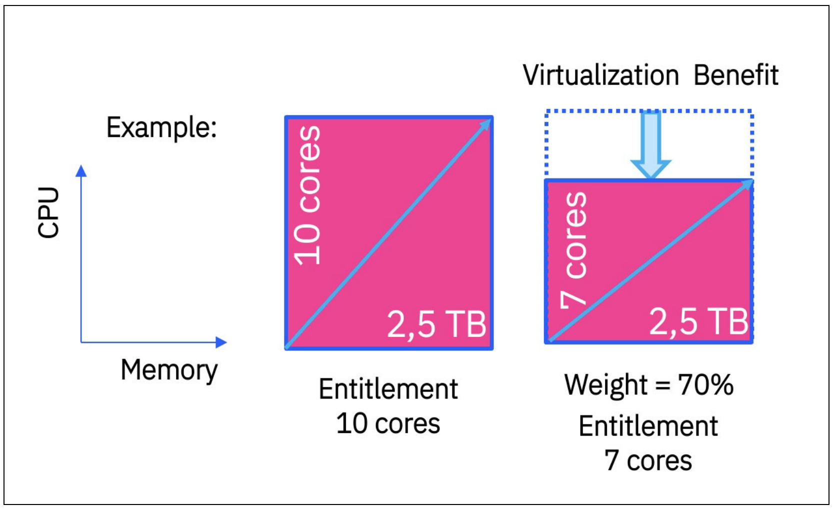


Figure 1-7 Virtualization Benefit

1.3.6 SAP HANA Native Storage Extension Advisor

Another feature of SAP HANA is memory optimization. Data temperature refers to the fact that not all data holds equal importance. Some data is considered *hot* and requires storage in the main memory for high performance. Other data is considered *warm* and is rarely accessed. Some data that is rarely accessed can be moved from the high-performance memory layer to the more cost-effective disk layers.

SAP HANA 2.0 SPS 04 introduced the Native Storage Extension (NSE) feature to handle warm data differently. NSE is a feature within the SAP HANA database, which means that the business applications built on top of SAP HANA need to support NSE. Figure 1-8 shows NSE considerations.

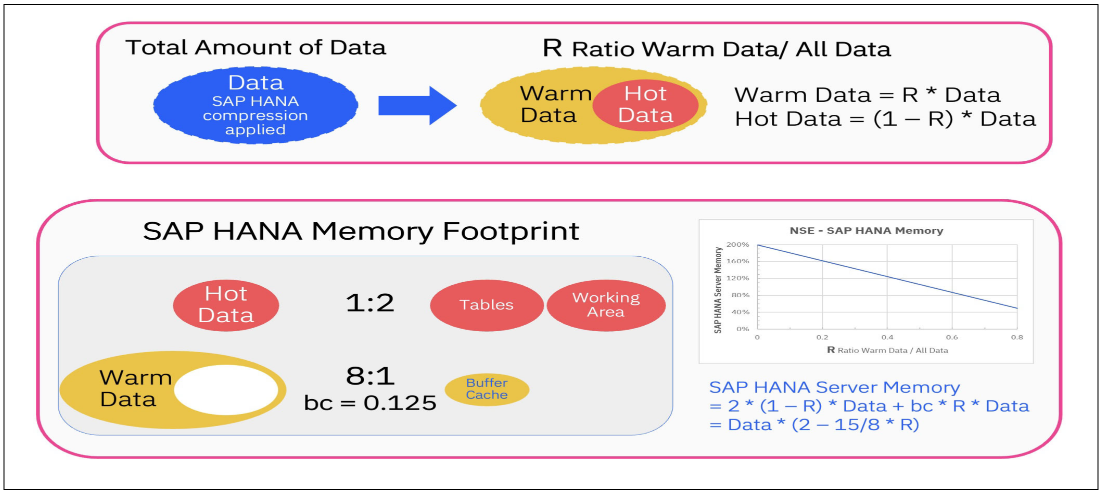


Figure 1-8 Native Storage Extension considerations

At the time of writing, NSE is applicable to SAP S/4HANA, SAP Business Suite systems and SAP BW/4HANA systems but not for SAP BW systems. Not all database objects related to the business application are eligible for NSE. The SAP business application defines additional boundary conditions.

With NSE, warm data is not loaded into memory. The majority remains on disks, and only a small portion is loaded into the buffer cache when needed. Therefore, instead of reserving twice the data footprint in memory for warm data, such as tables and working areas, only the buffer cache is allocated for warm data.

Assuming R as the relative amount of warm data, the amount of hot data is the full amount of data (100%) minus R. Typically, the buffer cache size is configured to be one-eighth of the size of the warm data. Therefore, the higher the relative amount of warm data R, the greater the memory savings.

SAP HANA provides the NSE Advisor, which is a tool to analyze and recommend whether data should be set to use NSE or not. The NSE Advisor is part of SAP HANA version 2.0 SPS04 and later. Clients can manage the NSE Advisor by using SAP HANA Cockpit 2.0 SP 11 or later or by using SQL commands.

To identify the objects, run the following exercise:

1. Configure the NSE Advisor and enable it in each SAP HANA system within the scope.
2. Run the representative workload, covering typical business processes and their access patterns.
3. Stop the NSE Advisor and analyze the results. It proposes candidates for warm data with a granularity of, for example, table, partition, or table column. Estimate the amount of warm data by using the output of the advisor run.
4. To gain memory benefits, alter the attributes of the identified objects from column loadable to page loadable. Continue running the representative workload to verify the settings.
5. Monitor the performance of statements, memory usage, and runtime of statements.
6. Repeat the exercise regularly to evaluate the proposal's effectiveness and identify any changes in workloads.

1.3.7 Using Enhanced Tetris Tool

This section describes the usage of the Tetris Excel spreadsheet.

IBM technical sellers and business partners can download the simplified sizing spreadsheet tool from the SAP HANA on Power Sales Kit page on Seismic:

Navigate to **SAP HANA Sizing**, and click **Enhanced Tetris Tool** to download it.

The streamlined sizing spreadsheet consists of multiple tabs as shown in Figure 1-9:

- The initial tab, LPAR Data, is where you input all the details regarding the SAP databases and application servers that are provided by the client. Also, you choose the proposed IBM POWER servers in this section.
- The “NSE optimization” tab is used to fine-tune the outcome by using the NSE information provided by the client.

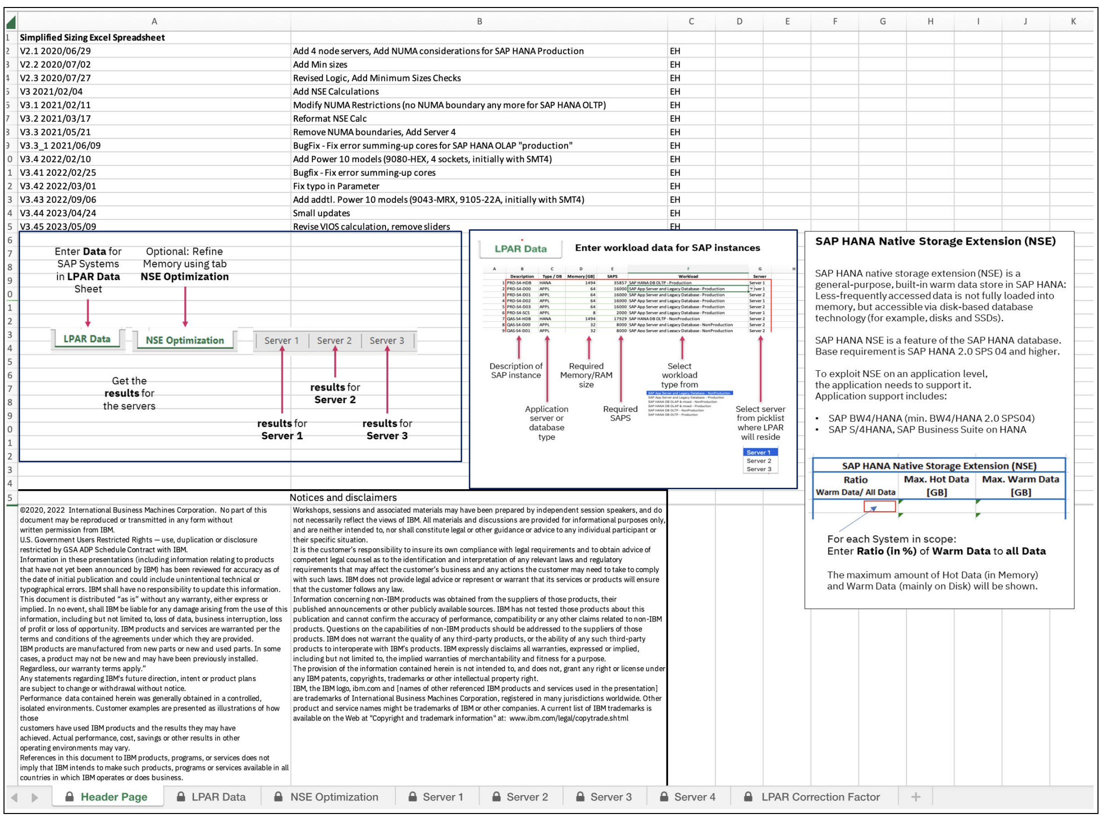
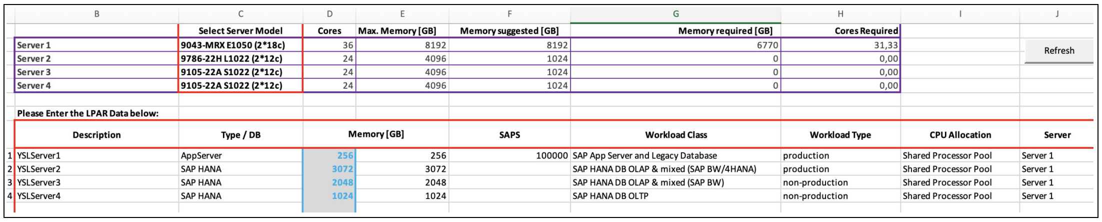


Figure 1-9 Tetris tool spreadsheet tabs

The following four tabs are used to display the results of the server CPU core and memory requirements for up to four IBM Power servers:

- The LPAR Data tab is the primary sheet where you input the relevant details of the SAP instances that you plan to deploy on the IBM Power servers. The following fields are mandatory and must be filled out as shown in Figure 1-10 :
  – Column B for a brief description of the SAP system or the SAP instance name.
  – Column C for the instance type, which can be APPL for the application server, HANA for the SAP HANA database instance, and ASE, Db2, ORA, or SQL for the legacy databases.

- – Column D/E for the required memory (RAM), which is available from the sizing reports.
- – Column F for the required SAPS, which is also available from the sizing reports.
- – Column G for the workload class, accessible from the picklist when clicking the down arrow on the right of the cell.
- – Column H for the workload type, also available from the picklist.
- – Column I for the CPU allocation, where you can choose between dedicated cores or a shared processor pool.
- – Column J for indicating the server in which this SAP instance is placed



	B	C	D	E	F	G	H	I	J
		Select Server Model	Cores	Max. Memory [GB]	Memory suggested [GB]	Memory required [GB]	Cores Required		
	Server 1	9043-MRX E1050 (2*18c)	36	8192	8192	6770	31,33		Refresh
	Server 2	9786-22H L1022 (2*12c)	24	4096	1024	0	0,00		
	Server 3	9105-22A S1022 (2*12c)	24	4096	1024	0	0,00		
	Server 4	9105-22A S1022 (2*12c)	24	4096	1024	0	0,00		
	Please Enter the LPAR Data below:								
	Description	Type / DB	Memory [GB]		SAPS	Workload Class	Workload Type	CPU Allocation	Server
1	YSLServer1	AppServer	256	256	100000	SAP App Server and Legacy Database	production	Shared Processor Pool	Server 1
2	YSLServer2	SAP HANA	3072	3072		SAP HANA DB OLAP & mixed (SAP BW/4HANA)	production	Shared Processor Pool	Server 1
3	YSLServer3	SAP HANA	2048	2048		SAP HANA DB OLAP & mixed (SAP BW)	non-production	Shared Processor Pool	Server 1
4	YSLServer4	SAP HANA	1024	1024		SAP HANA DB OLTP	non-production	Shared Processor Pool	Server 1

Figure 1-10 LPAR Tab

- ► In the NSE Optimization Tab, the objective is to redefine the actual quantities of hot data and warm data for SAP HANA, so you can adjust the required memory (Figure 1-11):
  - – The lines containing SAP applications where NSE is not applicable are shaded in gray.
  - – On the remaining lines:
    - • Column I labeled Total amount of data [GB] indicates the maximum data size to be stored in the SAP HANA System. Without NSE, this value is a maximum of half the size of the main memory, following SAP's guidelines for the working memory in addition to the data.
    - • If no adjustments are made, the Hot Data [GB] entry in Column G remains the same.
  - – Collaborate with the client to jointly run the NSE Advisor and determine the amount of warm data. If NSE is applicable, and you identified the amount of warm data, then use the spreadsheet as described in the following list:
    - • Enter the total data amount in Column D (using the value from Column I).
    - • Enter the total amount of warm data, relative to the total data, in Column E.
    - • Column G is then adjusted to show the new amount of hot data.
    - • Column H displays the entered amount of warm data.
    - • The required SAP HANA main memory is recalculated (Column J) and appears back in the LPAR Data sheet.

Note: There is no SAP NSE support for SAP BW and AppServers.

			SAP HANA Native Storage Extension (NSE)						
			Adjust			Summary			
			Total amount of data [GB]	Total amount of warm data [GB]	Ratio Warm Data/ all Data	Hot Data [GB]	Warm Data [GB]	Total amount of data [GB]	SAP HANA Memory [GB]
YSLServer1	AppServer	SAP App Server and Legacy Database							
YSLServer2	SAP HANA	SAP HANA DB OLAP & mixed (SAP BW/4HANA)	1536	230	15%	1306	230	1536	2641
YSLServer3	SAP HANA	SAP HANA DB OLAP & mixed (SAP BW)							
YSLServer4	SAP HANA	SAP HANA DB OLTP				512		512	1024

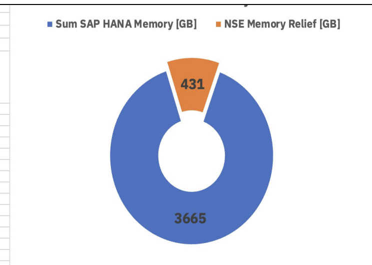


Figure 1-11 LPAR Tab after NSE optimization

- Select the LPAR Data sheet. After the NSE optimization, there is a revised value in Column D that represents the memory (RAM) size, which is used for the calculations. Column E retains the previous figure, but it is overridden by the NSE calculation, so it is displayed in gray because it is no longer applicable as shown in Figure 1-12.

	Select Server Model	Cores	Max. Memory [GB]	Memory suggested [GB]	Memory required [GB]	Cores Required		
Server 1	9043-MRX E1050 (2*18c)	36	8192	8192	6174	29,18		Refresh
Server 2	9786-22H L1022 (2*12c)	24	4096	1024	0	0,00		
Server 3	9105-22A S1022 (2*12c)	24	4096	1024	0	0,00		
Server 4	9105-22A S1022 (2*12c)	24	4096	1024	0	0,00		

Please Enter the LPAR Data below:

	Description	Type / DB	Memory [GB]		SAPS	Workload Class	Workload Type	CPU Allocation	Server
1	YSLServer1	AppServer	256	256	100000	SAP App Server and Legacy Database	production	Shared Processor Pool	Server 1
2	YSLServer2	SAP HANA	2641	3072		SAP HANA DB OLAP & mixed (SAP BW/4HANA)	production	Shared Processor Pool	Server 1
3	YSLServer3	SAP HANA	2048	2048		SAP HANA DB OLAP & mixed (SAP BW)	non-production	Shared Processor Pool	Server 1
4	YSLServer4	SAP HANA	887	1024		SAP HANA DB OLTP	non-production	Shared Processor Pool	Server 1

Figure 1-12 LPAR Tab after NSE optimization

- On the Server sheet you find aggregated information per physical server. The CPU and memory resources for SAP instances are combined based on their placement on the server as shown Figure 1-13.
  - The yellow section represents the total combined resources for production and nonproduction application servers and databases like Db2 and Oracle.
  - The green section displays the total combined resources for production and nonproduction SAP HANA databases used for OLAP and mixed OLAP applications such as SAP BW/4HANA and SAP BW, SAP SEM.
  - The blue section showcases the total combined resources for production and nonproduction SAP HANA databases used for OLTP applications such as SAP S/4HANA and SAP ECC.
  - The tan-colored cell indicates the number of LPARs that are configured on the IBM Power server.

SAP App Server and Legacy Database	Production	11,57	256
	Non-Production	0	**0**
SAP HANA DB OLAP & mixed	Production	18,87	**2640,75**
	Non-Production	7,32	**2048**
SAP HANA DB OLTP	Production	0	**0**
	Non-Production	2	**887,375**
Total Number of LPARs:	4		

Figure 1-13 Server1 Tab

– The initial table displays the resource requirements without optimization. In this example, the baseline requirement without optimization is approximately the same as the total memory required and remains constant at about 5.8 TB throughout the optimization phases as shown in Figure 1-14.

Server 1 - raw: 39,8 cores required		
Cores	Memory	VPs
11,6	256	12
0,0	0	0
18,9	2640,75	19
7,3	2048	8
0,0	0	0
2,0	887,375	2

Figure 1-14 LPAR Tab after NSE optimizationFigure –

– The second table at the bottom demonstrates the reduction in CPU cores that is achieved by implementing shared processor pools. As a result, the total number of CPU cores decreases from 40 cores to approximately 29 cores, which includes the VIOS (Figure 1-15.).

Server 1 - Shared Pool: 29,2 cores required (incl. VIOS)	
Entitlement	Memory
8,1	256
0,0	0
13,2	2640,75
3,4	2048
0,0	0
0,9	887,375

Figure 1-15 LPAR Tab after NSE optimization 40 cores

1.3.8 SAP HANA on IBM Power Systems additional considerations

An SAP HANA on Power Systems implementation is based on the Tailored Datacenter Integration (TDI) model, which allows customers to reduce costs by efficiently provisioning SAP HANA partitions that are ready for use with IBM PowerVM. In addition to the SAP HANA workload requirements, which can be resized based on business needs, on a Power Systems server you can run a maximum of eight SAP HANA DBs and non-SAP HANA workloads.

Although several dimensions of computational capacity are provided, the main one for the sizing of SAP HANA on Power Systems is the SAP HANA memory size requirements.

For more information about sizing, see SAP HANA on IBM Power Systems and IBM System Storage - Guides

Because you can combine several SAP HANA instances in one DB system, the overall management of operating system (OS) setup, administration, monitoring, and backup is simplified. Instead of performing the same activity in multiple systems, it is done only once. Another benefit is the reuse of resources in one server, a situation that typically requires less computational resources because you have only a single OS for several DB instances that are shared among two or more tenants.

SAP recommends a more conservative approach for sizing MDC, as described in SAP Note 2096000, where you use the additive sizing approach. When you use the additive sizing approach, each individual instance is added to an overall server, for example 4 TB for S/4HANA + 4 TB for BW/4HANA + 2 TB for a native SAP HANA application. However, the reality is that memory, vCPU, and storage resources can be saved by using the combination approach because the impact of the OS and DB can be shared. In addition, tailored sizing allows a much more demand-driven combination.

1.3.9 Moving or copying an SAP HANA tenant

By using the SAP HANA MDC technology, you can copy and move tenants across different SAP HANA systems, which adds flexibility to system management because you can adjust the SAP HANA tenant placement according to the current demand.

1.4 Scale-up and scale-out on Power Systems servers

The ability of computer systems to adapt over time to changing business demands (which usually involves growth) is referred to as *scalability*. Scalability is a typical challenge in the IT business because of the degree of uncertainty to predict growth rates of computational power requirements, which include processing, memory, storage, and bandwidth. The infrastructure scalability to handle the application demands is addressed through two main vectors, *scale-up,* or vertical scaling, and *scale-out* or horizontal scaling.

Scale-up is achieved by adding capacity to an existing infrastructure, like a single system node such as a server, where processors, memory, storage, and network interfaces can be added or replaced with ones with more capacity.

With scale-out, you combine multiple independent servers under a single system to distribute your SAP HANA data across multiple nodes or hosts. The main purpose of this approach is to overcome the limitations of a single server. You can also use the scale-out approach to gain performance or to meet a specific application requirement. Several considerations must be accounted for when you use this approach:

- Determine whether the application supports the scale-out approach.
- Based on your specific scenario and application, decide whether the data must be partitioned by schemas, tables, or table segmentation. The data is then divided among different server nodes.
- Assure that application-specific demands can be met, such as native applications with concurrent access that can benefit from distributed indexservers in different machines. If performance stress test data indicates that there are enough CPU cores on a single machine to process simultaneous threads, the overall performance can be better than multiple indexservers.

IBM Power Systems servers support both scale-up and scale-out modes for SAP HANA systems to address the demands of different workloads as shown in Figure 1-16 .

> **Note:** You can order a configuration with a high number of processor cores per socket. However, SAP certified IBM Power10 servers for SAP HANA production systems that contain a maximum of 20 cores per socket.

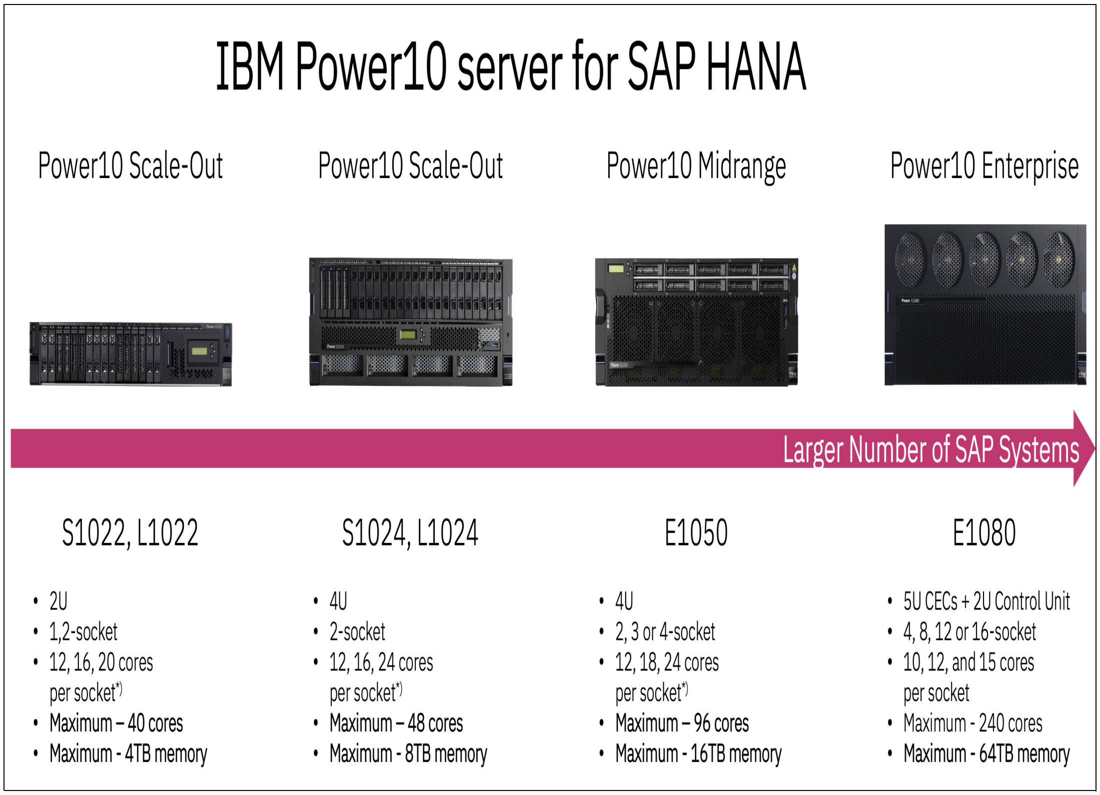


Figure 1-16 Supported Scale-up and Scale-out nodes

IBM Power offers the flexibility to allocate additional processor and memory capacity on-the-fly without the need for system shutdowns or LPAR reboots.

With the introduction of IBM Power10 for SAP HANA, IBM solidified its position as the leader among vendors supporting large SAP workloads. The latest enhancements to IBM Power solutions are to optimize the Power server experience for clients and streamlining and expediting their journey to SAP HANA.

Facilitating SAP HANA migration is now even smoother with the new 2–6TB bundles. As IBM Power10 provides up to 2.5x better per-core performance than comparable x86 servers, clients have the opportunity to achieve more, with fewer resources.

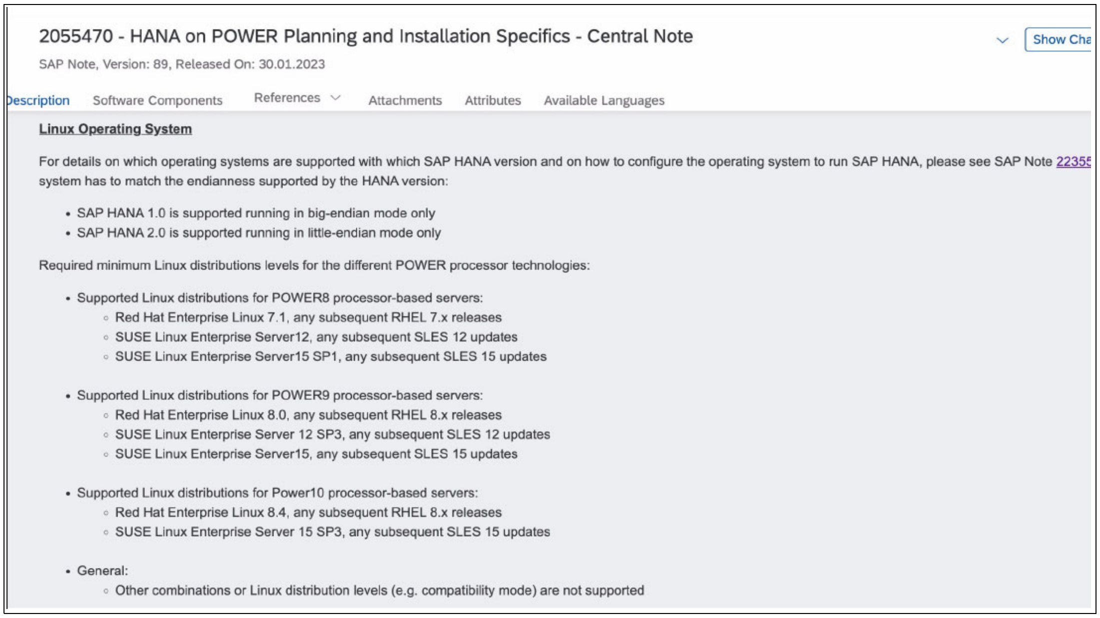
Because IBM recognizes the need for enterprise-class attributes in smaller SAP HANA workloads, IBM has introduced these new bundles, available through the Business Partners. The bundles include IBM Power L1022 and IBM Power L1024 servers ideally suited for smaller SAP HANA deployments, which offer competitive pricing and value.

In addition to the smaller workloads, IBM innovation extends to scaling SAP HANA OLAP workloads with unparalleled flexibility, which allows clients to allocate resources incrementally, starting as low as 0.01 cores and 1 GB memory. IBM announced SAP's support for a maximum of 40 TB (OLAP only) on IBM Power. At the time of writing, IBM is the first and only server vendor to receive this level of support for such large SAP HANA workloads.

1.5 Use cases for migration to SAP HANA on IBM Power10

The following section provides a practical description of different scenarios for possible migrations of SAP HANA on IBM Power. By doing so and based on experience, the scenarios can offer customers ideas that can aid in defining the migration strategy and ultimately, complete a successful project.

The first thing to consider for any SAP HANA migration project to IBM Power10 System is the minimum supported operating systems by IBM Power10 System. Refer to SAP Note: 2055470 - HANA on POWER Planning and Installation Specifics - Central Note.


2055470 - HANA on POWER Planning and Installation Specifics - Central Note

SAP Note, Version: 89, Released On: 30.01.2023

Show Cha

Description Software Components References Attachments Attributes Available Languages

Linux Operating System

For details on which operating systems are supported with which SAP HANA version and on how to configure the operating system to run SAP HANA, please see SAP Note 2235
system has to match the endianness supported by the HANA version:

- SAP HANA 1.0 is supported running in big-endian mode only
- SAP HANA 2.0 is supported running in little-endian mode only

Required minimum Linux distributions levels for the different POWER processor technologies:

- Supported Linux distributions for POWER8 processor-based servers:
  - Red Hat Enterprise Linux 7.1, any subsequent RHEL 7.x releases
  - SUSE Linux Enterprise Server12, any subsequent SLES 12 updates
  - SUSE Linux Enterprise Server15 SP1, any subsequent SLES 15 updates
- Supported Linux distributions for POWER9 processor-based servers:
  - Red Hat Enterprise Linux 8.0, any subsequent RHEL 8.x releases
  - SUSE Linux Enterprise Server 12 SP3, any subsequent SLES 12 updates
  - SUSE Linux Enterprise Server15, any subsequent SLES 15 updates
- Supported Linux distributions for Power10 processor-based servers:
  - Red Hat Enterprise Linux 8.4, any subsequent RHEL 8.x releases
  - SUSE Linux Enterprise Server 15 SP3, any subsequent SLES 15 updates
- General:
  - Other combinations or Linux distribution levels (e.g. compatibility mode) are not supported

Figure 1-17 SAPNote 2055470 - HANA on POWER Planning and Installation Specifics - Central

Note: As seen in Figure 1-17, the minimum required versions are SUSE Linux Enterprise Server 15 SP3 or later and Red Hat Enterprise 8.4 or later.

The following examples provide a customer situation, a challenge, and a general step-by-step approach for the different possible options to successfully migrate to an IBM Power10 System.

1.5.1 Case 1 – SAP ERP Business Suite 6.08 on SAP HANA

A customer has an SAP ERP Business Suite on SAP HANA, and its SAP system landscape is distributed:

- SAP HANA 2.0 revision 55 on IBM Power8® system with operating system SUSE Linux Enterprise Server 12 SP3
- SAP application servers (ASCS and PAS) on IBM Power8 system with operating system SUSE Linux Enterprise Server 12 SP3

Challenge: Migrate the SAP system to an IBM Power10 system with SUSE Linux Enterprise Server 15 SP4 and SAP HANA Revision 71.

Option 1: Updates to operating systems and SAP HANA

In the database node on Power8 system perform the following steps:

- Update SUSE Linux Enterprise Server 12 SP3 to SUSE Linux Enterprise Server 12 SP5 (1 online update).
- Upgrade SUSE Linux Enterprise Server 12 SP5 to SUSE Linux Enterprise Server 15 SP3 (1 offline upgrade). HANA 2.0 revision 55 is not supported on SUSE Linux Enterprise Server 15 SP4.
- Update SAP HANA to version 2.0 revision 71. This version is supported on SUSE 15 SP4.
- Update SUSE Linux Enterprise Server 15 SP3 to SUSE Linux Enterprise Server 15 SP4 (1 online update).
- Start SAP HANA and validate its operation.
- Run **saptune** (Solution HANA database) (SUSE Only).
- Execute Minicheck (General configuration).
- Stop SAP HANA database.
- Move the LPARs with IBM Live Partition Mobility (LPM) from Power8 system to Power10 system.

In the application server node, perform the following steps:

- Update SUSE Linux Enterprise Server 12 SP3 to SUSE Linux Enterprise Server 12 SP5 (1 online update).
- Upgrade SUSE Linux Enterprise Server 12 SP5 to SUSE Linux Enterprise Server 15 SP4 (1 offline upgrade).
- Update SAP HANA Client and SAP Host Agent.
- Run **saptune** (Solution Netweaver) (SUSE Only).
- Start SAP Application Server for testing.
- Move the LPARs with IBM LPM from Power8 system to Power10 system.

After the steps are performed, the LPARs are running on the IBM Power10 system.

Option 1 includes the following advantages and disadvantages:

- Advantages:
  - It is possible to run some tasks on the IBM Power8 system before moving to the new infrastructure running on the IBM Power10 system.
  - The update times are relatively short.
- Disadvantages:
  - The execution time to move the LPAR can vary depending on the storage size for each LPAR.
  - Because they are not clean installations, after the LPARs are migrated to IBM Power10 system, they retain all the residual information from the previous versions.

Option 2: New installation of SAP HANA and update operating system on App Server

In the database node on IBM Power10 system, perform the following steps:

- Install a new LPAR on IBM Power10 system with SUSE Linux Enterprise Server 15 SP4 for SAP.

- Install SAP HANA database 2.0 Revision 71.
- Run **saptune** (Solution HANA database) (SUSE Only).
- Perform a backup on the database node on IBM Power8 system.
- Perform a restore on the new SAP HANA 2.0 Revision 71 database node.

At the application layer on IBM Power8 system, perform the following steps:

- Update SUSE Linux Enterprise Server 12 SP3 to SUSE Linux Enterprise Server 12 SP5 (1 online update).
- Upgrade SUSE Linux Enterprise Server 12 SP5 to SUSE Linux Enterprise Server 15 SP4 (1 offline upgrade).
- Update SAP HANA Client and SAP Host Agent.
- Run **saptune** (Solution Netweaver) (SUSE Only).
- Start SAP Application Server for testing.
- Move the LPARs with IBM LPM from Power8 system to Power10 system.

Option 2 includes the following advantages and disadvantages:

- Advantages:
  – Clean installations on the new operating system, no junk, no bad practices
  – Short time for application layer updates
- Disadvantages:
  – Greater effort in testing SAP systems due to clean installations.
  – Availability of the new infrastructure is required from the beginning.
  – The time to move the LPAR varies and depends on the storage size of each LPAR.

Option 3: New installation to SAP HANA and SAP System Copy

At the database level, perform the following steps:

– Install a new LPAR on IBM Power10 system with SUSE Linux Enterprise Server 15 SP4 for SAP.
– Install SAP HANA database 2.0 Revision 71.
– Run **saptune** (Solution HANA database) (SUSE Only).
– Perform a backup on the database node on IBM Power8 system.
– Perform a Restore on the new SAP HANA 2.0 Revision 71 database node. This restore is automatically run by the SAP SWPM tool.

At the application layer, perform the following steps:

– Install a new LPAR on IBM Power10 system with SUSE Linux Enterprise Server 15 SP4.
– Perform the installation of the SAP system by using the System Copy option with the backup restore method. This results in a clean application layer installation.
– Run saptune (Solution Netweaver) (SUSE Only).

Option 3 includes the following advantages and disadvantages:

- Advantages:
  – Clean installations on the new operating system, no junk, no bad practices
  – Possible redesign of the architecture if necessary
- Disadvantages:
  – Greater effort in testing SAP systems due to clean installations
  – Requires availability of the new infrastructure from the beginning

Option 4: New installation of SAP HANA, SAP System Copy and SAP HANA System Replication

The following option is recommended for production environments with the goal of reducing the downtime window before production can start.

At the database level, perform the following steps:

- Install a new LPAR on IBM Power10 system with SUSE Linux Enterprise Server 15 SP4 for SAP.
- Install SAP HANA database 2.0 Revision 71.
- Run `saptune` (Solution HANA database) (SUSE Only).
- Perform a backup on the database node on IBM Power8 system.
- Run a Restore on the new SAP HANA 2.0 Revision 71 database node. This restore is automatically run by the SAP SWPM tool.

At the application layer, perform the following steps:

- Install a new LPAR on IBM Power10 system with SUSE Linux Enterprise Server 15 SP4 for SAP.
- Perform the installation of the SAP system using the System Copy option with the backup restore method. This results in a clean application layer installation.
- Run saptune (Solution Netweaver) (SUSE Only).
- Start SAP systems and test system functionality.
- Validate integrations with other systems such as File Server and Integrations.
- Stop SAP Application Server.

Additional steps are required:

- Configure SAP HANA System Replication between Node 1 (IBM Power8 system) and Node 2 (IBM Power10 system).
- Validate the synchronization of the replication.
- Wait for the scheduled maintenance window.
- On the day of the maintenance window, stop SAP systems on IBM Power8 system.
- Validate the synchronization of the SAP HANA database replication again.

SAP HANA on node 1 must be stopped and powered off before you use the new system:

- Stop and power off SAP HANA on Node 1 (IBM Power8 system).
- Execute a TakeOver of SAP HANA.
- Start new SAP systems on the new IBM Power10 system servers.

Option 4 includes the following advantages and disadvantages:

- Advantages:
  - Clean installations on the new operating system, no garbage, no bad practices.
  - Possible redesign of the architecture if necessary.
  - Reduction of downtime for go-live, thanks to HANA SR (System Replication).
  - It can be started with weeks of anticipation before go-live.
  - If there are any issues, it is possible to revert to the database on Node 1 (IBM Power8 system).

- Disadvantages:
  - Availability of the new infrastructure is required from the beginning.
  - Implementation requires more effort, but provides increased security.

1.5.2 Case 2 – SAP NetWeaver 7.5 on SAP HANA

A customer has an SAP NetWeaver 7.5 on SAP HANA, and its SAP system landscape is distributed:

- SAP HANA 2.0 revision 55 on IBM Power8 system with operating system SUSE Linux Enterprise Server 15 SP3
- SAP application servers (ASCS and PAS) on IBM Power8 system with operating system SUSE Linux Enterprise Server 15 SP3

Challenge: Migrate the SAP system to IBM Power10 system with SUSE Linux Enterprise Server 15 SP4 and SAP HANA Revision 71.

Option 1: Updates to operating systems and SAP HANA

In the database node on Power8 system, perform the following steps:

- Update SAP HANA to version 2.0 Revision 71. This version is supported in SUSE 15SP4.
- Update SUSE Linux Enterprise Server 15SP3 to SUSE Linux Enterprise Server 15SP4 (1 Online Update).
- Start SAP HANA and validate its functionality.
- Run saptune (Solution Hana database) (SUSE Only).
- Run Minicheck (General configuration).
- Stop SAP HANA database.
- Migrate the LPARs with IBM Live Partition Mobility (LPM) from Power8 system to Power10 system.

In the App Server node, perform the following steps:

- Update SUSE Linux Enterprise Server 15SP3 to SUSE Linux Enterprise Server 15SP4 (1 Online Update).
- Update SAP Hana Client and SAP Host Agent.
- Run `saptune` (Solution Netweaver) (SUSE Only).
- Start SAP Application Server for testing.
- Migrate the LPARs with IBM from Power8 system to Power10 system.

At the end of this process, the LPARs are running on IBM Power10 system.

Option 1 includes the following advantages and disadvantages:

- Advantages:
  - Some tasks can be run on the IBM Power10 system platform while waiting for the new IBM Power10 system infrastructure to be delivered.
  - Operating system update times can be relatively short.

Disadvantages:

– The run time to move the LPAR can vary depending on the storage size for each LPAR.
– Because these are not clean installations, After the LPARs are migrated to IBM Power10 system, they contain all the garbage information from previous versions.

Option 2: New installation SAP HANA and SAP System Copy

In the database node, perform the following steps:

- Install a new LPAR on IBM Power10 system with SUSE Linux Enterprise Server 15SP4.
- Install SAP HANA database 2.0 Revision 71.
- Run `saptune` (Solution Hana database) (SUSE Only).
- Perform a backup on the database node on IBM Power8 system.
- Perform a Restore on the new SAP HANA 2.0 Revision 71 database node. This restore is automatically run by the SAP SWPM tool.

In the application layer, perform the following steps:

- Install a new LPAR on IBM Power10 system with SUSE Linux Enterprise Server 15SP4.
- Perform an SAP system installation by using the System Copy option with the backup restore method. This provides a clean application layer installation.
- Run `saptune` (Solution Netweaver) (SUSE Only).

Option 2 includes the following advantages and disadvantages:

- Advantages:
  - Clean installations on the new operating system, avoiding clutter and bad practices
  - Possible architecture redesign if necessary
- Disadvantages:
  - Requires more effort in testing the SAP systems due to clean installations.
  - New infrastructure availability is required from the beginning.

1.6 Backup and recovery for SAP HANA systems

A true business continuity plan is always accompanied by a set of elements. The elements include hardware, software, services, and strategies that, together, guarantee organizations the ability to maintain their operations and functions that are truly critical during disruptions or disasters. Depending on the size and complexity of the business, these continuity plans can be implemented with small, medium, or very large solutions. Nevertheless, backups are part of any of the implemented solutions. Hence, the importance of ensuring backup and recovery solutions work effectively and reliably when they are needed.

Included in a discussion of backups are the components of an SAP HANA database. SAP HANA is an in-memory database, but it still relies on storage for data persistency, necessitating substantial disk space.

The required disk space includes enough space for the following data:

- Data in before-image and after-image versions
- Space for at least one process image in case of software failure
- Space for one data export

During regular database operation, modified data is automatically saved from memory to disk at regular savepoints. These savepoints are typically created every five minutes, even during a backup process. With a well-configured hardware setup, the performance impact of savepoints is negligible, and they do not disrupt transaction processing.

Transactions continue to run as usual during a savepoint, and new transactions can be initiated without any issues. Also, data changes are recorded in the redo log buffer's log segments. When a database transaction is committed, these log segments are saved to disk. Additionally, if the redo log buffer reaches its capacity, it is written to disk, regardless of whether a commit has been issued.

Apart from user data, SAP HANA requires disk space to store transaction logs, specifically the Redo logs, which are crucial for recovery in case of system failure or point-in-time recovery. SAP HANA logs are managed cyclically, and they reuse the same space until a backup of the log segments is completed, which clears the log space in memory.

There are many tools available for backup and recovery. Some are native to the operating system on which the SAP HANA system is installed, and others are third-party tools that provide advanced features such as compression and parallelism.

This section includes a discussion of the native tools provided by operating systems like SUSE Linux Enterprise Server, Red Hat Enterprise Linux, and some third-party tools that can be seamlessly integrated with our IBM Power System platform.

One of the advantages of having the technology of IBM Power System is the ability to use tools that cover from infrastructure to the application layer and which assist in performing tasks such as updates or migrations of the SAP HANA System. This enables the establishment of a quick rollback point before any system intervention. With Image Capture[2], for example, you can take copies or images of the main operating system disk or the entire LPAR. This tool performs a storage IBM FlashCopy® of the LPAR either locally or to the IBM Cloud Object Store.

As mentioned before, a backup strategy is a fundamental part of designing a good business continuity plan. Therefore, it is important to be aware of the different options that SAP HANA provides as backup types:

- Full Backup. A full backup is a complete copy of all the data in the SAP HANA database. It captures all the database objects, including tables, views, stored procedures, and other components. A full backup is essential to fully restore the database in the event of a complete data loss. The two types of Full Backup are Complete data backups and Data snapshots.
- Delta backup. Here, only the data that has changed since the last Full Data Backup is backed up. There are two types of delta backups:
    a. Incremental Backup. An incremental backup captures only the changes made since the last backup, whether it is a full backup or a previous incremental backup. This type of backup is faster and requires less storage space than a full backup because only the modified data is backed up. To restore the database by using incremental backups, the incremental backups must be applied sequentially until the recovery point is reached.
    b. Differential Backup. A differential backup captures the changes made since the last full backup. Unlike an incremental backup, a differential backup backs up all the changes made since the last full backup, rather than just the changes since the last incremental backup. This means that differential backups can be faster, larger, and require more storage space than incremental backups, but the restoration process can be faster because only the most recent full backup and the latest differential backup need to be restored.

[2] Backup strategies for IBM Power Systems Virtual Servers:
https://www.ibm.com/docs/en/power-systems-vs?topic=strategies-backup-power-systems-virtual-servers

- Redo Log backup. A type of backup that specifically captures the transaction log information in the system. The transaction log, also known as the Redo Log, records all the changes made to the database, which includes inserts, updates, and deletes. During normal database operations, SAP HANA continuously writes the transaction log records to the Redo Log files. These Redo Log files serve as a crucial component for data recovery and consistency in case of system failures or unexpected disruptions. Redo Log backups capture the contents of the Redo Log files at a specific point in time. By taking regular Redo Log backups, organizations can ensure the ability to restore the database to a consistent state up to the most recent Redo Log backup. These backups can be used during the recovery process to apply the transactions recorded in the Redo Log files and restore the database.

Note: By default, SAP HANA databases automatically run Redo Log Backups. Therefore, it is crucial to ensure that you have sufficient space for the complete backup cycle on disk or in the external tool you use for backups. This way, you can guarantee that no information is lost if you need to restore the database to a specific point in time.

Table 1-3 shows the differences between the two SAP HANA Delta Backup types.

Table 1-3 Differences between SAP HANA Delta Backup types

	Differential Backup	**Incremental Backup**
Type of data backed up	The data changed since the last full data backup	The data changed since the last full data backup or the last delta backup
Backup size	The amount of data to be saved with each differential backup increases.	If data remains unchanged, it is not saved to more than one backup. For this reason, incremental backups are the smallest of the backup types
Backup and recovery strategy	If the backup strategy is based on only full data backups and differential backups, only two backups are needed for a recovery: one full data backup and one differential backup	If the backup strategy is based on only full data backups and incremental backups, to recover the database, SAP HANA needs the following backups: ▸ The full data backup on which the incremental backups are based ▸ Each incremental backup made since the full data backup In some situations, many incremental backups might be needed for a recovery.

The SAP HANA database provides the necessary tools to ensure that backups are run. However, you can use third-party tools that add features for running and managing backup and recovery tasks that include some of the following main features or advantages:

- Data compression
- Encryption
- De-duplication
- Centralized backup management
- Flexible backup scheduling

Another important point to consider when designing the backup strategy is to understand that in SAP HANA, each system database is autonomous for backup and recovery tasks. This includes the system database, too. For each backup task and backup type, determine the number of tenants that need to be backed up, the size of each backup, and when each backup is run.

Integration with third-party tools is done through the Backint interface that each provider makes available, and its configuration is entirely dependent on the tool being configured. Figure 1-18 shows the configuration options for SAP HANA backup with Backint.

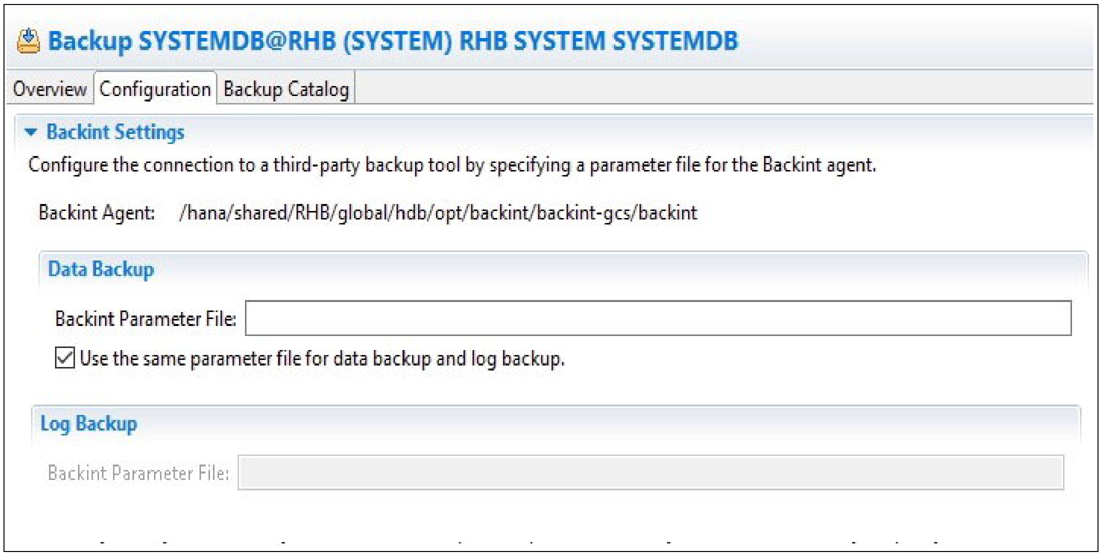


Figure 1-18 Configure options to SAP HANA backup with Backint

1.6.1 Basic backup tools

Sometimes, it is necessary to perform backups for certain components of the SAP HANA system that are not covered by the standard SAP HANA utility. Some issues that might occur include backups of the database configuration files, backups of logs for multiple reasons, or sometimes workarounds for disk space issues in the database. Basic tools provided by each operating system, Red Hat Enterprise Linux or SUSE Linux Enterprise Server, can help with these tasks.

The tools and utilities include command-line tools such as **`tar`**, **`cpio`**, and **`rsync`**, which are part of most Linux distributions. There are tools that work on specific operating systems, such as Snapper that works only on SUSE Linux Enterprise Server.

These tools can be a cost-effective and customizable option for performing backups of certain database components. However, they are not the official tool supported by SAP for a Full Backup or Incremental Backup of the data.

This section describes the benefits and drawbacks of some of the utilities that are available on the Linux distributions supported on the IBM Power System. Among them, cpio seems to have the best options, but it is not as easy to use as **`tar`**, which is the most commonly used backup tool of **`tar`**, **`cpio`** and **`rsync`**.

The tar utility

The **`tar`** utility collects many files into one archive file, which is often referred to as a .tar file. The utility is used for distribution of software among Linux users and for backups. The utility name comes from *tape archive* because it was originally developed for UNIX to write data to sequential I/O devices with no file system of their own. The current archive data sets that are created by **`tar`** contain various file system parameters, such as name, ownership, directory organization, timestamps, and file access permissions.

The **tar** version that is included in the SUSE Linux Enterprise Server for SAP Applications on IBM Power Systems is the GNU **tar** compilation. The GNU tar compilation eliminates many constraints that are present in regular **tar**, such as file name length and preservation of timestamps.

The **tar** utility is widely used in Linux due to its rich set of features and simplicity. A typical **tar** command looks like the following string:

```
# tar cvf - srcpath | (cd destpath && tar xv)
```

The zlib and zlibNX libraries

zlib is a free and open source software that is used for file compression and decompression. It falls under the category of lossless data compression. When zlib is used to compress files, it creates the .zlib extension. The compression ratio of zlib can vary from 2:1 to 5:1, depending on the data being compressed. The compression process in zlib employs the Deflate method, which encodes the data into compressed form. Conversely, the decompression process uses the Inflate method to decode the compressed data and generate the original file without any loss of data.

Note: One limitation of zlib is that it lacks a checksum mechanism to ensure data integrity.

For backup and archiving purposes, zlib compression can be applied to tar archives to save storage space and can maintain the integrity of the archived data. This is particularly useful when creating long-term backups or archives that need to be stored efficiently.

The zlibNX library is an enhanced version of the zlib compression library that supports hardware-accelerated data compression and decompression by using co-processors called Nest accelerators (NX) on IBM POWER10 processor-based servers. The zlibNX library is available for SUSE Enterprise Linux 15 SP4, Red Hat Enterprise Linux 8.6 and Red Hat Enterprise Linux 9.0 with BM POWER10 firmware level FW1010 or later.

The zlibNX library adheres to specific Request For Comment (RFC) standards, including RFC 1950 for the zlib compressed data format, RFC 1951 for the DEFLATE compressed data format, and RFC 1952 for the gzip file format. These RFC standards ensure consistency and compatibility among various implementations of the zlib library. Data compressed with zlibNX can be decompressed using the standard zlib library, and vice versa. This ensures seamless interchangeability of compressed data between the zlibNX and standard zlib implementations.

Using zlibNX with other tools for backing up and compressing SAP HANA data is considered a best practice due to its optimized performance and compatibility with IBM Power Systems, particularly IBM Power10. Because SAP HANA workloads can be resource-intensive, the use of zlibNX takes advantage of the hardware acceleration capabilities of IBM Power Systems, which provides faster compression and decompression speeds during the backup process.

The cpio utility

The **cpio** utility is a general file archiver that is the oldest backup utility. It can create and extract archives, and it has several options and features. You can use the tool to perform partial or full backups.

Table 1-4 on page 32 shows the **cpio** benefits and drawbacks.

Table 1-4 Strengths and weakness of cpio

Benefits	Drawbacks
Good for scripting with other tools, such as **find** and **gzip**.	Requires learning more about the specific version that is available because there are meaningful differences between releases.
Preserves hard links.	
File length limitation.	
Timestamps are preserved by default.	

The **cpio** utility can be a powerful backup tool. The main difference is its ability to accept the list of files to be backed up from standard input. However, this is also true for the **tar** GNU version that is embedded in the SUSE Linux Enterprise Server for SAP distribution.

> **Note:** Explore the tools options and secrets of **cpio** before using it on a production environment. Check the **man help** page for **cpio** for the specific Linux Enterprise Server distribution that you are using to understand its limitations and strengths.

The dd utility

The **dd** utility is used to convert and copy files. It is especially suitable for handling special device files. The **dd** command reads the *InFile* parameter or standard input, does the specified conversions, and then copies the converted data to the *OutFile* parameter or standard output. The input and output block size can be specified to take advantage of raw physical I/O.

> **Note:** The term *block* refers to the quantity of data that is read or written by the **dd** command in one operation, and is not necessarily the same size as a disk block.

The **dd** command is a low-level command that copies information from one place to another, and you must have **sudo** or `root` authorization to use it. The **dd** command does not require any knowledge of the structure of the data that it is copying. Therefore, unlike **tar** and **cpio**, it is not used to copy a group of files to a backup volume. It can copy a single file, a part of a file, a raw partition, or a part of a raw partition. It can also copy data from `stdin` to `stdout` and modify it while it is in route. Although it can copy a file, it has no knowledge of the file name or contents after it has done so.

> **Note:** The single biggest benefit of the **dd** command is that it can be used to copy unstructured data (raw data), such as a cold SAP HANA instance file system or set of files to a second place if necessary, and it can copy entire partitions among different disks.

To perform a copy, run **unmount** to unmount the partition, and then run the **dd** command as shown Example 1-1.

Example 1-1 Copying a file system by using the dd command

```
# dd if=/dev/mapper/VG_HANA1_OTHER-LV_HANACORE_SAP of=/dev/mapper/backup bs=1M
```

The **bs=1M** parameter ensures that **dd** copies the data in large chunks instead of issuing a request for each sector.

The pax utility

Another utility that is on SUSE for SAP Applications on Power Systems is the portable archive exchange (**pax**) utility. This utility produces a portable archive that conforms to the Archive/Interchange File Format that is specified in IEEE Std. 1003.1-1988. The **pax** utility extracts, writes, and lists members of archive files. The **pax** utility also copies files and directory hierarchies.

> **Note:** The pax utility maintains the integrity of sparse files during restoration by ensuring correct alignment and sizing of NULL-populated areas without unnecessarily allocating physical disk space for these parts. This approach optimizes disk usage and preserves the original file's size attributes.

A typical **pax** command creates a disk backup, as shown in Figure 1-19.

```
t@SAP2:~   root@hanna01:~
:~ # pax -wv -f /hana/backup/rd2_binnaries.pax /usr/sap/RD2
sr/sap/RD2/ directory
sr/sap/RD2/home/ directory
sr/sap/RD2/home/.config/ directory
sr/sap/RD2/home/.config/gconf/ directory
sr/sap/RD2/home/.fonts/ directory
sr/sap/RD2/home/.local/ directory
sr/sap/RD2/home/bin/ directory
sr/sap/RD2/home/.bash_history 50 bytes, 1 tape blocks
sr/sap/RD2/home/.bashrc 2256 bytes, 5 tape blocks
sr/sap/RD2/home/.profile 1791 bytes, 4 tape blocks
sr/sap/RD2/home/.i18n 305 bytes, 1 tape blocks
sr/sap/RD2/home/.xim.template 1952 bytes, 4 tape blocks
sr/sap/RD2/home/public_html/ directory
sr/sap/RD2/home/public_html/.directory 48 bytes, 1 tape blocks
sr/sap/RD2/home/.emacs 1637 bytes, 4 tape blocks
sr/sap/RD2/home/.inputrc 861 bytes, 2 tape blocks
sr/sap/RD2/home/.xinitrc.template 1112 bytes, 3 tape blocks
sr/sap/RD2/home/.muttrc 6043 bytes, 12 tape blocks
sr/sap/RD2/home/.gnu-emacs 18517 bytes, 37 tape blocks
sr/sap/RD2/home/.cshrc 2495 bytes, 5 tape blocks
sr/sap/RD2/home/.sapsrc.csh 176 bytes, 1 tape blocks
sr/sap/RD2/home/.sapsrc.sh 159 bytes, 1 tape blocks
sr/sap/RD2/home/.sapenv.sh 2688 bytes, 6 tape blocks
sr/sap/RD2/home/.sapenv.csh 2652 bytes, 6 tape blocks
sr/sap/RD2/home/Desktop/ directory
sr/sap/RD2/home/Desktop/sap-windows_cheat_sheet.desktop 440 bytes, 1 tape blocks
sr/sap/RD2/home/Desktop/sap-installation-wizard.desktop 673 bytes, 2 tape blocks
sr/sap/RD2/home/Desktop/customer_center.desktop 574 bytes, 2 tape blocks
sr/sap/RD2/home/Desktop/suse-connect-program.desktop 348 bytes, 1 tape blocks
sr/sap/RD2/home/.sap-icons 0 bytes, 0 tape blocks
sr/sap/RD2/home/.hdb/ directory
sr/sap/RD2/home/.hdb/SAP2/ directory
sr/sap/RD2/home/.hdb/SAP2/SQLDBC.shm 26656 bytes, 53 tape blocks
sr/sap/RD2/home/.hdb/SAP2/SSFS_HDB.KEY 92 bytes, 1 tape blocks
sr/sap/RD2/home/.hdb/SAP2/SSFS_HDB.DAT 675 bytes, 2 tape blocks
sr/sap/RD2/SYS/ directory
sr/sap/RD2/SYS/exe/ directory
sr/sap/RD2/SYS/exe/hdb -> /hana/shared/RD2/exe/linuxppc64le/hdb
sr/sap/RD2/SYS/profile -> /hana/shared/RD2/profile
sr/sap/RD2/SYS/global -> /hana/shared/RD2/global
```

Figure 1-19 using pax command to create a disk backup

The **pax** utility also supports incremental backups, and like most tools on Linux, it can be combined with other commands by using pipes.

The rsync utility

Similar to the standard UNIX or Linux **cp** command, **rsync** copies files from a source to a destination. The **rsync** command is typically used for synchronizing files and directories between two different servers, but it can also perform backups within a single server. With **rsync**, data backups can be performed locally and remotely, across disks and networks, and among mirroring servers.

The **rsync** utility determines which files differ between the sending and receiving systems by checking the modification time and size of each file. If the time or size are different between the systems, **rsync** transfers the file to the receiving system. This action requires only reading file directory information, so it is quick, but it misses any unusual modifications. However, the tool can perform a slower but comprehensive check if it is run with **--checksum**. This flag forces a full checksum comparison of every file that is present on both systems. Barring rare checksum collisions, this check avoids the risk of missing changed files at the cost of reading every file that is present on both systems.

The **rsync** utility supports copying several attributes of files and directories:

- Devices
- Links
- Owners, groups, and permissions

> **Note:** The **rsync** utility has good performance on most systems, but it should be tested in the specific context to be used to determine its capabilities.

The basic syntax of the **rsync** command is shown in Example 1-2.

Example 1-2 Syntax of the rsync command

```
# rsync options source destination
```

Here are the common options that are used with the **rsync** command:

-v Verbose.

-r Copies data recursively, but does not preserve time stamps and permissions while transferring data.

-a Archive mode allows copying files recursively, and it preserves symbolic links, file permissions, user and group ownerships, and time stamps.

-z Compresses file data.

-h Output numbers in a human-readable format.

For more information about **rsync**, including guidelines and examples to perform a backup, see Speaking UNIX: Advanced applications of sync. Although written for IBM AIX, the information at this resource has useful information that you can apply to a SUSE Enterprise Linux Server for SAP Applications and Red Hat Linux Enterprise for SAP Solutions on Power Systems.

The Snapper utility

Snapper is a utility that is embedded in the Linux distribution that manages Btrfs snapshots. Btrfs is a new copy-on-write (COW) file system for Linux that is aimed at implementing advanced features and focuses on fault tolerance, repair, and easy administration. BtrFS supports file system snapshots of subvolumes snapshots.

Snapper is available as a command-line interface (CLI) tool and a YaST module.

1.6.2 Advantages and features for backup and recovery on IBM Power Systems

In the context of Backup and Recovery on the IBM Power System platform, there are advantages and characteristics compared to other platforms on which SAP HANA can run:

- Hardware Integration. IBM Power Systems are designed to work optimally with SAP HANA. The close integration of hardware and software allows for efficient and reliable backup and recovery operations by using the capabilities of Power Systems' processors, memory, and storage.
- Granular Resource Allocation. IBM Power Systems offer granular resource allocation, allowing you to allocate CPU, memory, and storage resources precisely to SAP HANA instances. This ensures that backup operations can be performed without impacting the performance of other workloads running on the system.
- Snapshots and FlashCopy Support. Power Systems often use advanced storage technologies like IBM Storage Virtualize to support efficient snapshots and FlashCopy features. These technologies enable rapid and space-efficient point-in-time copies of SAP HANA data for backup and recovery purposes.
- High Availability Features. Power Systems provide high availability features, such as IBM PowerHA®, that can be used to create highly available SAP HANA configurations. These features ensure continuous availability and quick recovery during hardware or software failures.
- Integration with IBM Storage Protect. IBM Storage Protect (formerly known as Tivoli® Storage Manager) is a robust data protection solution that integrates well with IBM Power Systems and SAP HANA. It allows for centralized backup management, data deduplication, and offsite storage for disaster recovery.
- SAP HANA System Replication. SAP HANA System Replication is supported on IBM Power Systems and enables asynchronous replication of data between different Power Systems servers. This adds an extra layer of protection and redundancy for SAP HANA databases.
- Support and Expertise. IBM has a wealth of experience and expertise in SAP HANA deployments on Power Systems. This means that customers can benefit from the support and guidance of IBM professionals with in-depth knowledge of both SAP HANA and Power Systems.

1.6.3 Integration with third party tools for backup and recovery

Third-party tools that integrate with SAP HANA offer several features that improve the efficiency, reliability, and ease of managing backup and recovery. Some common features that they typically provide include:

- Fast Backup and Restore. These tools allow for quick backups of the SAP HANA database, which minimizes downtime and improves backup and recovery efficiency.
- Compression and Deduplication. Many third-party tools offer compression and deduplication capabilities to reduce the storage space required for backups, optimizing resources, and reducing storage costs.
- Integration with External Storage Infrastructure: Several tools integrate with external storage systems, such as SAN or NAS, using advanced storage capabilities and snapshot technologies for more efficient data protection.
- Flexible Backup Scheduling. These tools often provide options for flexible backup scheduling based on business needs, which allows for customized backup schedules, frequencies, and retentions.
- Centralized Management. Third-party tools provide centralized management interfaces to administer and monitor backups across multiple SAP HANA instances on different servers, simplifying administration and monitoring.
- Point-in-Time Recovery. These tools enable recovery to specific points in time, facilitating data restoration to earlier states in case of errors or issues.

- Integration with Automation and Orchestration Tools. Some third-party tools integrate with automation and orchestration platforms, enabling automated backup and recovery tasks to improve operational efficiency.

The following are some of the third-party tools supported by IBM Power Systems to back up SAP HANA include:

- IBM Storage Protect (formerly known as Tivoli Storage Manager) is a third-party backup tool that can use the SAP HANA Backint API interface. What sets Storage Protect apart is its exceptional scalability compared to other backup systems currently available. With Storage Protect, a single server can perform the same amount of work that would typically require dozens of servers with competitor products. Key metrics used for this comparison, especially against scale-out competitor products, include data density measured in kilowatts per terabyte (KW/TB) and rack unit per terabyte (U/TB) served. In a large multi-server environment, Storage Protect operates at the individual server level. Therefore, the failure of a single server does not disrupt the entire infrastructure's operation. For scale-out storage solutions, Storage Protect integrates seamlessly with Storage Scale or IBM Cloud Object Storage. Storage Protect offers infrastructure cost reduction through its efficient scale-up capabilities. The server supports large data volumes and storage with just a single node, and also makes effective use of tape, block, or object storage media. A unique aspect of Storage Protect is that it does not require any additional hardware, like appliances, to enable features or scale.
- Veritas NetBackup is a backup and recovery solution that has also been certified for use with SAP HANA on Power Systems. Veritas NetBackup is a widely used enterprise-level data backup and recovery software solution developed by Veritas Technologies, a data management company. NetBackup is designed to provide robust data protection, disaster recovery, and data management capabilities for diverse IT environments, including physical and virtual servers, databases, applications, and cloud-based systems.
- Cohesity DataProtect for SAP HANA is a comprehensive data protection and backup solution designed specifically for SAP HANA databases that integrate with SAP HANA's native backup and recovery capabilities to provide advanced data protection features and streamline the backup process.
- Dell EMC NetWorker is a backup and recovery software solution that supports SAP HANA databases. It provides data protection and backup management capabilities, including full and incremental backups, data deduplication, and snapshot-based backups for SAP HANA environments.
- Commvault is a data management and protection solution designed to support SAP HANA databases. Commvault is a leading data management software company that offers a comprehensive range of solutions for data backup, recovery, archiving, and analytics.
- SEP Sesam is an enterprise-level data backup and recovery software solution developed by SEP AG, a German-based software company. SEP Sesam provides comprehensive data protection and disaster recovery capabilities for various IT environments, including physical and virtual servers, databases, applications, and cloud-based systems.
- Data protection and disaster recovery capabilities for various IT environments, including physical and virtual servers, databases, applications, and cloud-based systems.

Table 1-5 Supported third-party tools for SAP HANA on IBM Power Systems

Vendor	Backup tool	Power SAP HANA 1.0	Power SAP HANA 2.0
Commvault	Commvault (formerly Simpana) Hitachi Data Protection Suite	Yes	Yes

Vendor	Backup tool	Power SAP HANA 1.0	Power SAP HANA 2.0
Cohesity	DataProtect for SAP HANA	No	Yes
EMC	Networker Data Domain Boost for Databases and Applications Data Domain Boost for Enterprise Applications	Yes	Yes
IBM	Storage Protect for Enterprise Resource Planning (formerly Tivoli Storage Manager)	Yes	Yes
SEP	SEP Sesam	Yes	Yes
Veritas	NetBackup	No	Yes

Note: It is recommended that you periodically review SAP Note 2031547 - Overview of SAP-certified 3rd party backup tools and associated support process.

1.6.4 Data Backup and Data Snapshot

This section compares data backup and data snapshot. 1.7, “High availability and disaster recovery” on page 38 provides a comparison of the benefits, drawbacks, and backup durations for two delta backup types and a data snapshot.

Table 1-6 Data Backup and Data Snapshot comparison

	Data Backup to File	Data Backup Using Backint	Data Snapshot
Advantages	► Integrity checks at block level ► Can be encrypted	► Integrity checks at block level ► Integrated into existing data center infrastructure ► Third-party backup tool offers additional features ► Backups are immediately available for recovery	► Fast ► Negligible impact on network ► Can be encrypted ► No third-party backup integration
Disadvantage	► Requires additional storage ► Generates additional network load ► File system must be monitored ► More time is needed to make backups available for recovery	Generates additional network load	No integrity checks at block level

	Data Backup to File	Data Backup Using Backint	Data Snapshot
Backup duration	► IO-bound (reading from data volume, writing to target. ► Network-bound (writing to target file system).	► IO-bound (reading from data volume) ► Network-bound (writing to backup server	Negligible

1.7 High availability and disaster recovery

Although business continuity plans are a critical requirement for successful businesses, some customers have no clear picture of the key considerations for designing a high availability (HA) and disaster recovery (DR) (HADR) strategy. Power Systems servers have a set of features for business continuity and fault-tolerance, and SAP HANA has a solution for business continuity, but to design a proper solution, you must analyze and define the business requirements.

When creating a strategy to decrease downtime and eliminate single points of failure, the most important question is "What are the availability requirements?" Most customers say that their systems cannot be unavailable at all. When the consultants show that every improvement increases costs, the customer reassesses and defines their real requirements.

Availability is measured by a service-level agreement (SLA). To obtain your availability SLA, you must follow a structured assessment line:

- List the critical business systems
- Determine these systems' operational hours
- Assess the systems' utilization beyond operational hours

The answers to the assessment make it possible to map system usage based on its consumption. As expected, the usage is higher during business hours, but with remote access, usage often occurs outside of business hours too. Assessing the number of users per hour can help to identify the criticality of the system throughout the day.

Other considerations that need to be assessed are service consumption by other systems, and background processing to generate reports or to use an ETL process on data. Some solutions that use SAP S/4HANA include an embedded ETL process, which forks use cases that might be important to map.

All of this information must be collected to determine the system's true availability.

The SLA defines the contracted auditability of the system. You can use the information shown in Table 1-7 to determine the maximum allowed system outages per year based on your SLA target.

Table 1-7 Typical service-level agreements

Number of nines	% Uptime	Maximum Annual downtime
Six (6)	99.9999	32 seconds
Five (5)	99.999	5 minutes 16 seconds
Four (4)	99.99	52 minutes 36 seconds

Number of nines	% Uptime	Maximum Annual downtime
Three (3)	99.9	8 hours 45 minutes 57 seconds
Two (2)	99.0	3 days 15 hours 39 minutes 29 seconds
One (1)	90.0	36.5 days

SLAs primarily focus on addressing unplanned downtime rather than planned outages. The level of uptime guaranteed in the SLA directly correlates to the amount of time the service remains unavailable. Each additional “9” after the decimal point indicates a significant reduction in downtime. Resilience takes various forms.

For instance, IBM Power servers are designed with built-in redundancy, boasting a 99.999% hardware availability. Similarly, IBM Flash storage offers a 99.9999% availability. However, if the switch connecting the fiber cables between the server and storage has only a 99.9% availability, the overall infrastructure availability is limited to that of the lowest component. To address this, many clients opt to install a second switch on an entirely different fiber channel path to ensure higher availability.

Some other requirements need to be assessed:

- Size of the system. Larger systems are slower to restart (if needed) than smaller ones. Factors such as running transactions and the number of users are also included. Sometimes, the system administrator must wait for a set of transactions to finish its processing to avoid a bigger problem affecting the entire system.
- Complexity. Another component is the complexity of the overall solution. The Power Systems servers on which the SAP HANA is running are used by at least one application server and other components. Good documentation about the relationship among the systems is fundamental to ensure a quicker response time for problem determination and resolution.
- Extra factors. Other factors might affect the resolution of the issue and the SLA, such as availability of the personnel to support the system, problems with the SAP HANA software that require SAP development support, and conflicts between support teams. Although not all aspects of the design can be mapped, include as much information as possible.

In addition to the SLA for HA systems, consider key performance indicators (KPIs) that are related to outages that require DR. The two standard KPIs are (Figure 1-20):

- Recovery Point Objective (RPO). The maximum amount of data loss that an organization can tolerate, it is measured in time.
- Recovery Time Objective (RTO). The maximum period that is required to recover the system.

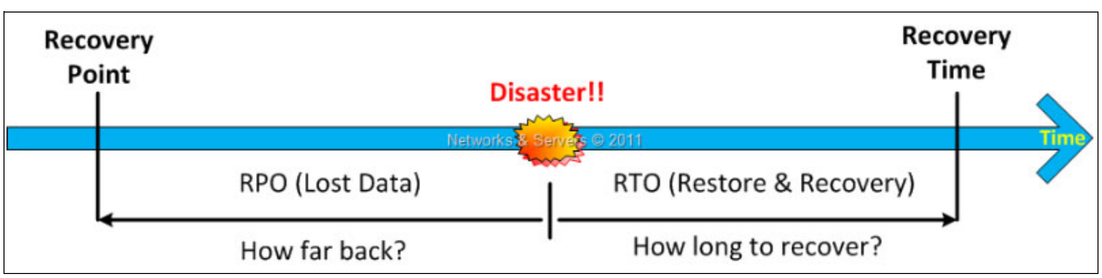


Figure 1-20 Disaster recovery: RPO and RTO

After the availability requirements are defined from a business point of view, there are options to implement the requirement for SAP HANA on Power Systems. Although there are many considerations for planning HADR, this publication focuses on the options that are relevant for

SAP HANA on Power Systems instead of every component, such as CPUs, storage devices, input/output devices, firmware, applications, power, and cooling.

1.7.1 SUSE Linux Enterprise High Availability Extension

When running your SAP HANA environment on IBM Power and using SUSE Linux Enterprise Server, the recommended HA solution is SUSE Linux Enterprise High Availability Extension, which is a feature that was added in SUSE Linux Enterprise Server 11.

SUSE Linux Enterprise High Availability Extension is based on open source code and is included by default on SUSE Linux Enterprise Server for SAP Applications. It includes HA service and application clustering, file systems and clustered file systems, network-attached storage (NAS), network file systems, volume managers, storage area network (SAN) and drivers, and tools to manage these components.

SUSE Linux Enterprise High Availability Extension is an important part of building an HA solution with SAP HANA. SUSE Linux Enterprise High Availability Extension supports the following application servers:

- SAP WebAS ABAP Releases 6.20 - 7.30.
- SAP WebAS Java Releases 6.40 - 7.30.
- SAP WebAS ABAP + Java Add-In Release (For release 6.40, see SAP Note 995116).
- For releases 6.20 - 7.30, Java is not monitored by the cluster.

The SAP resource agent starts, stops, and monitors DB instances (such as SAP BW solutions with a main instance on SAP HANA, and JAVA instances DBs on SAP Sybase ASE).

1.7.2 Red Hat Enterprise Linux High Availability Solution for SAP HANA

Red Hat Enterprise Linux for SAP Solutions offers Pacemaker resource agents that facilitate the automated transition from a primary to a secondary data center. This solution encompasses concepts, planning, and high-level instructions for implementing an automated SAP HANA scale-out replication solution by using RHEL for SAP Solutions. By using RHEL for SAP Solutions, native SAP HANA replication and failover mechanisms can be further enhanced through the integration of resource agents that facilitate the automated switchover from the primary to the secondary data center.

The SAP HANA scale-out solution can be expanded to encompass multiple independent scale-out solutions, acting as additional mirrors. The process of system replication employs the active/passive method, ensuring optimal performance, and relies solely on network communication, which eliminates the need for extra infrastructure components.

Pacemaker automates the system replication process in case of critical component failures. It evaluates data from both the scale-out environment and the system replication process to ensure uninterrupted operation. The cluster manages the primary IP address, which serves as the connection point for clients accessing the database. Therefore, when database takeover is triggered by the cluster, clients can still connect to the active instance.

Red Hat Pacemaker for SAP HANA has several key advantages:

- Red Hat Pacemaker provides a robust and reliable high-availability solution for SAP HANA. It ensures continuous availability of the SAP HANA database, which minimizes downtime and maximizes system uptime.
- Pacemaker automates the failover process in case of component failures, such as hardware or network issues. This automated failover mechanism reduces the need for

manual intervention and helps to ensure a quick and seamless transition to backup resources.

- You can use Red Hat Pacemaker to scale SAP HANA systems by enabling the implementation of scale-out solutions. It can manage multiple independent scale-out environments that can act as additional mirrors to increase capacity and performance.
- Pacemaker efficiently manages critical resources, such as IP addresses and database instances. During a failover, it ensures that clients can continue to connect to the active instance by maintaining uninterrupted access to the SAP HANA database. The Pacemaker solution for SAP HANA has undergone extensive testing and is a proven and reliable solution. It has been designed to work seamlessly with SAP HANA, providing a stable and optimized environment.
- Red Hat Pacemaker offers flexibility and customization options to tailor the high-availability setup according to specific requirements and preferences. It can be configured to adapt to varying workload demands and system configurations.

To understand the process of setting up Automated HANA System Replication in Scale-Up within a Pacemaker cluster on RHEL releases that are supported, see Automating SAP HANA Scale-Up System Replication using the RHEL HA Add-On.

1.7.3 SAP HANA HADR options on Power Systems servers

SAP HANA has a tool-set of options that enables HADR. The solution is resilient because of the following set of features:

Host Auto-failover

With this option, one or more standby hosts are added to an SAP HANA system. If one of the active hosts fails, then the SAP HANA watchdog service assesses and automatically starts the failover process on the standby server. Multiple standby hosts can be added to increase the SLA goals of the system, which addresses full system outages by providing a hardware-ready node, as shown in Figure 1-21.

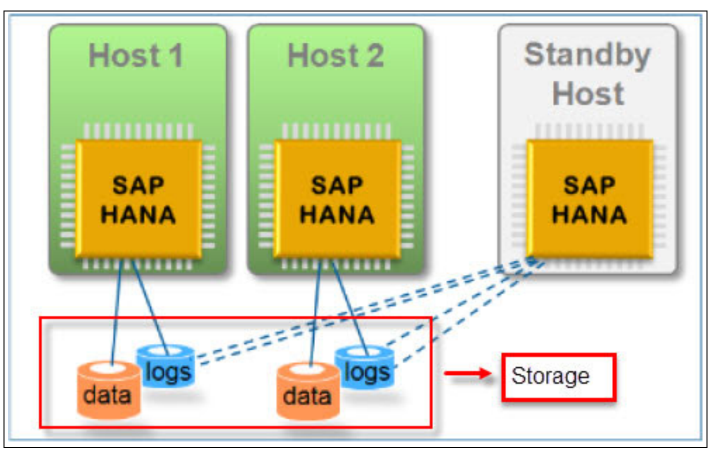


Figure 1-21 SAP HANA: Host-Auto Failover

Because the data is preinstalled in the standby system, both RPO and RTO on this scenario tend toward zero.

System Replication

In this architecture, a fully redundant standby system is configured and SAP HANA replicates all the data to that secondary system. It is a standard SAP HANA feature because data is frequently updated in the target standby system, as shown in Figure 1-22.

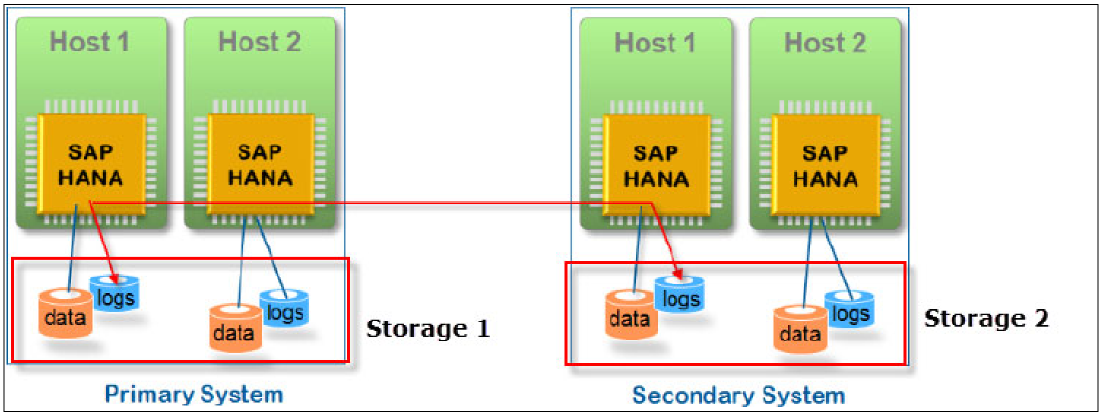


Figure 1-22 SAP HANA: System Replication

This scenario is similar to shadow DB solutions. The automation of the failover is done by using SUSE Linux Enterprise High Availability Extension or Pacemaker on Red Hat Enterprise Linux Power Systems. Additionally, this scenario is used for HADR situations. This solution has a low RTO and RPO.

System replication facilitates a rich set of configurations. For more information about data loading with synchronous and asynchronous options, see SAP Note 2407186.

Storage Replication

Storage Replication is another solution that is used for DR scenarios. IBM Storage systems can integrate that solution, which allows integrity checks with SAP HAN. For synchronous replication, IBM Storage systems commit transactions after the replication is complete. The distance between the primary and secondary sites is limited to a maximum of 100 kilometers with less than approximately 5 microseconds of latency per kilometer.

As with System Replication, this solution offers an excellent RPO, and covers large hardware or even entire site failures, but it requires stable and broad bandwidth.

When you use the Storage Replication solution, SAP system management can be automated with the SUSE Linux Enterprise High Availability Extension or Pacemaker with Red Hat Enterprise Linux, as shown in Figure 1-23 .

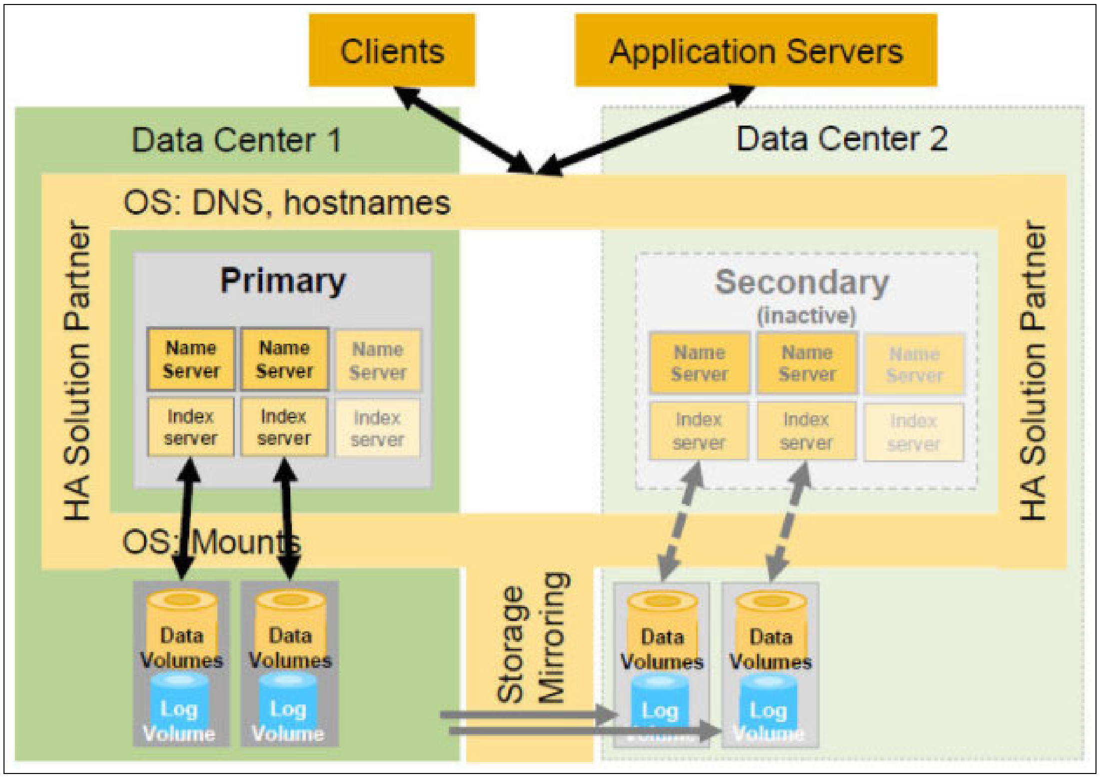


Figure 1-23 SAP HANA: Storage replication

IBM VM Recovery Manager

IBM VM Recovery Manager (VMRM) is a widely used VM management software offered by IBM that serves both High Availability (HA) and Disaster Recovery (DR) purposes. For High Availability, VMRM can be used with Logical Partition Mobility (LPM) to seamlessly transfer LPARs from one server to another during planned maintenance. Additionally, VMRM can automatically restarting servers or LPARs, if required, and it can monitor applications through the VM agent, which helps to ensure prompt application restart when necessary.

The Disaster Recovery solution does not require the deployment of backup virtual machines (VMs) for disaster recovery purposes. Instead, only the storage logical volumes containing the OS, application binaries, and user data are blocked and replicated to the DR site. In case of a disaster, these replicated volumes are used to rebuild the VM, which can lead to reduced software license and administrative costs. VMRM DR supports both planned and unplanned DR events and can be valuable for DR rehearsal functionalities as shown in Figure 1-24

Enabling VMRM HADR involves the use of the following subsystems:

- KSYS or controller system is installed in an AIX LPAR and is responsible for monitoring the entire environment and triggers HA and DR operations when necessary.
- HMCs are located in the active site and the backup site and manage the IBM Power Systems servers.
- Various VIOS partitions within hosts of the active site and the backup site virtualize and manage storage resources for the hosts' virtual machines.

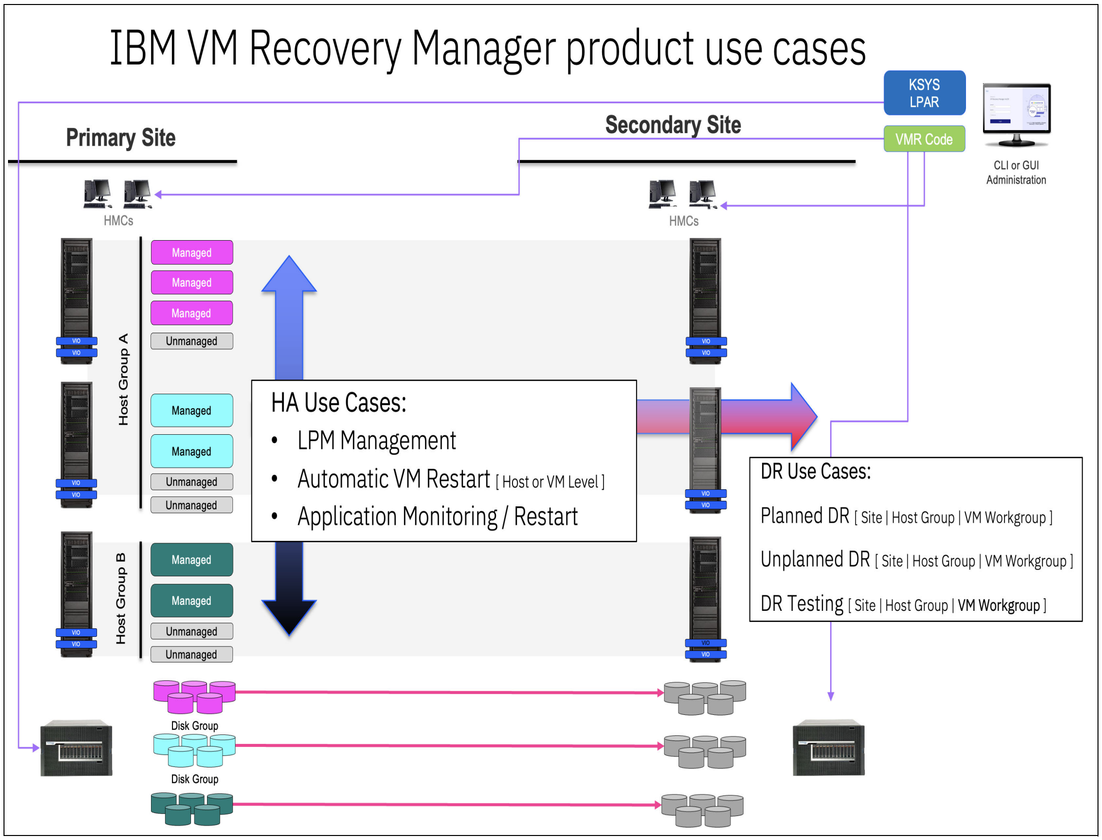


Figure 1-24 IBM VM Recovery Manager use cases

VM Recovery Manager HA

IBM VM Recovery Manager HA for Power Systems is a user-friendly high-availability solution designed to facilitate the recovery of virtual machines (VMs), also known as logical partitions (LPARs), in an automated manner. The VM Recovery Manager HA solution employs VM restart technology to accomplish VM recovery. This technology relies on an out-of-band monitoring and management component that initiates VM restart on a different server when a host infrastructure failure occurs. Unlike conventional cluster-based technology that deploys redundant hardware and software components for real-time failover, the VM restart technology streamlines the process without adding complexities associated with clustering. The VMRM HA solution is particularly suitable for ensuring high availability for numerous VMs. In the depicted scenario, SAP HANA can be replicated to the failover server through either HSR or storage-level replications of the SAP HANA data and transaction logs by using block-level storage replication. In the context of SAP HANA with HSR configuration and a VM Recovery Manager HA topology, an integrated VM agent for SAP HANA is available. Assuming two SAP HANA LPAR instances with the VM Agent for HANA installed on both primary and secondary LPARs, the agent integration enables monitoring of the virtual machine status, SAP HANA database, and HANA HSR status by the KSYS. KSYS can manage failures, such as HSR role changes, virtual IP redirection, and replication adjustments.

For more detailed VMRM HA implementation, additional information can be accessed in these sources:

- SAP HANA Agent for IBM VM Recovery Manager
- *IBM VM Recovery Manager for Power Systems Cookbook*, SG24-8539
- *IBM Power Systems High Availability and Disaster Recovery Updates: Planning for a Multicloud Environment*, REDP-5663

When you use block-level replication, it is not mandatory to have an existing SAP HANA VM instance on the failover server. The KSYS deploys the VM when it is notified by the VM Agent about a primary VM failure. The benefit of this approach is that the failover server can be used for other applications and remains resource-free until a failure in the primary VM occurs. However, it might take longer for VM deployment and restart because of the database size. The time span can be potentially tens of minutes or an hour. In this deployment, VM Recovery Manager can be configured in a DR-type topology and provide DR orchestration without monitoring application-level failures. Alternatively, a VMRM HADR topology can be implemented in which both the VM agent capability and DR orchestration with block-level replication work together.

Note: VMRM HA topologies with the VM Agent for SAP HANA offer a straightforward HA solution, but some clients prefer to implement Linux Pacemaker clusters for high availability and use VMRM for DR instead. For more information, see Table 1-8.

VMRM – DR for SAP HANA

In a DR exclusive topology, the KSYS would enable user-initiated DR operations, as opposed to the automated operations in an HA topology. In this setup, no remote LPARs need to be configured. Instead, a designated server acts as a landing zone with sufficient CPU and memory capacity to accommodate managed LPARs if the administrator triggers a planned or unplanned operation. As previously mentioned, one significant advantage of VMRM is that there is no pre-deployed LPAR (VM) on the secondary server until a failure occurs. This feature allows the resources at the secondary site to be used for alternative VM instances, such as Test or QA systems.

Table 1-8 compares Pacemaker with VMRM HA and VMRM DR for SAP HANA on IBM Power Systems.

Table 1-8 Pacemaker vs VMRM HA and DR

Solution	Pacemaker for SAP HANA	VM Recovery Manager – HA for SAP HANA	VM Recovery Manager – DR for SAP HANA
Base technology	Linux OS	IBM PowerVM and shared storage	IBM PowerVM and Storage Replication
Supported by SAP	Yes	Yes	Yes
Automated failover	Yes	Yes	No
Ideal for DR	No	No	Yes
Ideal for HA	Yes	Yes	Yes
Ease of Management	Easy	Advanced	Moderate
Reuse secondary for nonproduction workload	Partly	Partly	Yes
DR Rehearsal feature for nondisruptive DR testing	No	No	Yes
Cloud support	Yes	No	No
License cost	Part of Linux OS	Low cost	Moderate cost

HA and DR are a critical aspect of SAP applications, especially for the enterprise-class customers that demand 24x7 operations spanning the globe. SAP and IBM recognize the importance of providing robust HADR solutions, and this is equally true for the SAP HANA database. To better understand and compare the HADR capabilities available for SAP systems, refer to the Table 1-9 which summarizes the advantages and disadvantages of each option.

Table 1-9 Scenario summaries

Scenario	Advantages	Disadvantages
Host Auto-Failover	► Used to complement other solutions or by itself ► Automatic detection and failover with SAP HANA internal tools	► Requires access to the DB storage of the standby host
System Replication	► Supports HADR ► Active/Active ► RPO=0 (when synchronous) ► RTO<=1 ► Full performance when the takeover completes ► Supports a single-host system (shared storage not required)	► Requires a dedicated live standby system ► Requires stable and broad bandwidth ► Requires a solution for client connection recovery upon failover (DNS or Virtual IP address based)
Storage Replication	► Supports DR. ► RPO=0 (with sync replication). ► Secondary system can be used for other purposes.	► Requires a complementary solution for starting the system on the DR side ► RTO varies based on the other components solution ► Requires stable and broad bandwidth ► Limitation on the distance between DBs
VMRM	► Supports HADR ► Centralized administration ► Reduced RTO ► Supports reduced capacity on the DR site ► Flexible capacity between source and target	► Longer RTO due to SAP HANA cold start ► Not optimal initial performance

1.7.4 IBM Technology Lifecycle Services for SAP HANA on IBM Power

IBM Technology Lifecycle Services (TLS) has developed a comprehensive support solution that is tailored for SAP HANA on IBM Power Systems. These services can be used as a stand-alone offering to support SAP HANA on Power or integrated into a broader multi-vendor support solution provided by IBM.

The use of IBM Proactive Support for SAP HANA on Power has several benefits:

- It helps clients prevent unplanned downtime by offering proactive service from a leading industry expert with extensive experience supporting both IBM and multi-vendor products.
- Clients can protect their brand reputation and ensure customer satisfaction through fast, 24-hour access to experienced technical support because of IBM's global presence.
- The support solution simplifies the clients' IT operations by augmenting their in-house IT staff, which allows them to focus on innovation, growth, and digital transformation.

- In case issues do occur, IBM Proactive Support enables clients to resolve problems faster, which reduces downtime and costs.
- With a holistic set of software and hardware support services, IBM can identify dependencies across the clients' entire IT portfolio, optimizing IT and user productivity to drive business transformation.

To benefit from IBM Proactive Support for SAP HANA on Power, clients must fulfill certain prerequisites:

- They should have a 24x7 IBM Hardware Maintenance contract or a 24x7 warranty service upgrade for IBM Power Systems running SAP HANA.
- Red Hat Enterprise Linux (RHEL) for SAP Solutions or SUSE Linux for Enterprise Servers for SAP applications subscription should be installed and operational.
- Clients need an IBM TLS Proactive Support contract.
- An IBM TLS Support Line for the Linux operating system service contract.

The premium service is designed to simplify clients' IT operations, reduce costs, increase availability, and optimize system performance.

Clients will receive proactive support from a dedicated account manager in IBM's Center of Excellence, acting as their single point of contact.

IBM offers a comprehensive support experience covering the entire solution stack, with a 30-minute committed response time for Severity 1 issues, and a 2-hour response time for Severity 2, Severity 3, and Severity 4 tickets. If the problem is not directly related to infrastructure or software, IBM collaborates with the SAP HANA application support team to resolve the issue.

The service also includes health check analyses for hardware, Hardware Management Console (HMC), PowerVM, operating system, and the SAP HANA infrastructure stack. IBM ensures microcode and release management during customer prime-shift hours and offers extended IBM Center of Excellence Support (Hot Standby).

1.7.5 Sustainability achieved through IBM Power Systems Migration

IBM Power provides businesses with the ability to respond rapidly to customer demands through scalable computing solutions. With the Power10 processor, IBM has achieved a reduction in carbon footprints compared to older Power servers, and simultaneously delivers a maximum of twice the compute scalability. This means that data centers running on IBM Power10 require fewer cores and less power to achieve the same or even greater computing capability compared to other platforms.

For instance, the IBM Power E1050 consumes up to 59% less energy for the same workload when compared to the IBM Power E850, and 47% less energy when compared to the IBM Power® System E950 as shown in Figure 1-25 .

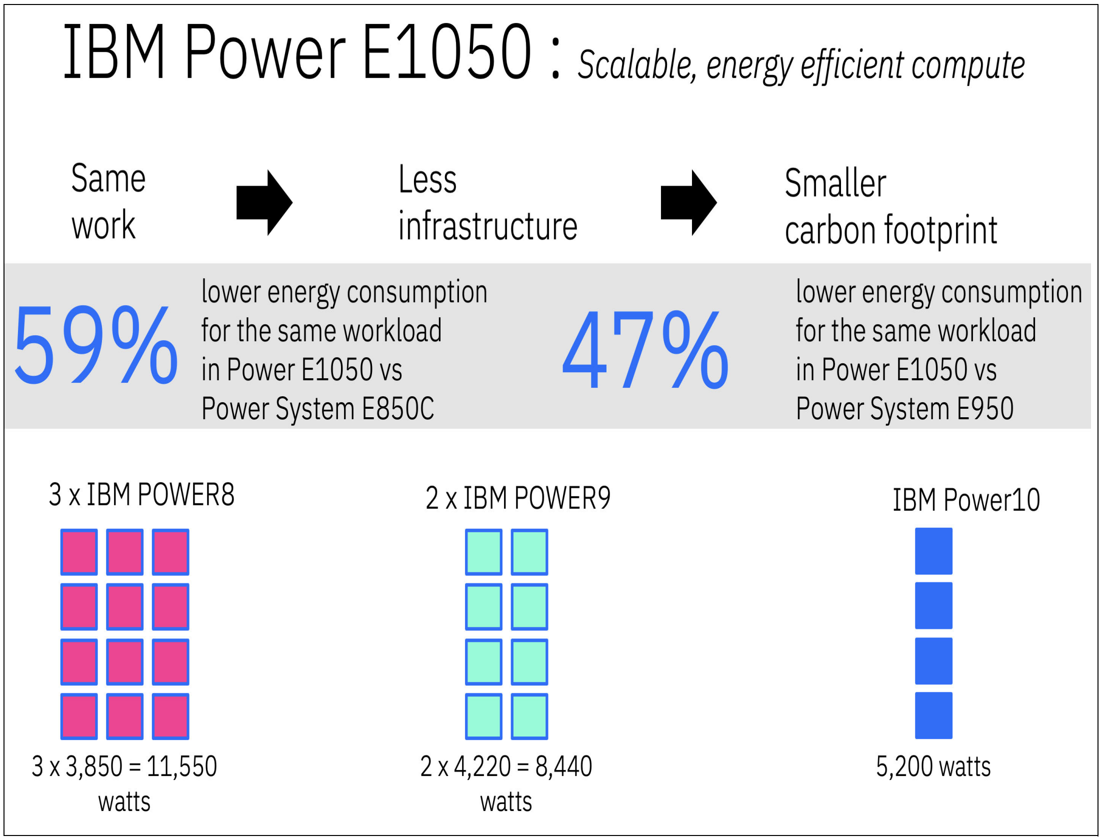


Figure 1-25 IBM Power E1050 vs E950 and E850 energy consumption

These energy savings contribute to supporting mission-critical environments in businesses, and also enables consolidation capabilities. IBM Power is committed to supporting clients in reusing many components from previous Power servers, including I/O adapters and internal storage.

By doing so, the organization not only reduces environmental stress by minimizing the production of new components but also minimizes waste from discarded components. Some components can even be used across multiple generations for up to 8 years from their introduction. The highly resilient nature of IBM Power systems further enhances sustainability efforts. As the most reliable non-mainframe server in the industry[3], IBM Power10 leads to less downtime and requires fewer redundant components and systems.

Additionally, IBM Power10 processor-based servers are designed to be environmentally friendly from end-to-end. IBM has set a goal to achieve net-zero greenhouse gas emissions by 2030, and IBM Power takes pride in driving technology advancements to deliver highly efficient, resilient, and sustainable products for its clients.

IBM Power and SAP HANA play a crucial role in making use of sustainability through various approaches and practices:

- IBM Power Systems are designed to be highly energy-efficient. They use advanced technologies such as IBM Power Architecture and energy-saving features such as dynamic frequency scaling to optimize power consumption. By running SAP HANA workloads on energy-efficient infrastructure, organizations can reduce their carbon footprint and can lower energy costs.

[3] https://www.scribd.com/document/679648841/itic-2022-global-server-hardware-server-os-reliability-report-feb-2023

- IBM Power Systems support server consolidation and virtualization, which allow multiple SAP HANA instances to run on a single physical server. This reduces the overall hardware footprint and leads to lower power consumption and better use of resources.
- IBM has made significant strides in building green data centers that prioritize energy efficiency and environmental sustainability. By deploying SAP HANA on IBM Power in these data centers, businesses can further contribute to sustainability efforts.
- Cloud-based deployments of SAP HANA on IBM Power Systems offer the advantage of shared resources and dynamic provisioning. This flexibility enables organizations to scale resources based on demand, which optimizes resource usage and reduces waste.
- SAP HANA, as an in-memory database, enables real-time data processing and analytics. This speed and efficiency help organizations make data-driven decisions quickly, leading to more sustainable business practices.
- SAP HANA on IBM Power Systems allows for greater data center efficiency through workload optimization and resource allocation. This leads to better use of resources, reduced power consumption, and lower operational costs.
- By streamlining business processes and enabling real-time insights, SAP HANA on IBM Power Systems can help organizations optimize their supply chains and reduce carbon emissions associated with transportation and logistics.
- SAP HANA and IBM Power support various analytics and reporting capabilities that help organizations identify opportunities for sustainable practices. Businesses can analyze their operations, supply chain, and energy consumption to make informed decisions to reduce their environmental impact.

The combination of IBM Power Systems and SAP HANA offers businesses the potential to improve their sustainability practices by optimizing energy usage, reducing hardware requirements, and enabling smarter and more sustainable business operations.

1.8 Cloud

IBM Power systems are ideal for businesses that need high-performance, scalability, and a secure infrastructure for their mission-critical applications. IBM Power Virtual Server (PowerVS) is a fully managed service offering from IBM, which includes all the hardware, software, and maintenance.

The expansion of the IBM Power10 family is engineered to establish one of the industry's most flexible and broadest range of servers for data-intensive workloads, such as SAP S/4HANA, from on-premises workloads to hybrid cloud.

RISE with SAP is a business transformation as a service (BTaaS) offering from SAP that provides businesses with a comprehensive set of tools and services to help them accelerate their digital transformation. RISE with SAP helps businesses to become more agile by moving their SAP applications to the cloud, which allows businesses to quickly adapt to changes in the market and customer demands. IBM's RISE with SAP, premium supplier option offers clients the choice of a single partner to manage everything within RISE with SAP, from software, support, and infrastructure, such as IBM Power in IBM Cloud to technical managed services.

RISE with SAP, premium supplier option combines the power of SAP HANA and IBM Power servers on IBM Cloud, which helps businesses to accelerate their digital transformation and achieve their business goals.

At the time of writing, IBM is enabling IBM Power10 for the RISE with SAP, premium supplier option in IBM Cloud. This offer provides the following advantages for clients:

- Increased performance. Power10 can deliver a maximum of 2.5x better performance than previous generations of SAP HANA deployments. This means that businesses can run their SAP HANA applications faster and more efficiently.
- Improved scalability. Power10 can support larger SAP HANA deployments. For example, Power10 can support 150% more memory compared to Power. This means that businesses can grow their SAP HANA deployments without having to worry about performance or scalability limitations. Additionally, it helps SAP customers to consolidate their infrastructure footprint.
- Enhanced security. Power10 enhanced security features help businesses protect their SAP HANA data from unauthorized access.
- Lower Cost. Power10 helps businesses to reduce their IT costs by moving their SAP applications to the cloud, which includes SAP HANA, which is the core of most SAP Applications. Businesses no longer need to invest in and maintain their own on-premises infrastructure. IBM Power10 supports both cloud and on-premises.
- Easier Management. With Power10, IBM maintains the infrastructure, so the customer can focus on business processes without regard to the underlying infrastructure.

For the product RISE with SAP, Premium Supplier Option on Power10, pricing is based on the number of SAP HANA users and the amount of storage that is required. IBM also offers various pricing options, including pay-as-you-go and subscription-based pricing.

IBM Cloud also offers IaaS as high-performance SAP-certified infrastructure to run production, dev/test, and HADR environments. The IaaS portfolio of IBM Cloud consists of the following features:

- IBM Power Virtual Server (only SAP-certified IaaS on Power for SAP HANA deployments)
- IBM Cloud for VMware Solutions (only SAP-certified IaaS on VMware)
- IBM Cloud Virtual Servers
- IBM Cloud Bare Metal Servers

The use of the IBM Cloud for SAP Solution includes certain benefits:

- Accelerate time to value
- Mitigate IT infrastructure risk
- Govern resource configurations and centrally manage compliance
- Gain complete authority over data at rest, in transit and in use
- Protect data across the compute lifecycle
- Deploy locally and scale globally with resilient and highly available infrastructure
- Empower an intelligent, more sustainable enterprise
- To find more details and offers from IBM Power for SAP HANA, see:
  - Accelerate SAP ERP modernization on IBM Cloud
  - *SAP HANA on IBM Power Systems Virtual Servers: Hybrid Cloud Solution*, REDP-5693
  - IBM Training: IBM Cloud for SAP Specialty

2

Hardware advantages of IBM Power9 servers

IBM POWER processor-based servers provide many technological capabilities for running SAP HANA on Power Systems. The IBM Power9 processor has capabilities that enhance analytic workloads, high-density cloud virtualization, emerging artificial intelligence (AI), and large in-memory database (DB) solutions. Advantages that the IBM Power9 processor-based servers provide include high memory bandwidth architecture, simultaneous thread concurrency at the core, large processor data caching, automated error checking and correction, and 24x7 system reliability.

This chapter covers the following topics:

2.1 IBM Power9 processor

The Power9 processor cache hierarchy includes a 32 KB L1 instruction and 64 KB L1 data cache, a 512 KB L2 cache per core, and a 10 MB L3 cache. This cache architecture is equipped with enhanced logic for enhanced cache coherency, and data type awareness for better accessibility of critical data.

There is a 7 TBps interconnect bus that connects the cores to the high-bandwidth DDR memory bus, the PCIe Gen4 bus, CAPI 2.0 bus, NVLink 2 bus, and other processors in the system. The cores of the Power9 processor are connected to this bus by dedicated 256 GBps data lanes, as shown in Figure 2-1. This architecture provides a high capacity, low latency architecture for the largest applications, which includes SAP HANA.

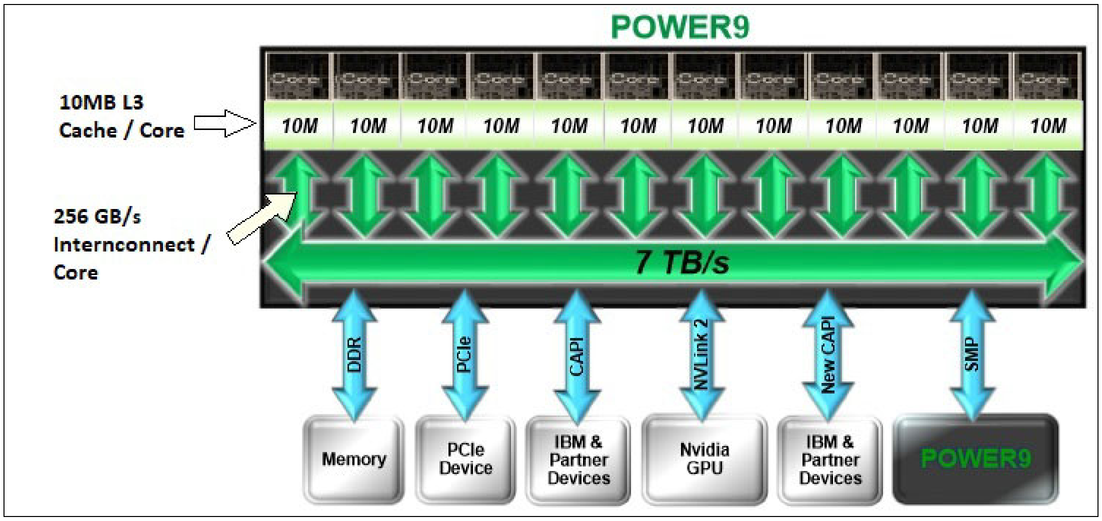


Figure 2-1 IBM Power9 processor bus

The Power9 processor family is fully binary compatible for applications that are optimized for earlier POWER processors, although it provides for new capabilities and optimization for performance for new processor versions.

The IBM Power8 and Power9 processors support dual-Endian architectures. Older applications that were developed with Big Endianness (for example, SAP HANA V1) can run with LPARs that are running applications that are developed with Little Endianness (for example, SAP HANA V2). This dual Endianness allows customers to migrate applications to newer versions when the opportunity presents instead of forcing a complex migration. It is also possible to directly restore SAP HANA V1 and V2 DBs from an Intel processor-based server (Little Endian) to a Power9 processor-based server running SAP HANA V2. Little Endian compatibility between the Intel architecture and POWER architecture can help make it easier to rehost an SAP HANA installation running on an Intel processor-based server to a Power platform.

2.2 Memory architecture

The IBM Power9 processor addresses the processing and memory address needs of platforms such as SAP HANA. The scale-up servers, such as IBM Power System E950 and IBM Power System E980 servers, incorporate eight buffered memory channels, which can provide up to 230 GBps sustained bandwidth, with a maximum capacity of 8 TB per socket. With available DIMM sizes, a system can be configured with a memory capacity of 16 x 128 GB (4 TB) per socket to produce a fully configured Power9 processor-based server that can accommodate up to 64 TB of RAM. SAP supports productive LPARs running SAP HANA up to 32 TB.

The Power9 processor-based memory interface is compatible with Power8 processor-based memory modules and new emerging memory solutions.

2.3 PCIe

Power9 processor-based servers use the PCIe Gen4 bus slot technology, which doubles transfer rates and throughput compared to the PCIe Gen3 design. For large-scale systems, PCIe Gen4 buses provide higher data rates among external data sources like storage area network (SAN and RAM where the data is accessed, for applications such as SAP HANA. Devices include 100 GB Ethernet, InfiniBand, CAPI-enabled adapters, flash memory adapters, SAN-attached flash storage, and others. Also, the increased rates allow for higher throughput and lower latency between PCIe-attached devices.

2.4 Virtual Persistent Memory storage

Power9 processors support DRAM-based persistent storage that is called Virtual Persistent Memory (vPMEM). vPMEM is created from the system DRAM and allocated to a partition. For SAP HANA applications, vPMEM storage can be used to mirror the table data that is stored on disk onto a DRAM-based storage volume that acts like a regular disk device. SAP HANA can address these devices directly, bypassing the need to load data from persistent disk into RAM. vPMEM decreases the start times of the SAP HANA DB and keeps data consistent by synchronizing to disk-based storage for data integrity and persistence.

vPMEM uses system memory and allocates the vPMEM memory on the sockets from which the LPAR core and memory are allocated.

With the Power9 enhanced memory architecture, access to data that is stored in vPMEM devices is provided at DRAM speeds. No special hardware memory devices are needed by SAP HANA DB to use this technology.

2.5 Central electronics complex expansion architecture

The scale-up capability of the Power9 architecture connects four socket-compute enclosures to provide up to 4 enclosures, 16 sockets, and up to 192 cores. These enclosures are interconnected by a 25 GBps optical-style interconnect signal technology, which is four times faster than the IBM Power8 architecture. The interconnect communication design provides at most 2-hop connectivity between processors.

2.6 Virtualization

The IBM POWER architecture is designed for virtualization. The IBM PowerVM Hypervisor is built into the system and is designed for efficient allocation of hardware resources, such as processors, memory, and I/O adapters, to the running LPARs.

2.6.1 Dedicated processor partitions

There are two types of LPARs that can be defined. The first is a *dedicated processor LPAR*, where CPU and memory resources are dedicated to the running LPAR and are not shared with other LPARs running on the system. These LPARs receive CPU resources in increments of full cores and dedicated memory regions. With SAP HANA, SAP supports up to 16 productive dedicated processor LPARs per system.

2.6.2 Shared processor partitions

The other type of LPAR is the *shared processor LPAR* (SPLPAR), where a set number of cores is assigned to a pool, and LPARs are defined to use core resources from this pool. Multiple LPARs are defined to use the core resources from this compute pool. Defining these types of LPARs provides the flexibility to allow LPARs to use, then return, compute resources from the pool.

Among other attributes, an SPLPAR can be defined to use down to 1/100th of a core, which provides fine-grained resource allocation to SPLPARs. A weighting, or priority, value can be assigned to an SPLPAR to allow for high priority applications running in an SPLPAR to get the resources it needs when they are needed.

2.7 Capacity on Demand

The PowerVM environment can dynamically activate idle processor and memory resources when workload peaks occur. This capability, which is called Capacity on Demand (CoD), allows a server to be configured with more processor and memory than is needed for the current workload. These resources are not initially licensed, but put in reserve if needed in the future.

When peak processing demand exceeds the currently licensed resources, this idle capacity can be activated either temporarily or permanently. This CoD allows for a rapid response to unexpected processing demands without the delay of ordering new hardware components and waiting for them to be delivered and installed.

If the resources that are used are only temporary, then they can be removed from the LPAR and returned to the idle state to wait for another peak in demand in the future.

2.8 Live Partition Mobility

Live Partition Mobility (LPM) is a feature of IBM Power that can move active and inactive AIX, Linux, and IBM logical partitions from one frame to another without any downtime. SAP HANA on IBM Power10 systems can enhance the availability, scalability, and flexibility of the SAP applications through LPM. This feature can be used to do the following tasks:

- Migrate SAP HANA systems to new hardware without disrupting production workloads

- Balance SAP HANA workloads across multiple systems
- Perform maintenance on SAP HANA systems without interrupting users

To use LPM for SAP HANA on IBM Power10 systems, the SAP HANA system must meet the following requirements:

- Prerequisites from SAP Note 2055470 are fulfilled.
- The source and destination systems must be IBM Power10 systems.
- The source and destination systems must have the same compatibility mode (POWER10 or default)
- The source and destination systems must have the same amount of memory.
- The SAP HANA system must be running in a logical partition
- The Operating system is not reflecting NUMA topology changes (proc/powerpc/topology_updates is set to off)
- Appropriate licensing for LPM and infrastructure setup is available.

The LPM migration process involves the following steps:

- Analyze workload. Identify partitions or table segments to be moved based on workload analysis.
- Prepare target server. Ensure that the target server meets the hardware and software requirements for SAP HANA.
- Initiate migration. Use SAP HANA administration tools to initiate the migration, selecting the partitions and specifying the target server.
- Data replication. Replicate data from the source server to the target server, ensuring consistency.
- Switch over. Redirect active processes accessing the partitions to the target server seamlessly.
- Verification and monitoring. Monitor performance and verify the integrity of migrated partitions.
- After LPM, performance degradation might occur if SAP HANA is not notified by the operating system about NUMA topology changes. A reboot of the partition updates the NUMA topology.

The following configuration aspects are additional considerations for LPM of SAP HANA partitions:

- Ensure that the VIOS servers that are configured as the Mover/Server Partitions (MSPs) have additional processor and memory resources to drive the LPM traffic at the configured network speeds. Use dual MSPs to provide better performance and reliability, such as a redundant network connection.
- The network configuration is critical to ensuring the minimal impact on performance for LPM:
  - The faster the network connection, the quicker the migration operation finishes. When you migrate larger partitions of 1 TB or more of memory, it is recommended to use 40 Gbps or faster network connections.
  - It is recommended that a separate network adapter and port be used for LPM data transfer between the source MSPs and target MSPs. LPM requires significant amounts of data to be transferred and can affect other traffic on a shared connection.
  - For the LPM network adapters, configure Jumbo frames, Large Send Offload (LSO), and Large Receive Offload (LRO).

- Consider the size of the Physical Page Table (PPT) that is used for LPM to improve SAP HANA application performance when LPM is in progress. For more information, see Trust in your AI journey.

As new features are introduced for LPM, these are documented in IBM Docs and the IBM Power VM Community.

2.9 Additional resources

For more information about technical resources for the IBM POWER architecture, see “Related publications” on page 87.

3

Hardware advantages of IBM Power10 Systems servers

IBM POWER processor-based servers provide many technological capabilities for running SAP HANA on Power Systems. The Power10 processor has capabilities that enhance analytic workloads, high-density cloud virtualization, emerging artificial intelligence (AI), and large in-memory database (DB) solutions. Advantages that the Power10 processor-based servers provide include high memory bandwidth architecture, simultaneous thread concurrency at the core, large processor data caching, automated error checking and correction, and 24x7 system reliability.

This chapter covers the following topics:

3.1 Power10 advantages over Power9

Chapter one of this book described how IBM POWER is a superior platform on which to run SAP workload, and chapter two described the IBM Power9 server advantages. This chapter describes how the new technologies in IBM Power10 servers provide an even better solution for SAP.

This chapter includes details on the newest IBM POWER processor based servers that use IBM Power10 processors. Included in the discussion are details of the hardware of Power10 at the architecture level and the benefits of the IBM Power10 processor over IBM Power9.

3.1.1 Benefits of Power10 processor based servers

Protecting data from core to cloud

IBM Power10 protects data from core to cloud with memory encryption at the processor level and four times more crypto engines in every core compared to Power9.

With data residing in increasingly distributed environments, you cannot define a perimeter for it anymore. This reinforces the need for layered security across your IT stack. The Power10 family of servers introduces a new layer of defense with *transparent memory encryption*. With this feature, all stored data remains encrypted when in transit between the memory storage and processor. Because this capability is enabled at the hardware level, there is no additional management setup or performance impact. Power10 includes four times more crypto engines in every core compared to IBM Power9 servers to accelerate encryption performance across the stack. These innovations, along with new in-core defense for return-oriented programming attacks and support for post quantum encryption and fully homomorphic encryption, helps make one of the most secure server platforms better.

Streamlining insights and automation

The IBM Power10 processor streamlines insights and automation with four Matrix Math Accelerators per core for faster AI inferencing.

As more AI models are deployed in production, the challenges around AI infrastructure are increasing. A typical AI deployment involves sending data from an operational platform to a GPU system. This usually induces latency and can increase security risks by leaving more data in-network. Power10 addresses this challenge with core AI inferencing and machine learning. The Matrix Math Accelerator (MMA) in Power10 cores provide the computational strength to tackle demanding AI inferencing and machine learning at multiple levels of precision and data bandwidth.

Delivering better memory reliability and availability

The IBM Power10 processor uses Active Memory Mirroring to deliver two times better memory reliability and availability than industry-standard DIMMs.

Power S1022s makes the most reliable server platform in its class even better with advanced recovery, diagnostic capabilities, and open memory interface (OMI) attached advanced memory differential DIMMs (DDIMMs). The continuous operations of today’s in-memory systems depend on memory reliability because of their large memory footprint. Power10 DDIMMs deliver two times better memory reliability and availability than industry-standard DIMMs. Power10 can also provide an increase in uptime and improve availability by implementing Active Memory Mirroring.

Improving performance when running SAP HANA

For customers running SAP HANA on POWER, Power10 can provide a significant performance improvement that helps businesses run faster and more efficiently. POWER has improved per-core performance on-time with each successive generation, which includes Power10. The increase in per core performance for Power10 means that fewer cores are needed to run workloads like SAP HANA. This can result in reduced software acquisition costs and reduced recurring maintenance costs for software that is licensed by the core. Reduced costs for acquisition and maintenance can significantly decrease both total cost of acquisition (TCA) and total cost of ownership (TCO) as shown in Figure 3-1.

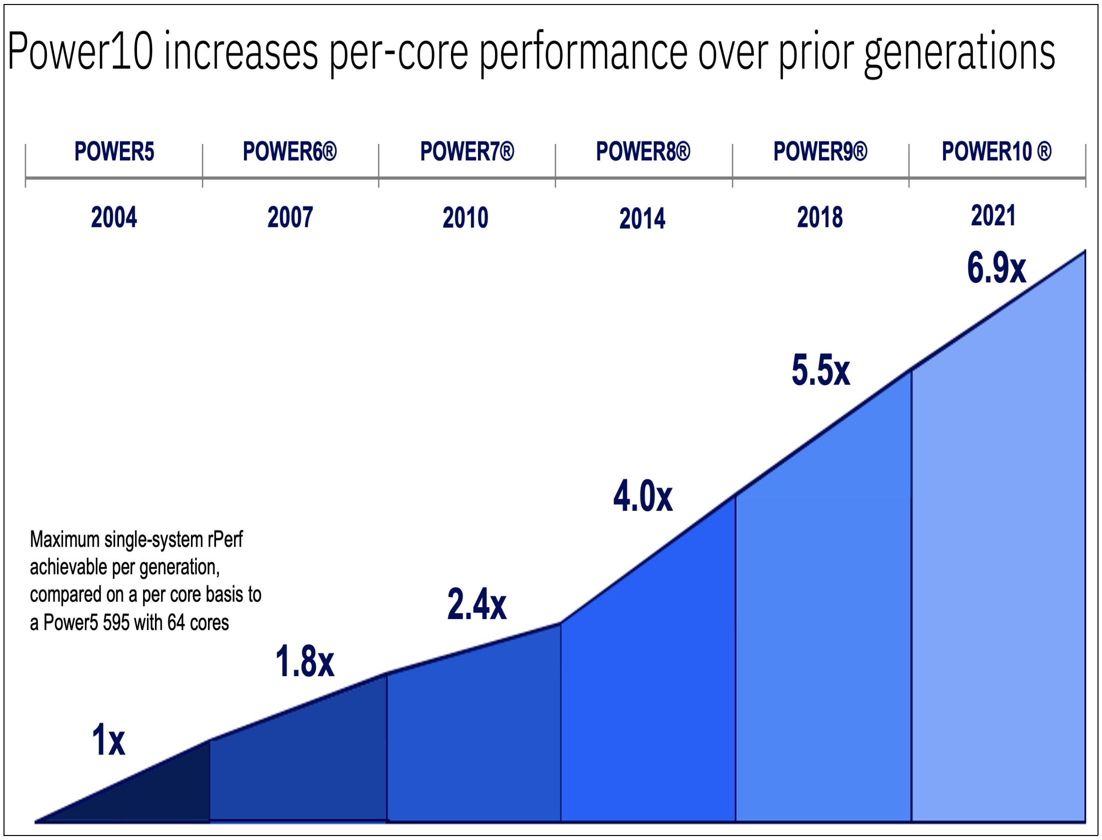


Figure 3-1 IBM Power10 processor performance over prior generations

New memory technologies in Power10 processor based systems

IBM's Power10 introduction included DDIMMs for both Scale-Out (SCO) and Enterprise servers. Also, unlike the Power9, IBM implemented the Dual Chip Module (DCM) to enhance performance by offering more high-performing cores per socket on SCO servers. This advancement enables the consolidation of a greater number of SAP database and application virtual machines on a 2U scale out server.

Memory failure remains a major contributor to system reliability challenges. With the introduction of the Power10 platform, IBM provides innovative solutions such as the Open Memory Interface (OMI) and DDIMMs. These advancements are included to enhance system usage for critical applications and mitigate interruptions by incorporating spare memory modules in DDIMMs to dynamically deallocate memory in response to predictive errors. The new memory architecture of Power10 is engineered to deliver an improvement of twice the availability compared to industry-standard alternatives.

For seamless operation during planned maintenance and to ensure continuous application availability during peak demands, Power Private Cloud with Dynamic Capacity offers remarkable flexibility. By employing a pay-per-use pricing model across a pool of systems, users gain the ability to reallocate resources during support maintenance or access on-demand capacity during peak periods. The OMI memory architecture of Power10 sets it apart from x86 systems that rely on industry-standard DIMMs. IBM's OMI DIMMs demonstrate twice the memory reliability and availability compared to traditional DIMMs used in x86 systems.

This distinction is critical as x86 DIMM failures can lead to system outages. Power10's buffered memory DIMM architecture, found in the scale-out servers, effectively isolates failures to the DIMM level, avoiding system-wide outages. This focus on memory reliability becomes increasingly crucial for large in-memory configurations like SAP HANA and Oracle, as it significantly impacts overall system reliability. This level of resilience demonstrates a commitment to maximizing system reliability.

Virtual Persistent Memory

Virtual Persistent Memory (vPMEM) is an enhancement of IBM's advanced virtualization platform (PowerVM) introduced with the Power9 family and continued in the Power10 family.

With vPMEM, you can configure persistent volumes by using the conventional DRAM memory modules available in every IBM Power9 Systems. Therefore, no special or additional hardware components are required.

Because vPMEM is built on DRAM technology, it has the same performance characteristics as DRAM. The vPMEM enhancement on IBM Power9 enables faster restarts of SAP HANA during planned maintenance and during unplanned outages, without compromising performance in production. The vPMEM volumes are managed on the system Hardware Management Console (HMC). They are defined per LPAR and are not directly sharable or transferable to other LPARs.

The SAP HANA database is designed to identify the presence of Persistent Memory and store its Main table fragment of columnar data in vPMEM and store Delta fragment, Row Store in DRAM. The main data fragment is stored in vPMEM and represents 95% of the database data. Because the data is stored in vPMEM, it takes less time to load data when SAP HANA is restarted or the server is rebooted. Therefore, SAP HANA does not have to wait for data to be loaded into memory because most of the data is already in the memory. This enables rapid restart and recovery times with full performance benefits of SAP HANA. The use of vPMEM includes the following benefits:

- Faster Restart

  With the improved reliability of IBM Power, hardware issues are seldom the reason for SAP HANA restarts. Instead, the use of persistent memory can decrease the restart time of the database. The faster restart can improve your availability by reducing the amount of time required to do planned maintenance such as patches, upgrades. The faster restart can also reduce the time to recover from unplanned outages.

- Faster Shutdown

  This is also an important aspect to help manage your maintenance window. You can bring down the environments more quickly to perform the maintenance work and then restart the environments quickly to reduce the business impact.

Power10 sustainability

IBM Power10 systems are positioned as a fully eco-friendly solution. The focus is on delivering highly efficient, robust, and sustainable products that cater to the needs of IBM clients. Table 3-1 shows a comparison of expected power usage between Power10 servers and Power9 and Power8 servers. Calculations are based on rPerf/Watt ratio.

Table 3-1 Improvements in energy use Power10 systems

Power10 Server	Power9 Server	Power8 Server
Power E1050	Power E950	Power E850
	47% less energy consumption	59% less energy consumption
Power S1024	Power S924	Power S824
	56% less energy consumption	64% less energy consumption
Power S1022	Power S922	Power S822
	41% less energy consumption	56% less energy consumption
PowerE1014	Power S914	Power S814
	52% less energy consumption	59% less energy consumption

IBM POWER remains committed to assisting clients in repurposing various components from previous Power Systems, including I/O adapters and internal storage. This approach helps reduce the environmental impact by minimizing the production of new components and disposing of non-functional ones. Furthermore, the focus on building highly resilient systems contributes to enhanced sustainability.

3.2 IBM Power10 processor

The Power10 processor is a superscalar symmetric multiprocessor (SMP) designed for use in servers and large-cluster systems. It uses complementary metal–oxide–semiconductor (CMOS) 7 nm technology with 18 metal layers. Each processor die building block contains a maximum of 30 high-performance, 4-threaded general-purpose cores in the 4-way simultaneous multithreading variant. Each processor has 1 MB of low-latency, high-bandwidth alert Level 2 cache or fifteen 8-threaded general-purpose cores in the SMT8 variant, each with 2 MB of L2 cache. The cores share a maximum of 120 MB of latency optimized Non-uniform cache access (NUCA) L3 cache. Each processor die building block contains the following features:

- A maximum of 16 × 8 lanes of open memory interface (OMI) memory at a maximum of 32 gigatransfers per second (GTps)
- A maximum of 14 × 9 lanes of SMP fabric interconnect at a maximum of 32 GTps
- A maximum of 12 × 8 lanes of OpenCAPI attach at a maximum of 32 GTps
- A maximum of 32 lanes of peripheral component interconnect express (PCIe) Gen 5 attach at a maximum of 32 GTps.

Figure 3-2 on page 62 provides a block diagram of the Power10 processor.

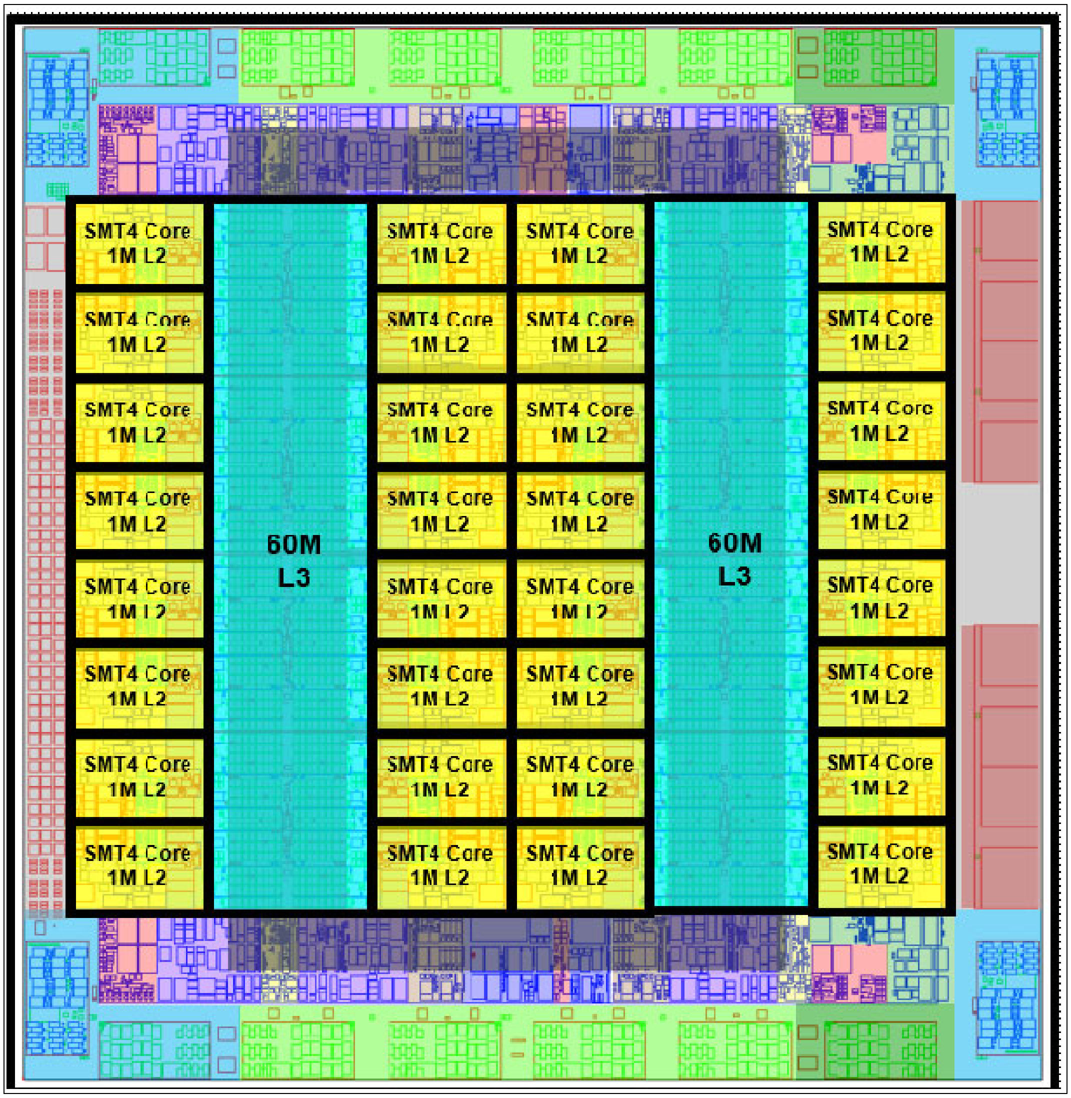


Figure 3-2   IBM Power10 processor

Because it is targeted for a broad range of applications, flexibility and configurability are key themes that govern the Power10 design point.

3.2.1 Interface types

A single, highly-optimized, 32 GTps differential signaling technology can service most functional interfaces connecting the processor to other system elements. The only other functional signaling technology is the industry-standard PCIe Gen5 technology that is integrated into the Power10 processor. The 32 GTps signaling technology is deployed as two interface types, both of which support a myriad of protocols and configurations. Figure 3-3 on page 63 illustrates the Power10 signaling technology.

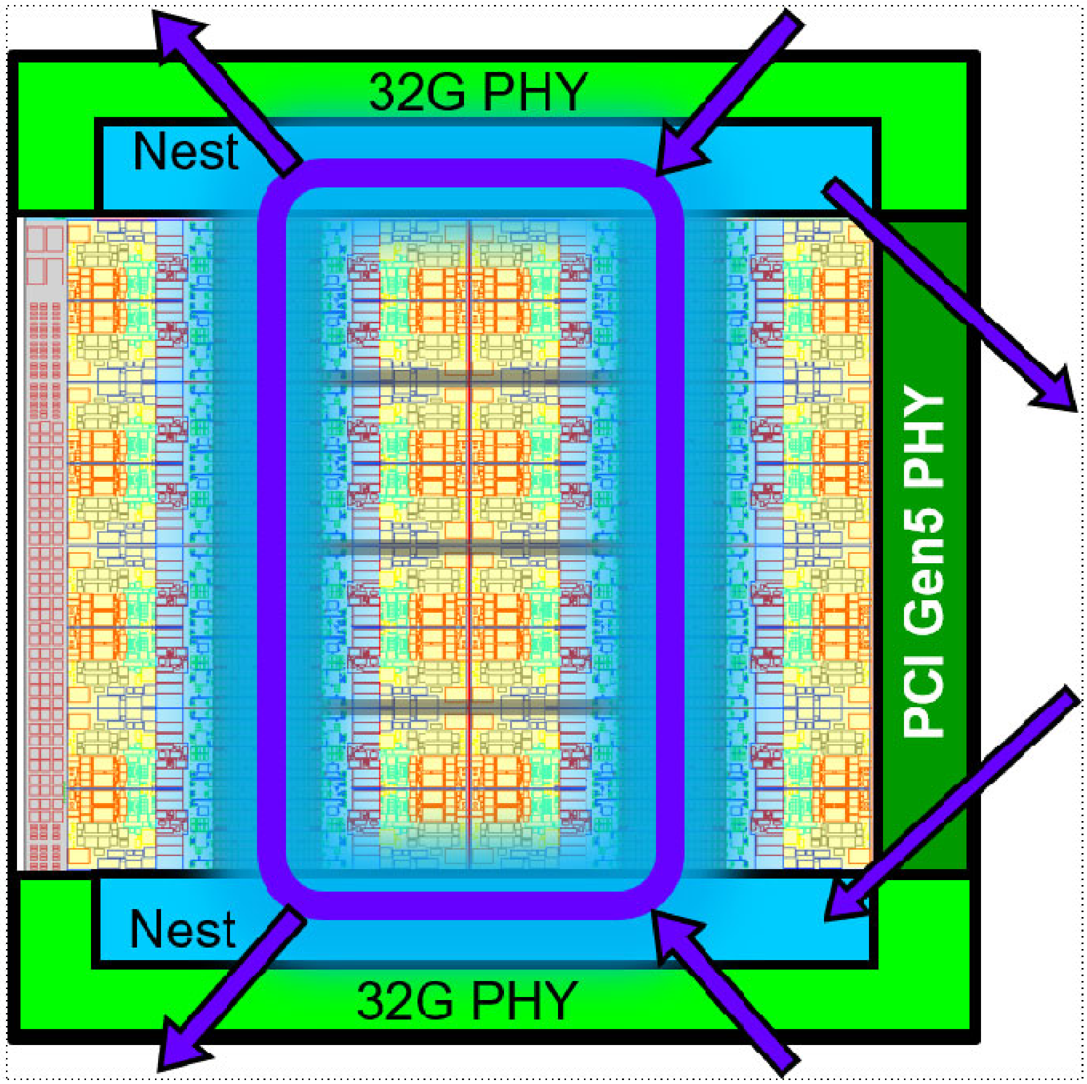


Figure 3-3 Power10 Signaling Technology

OMI

The first type of interface, the OMI, provides a single physical interface that enables low-latency, high-bandwidth, technology-independent host memory semantics. The interface can attach both established and emerging memory elements to the processor, which includes the following memory types:

- Main tier, low-latency, commodity, and enterprise-grade double data-rate memory interface, 4th generation (DDR4) and DDR5
- Near-tier. high-bandwidth Graphics double data-rate (GDDR) and High-bandwidth memory (HBM)
- Storage-tier, high capacity, persistent phase-change, and flash-derivative technologies

Figure 3-4 illustrates the OMI.

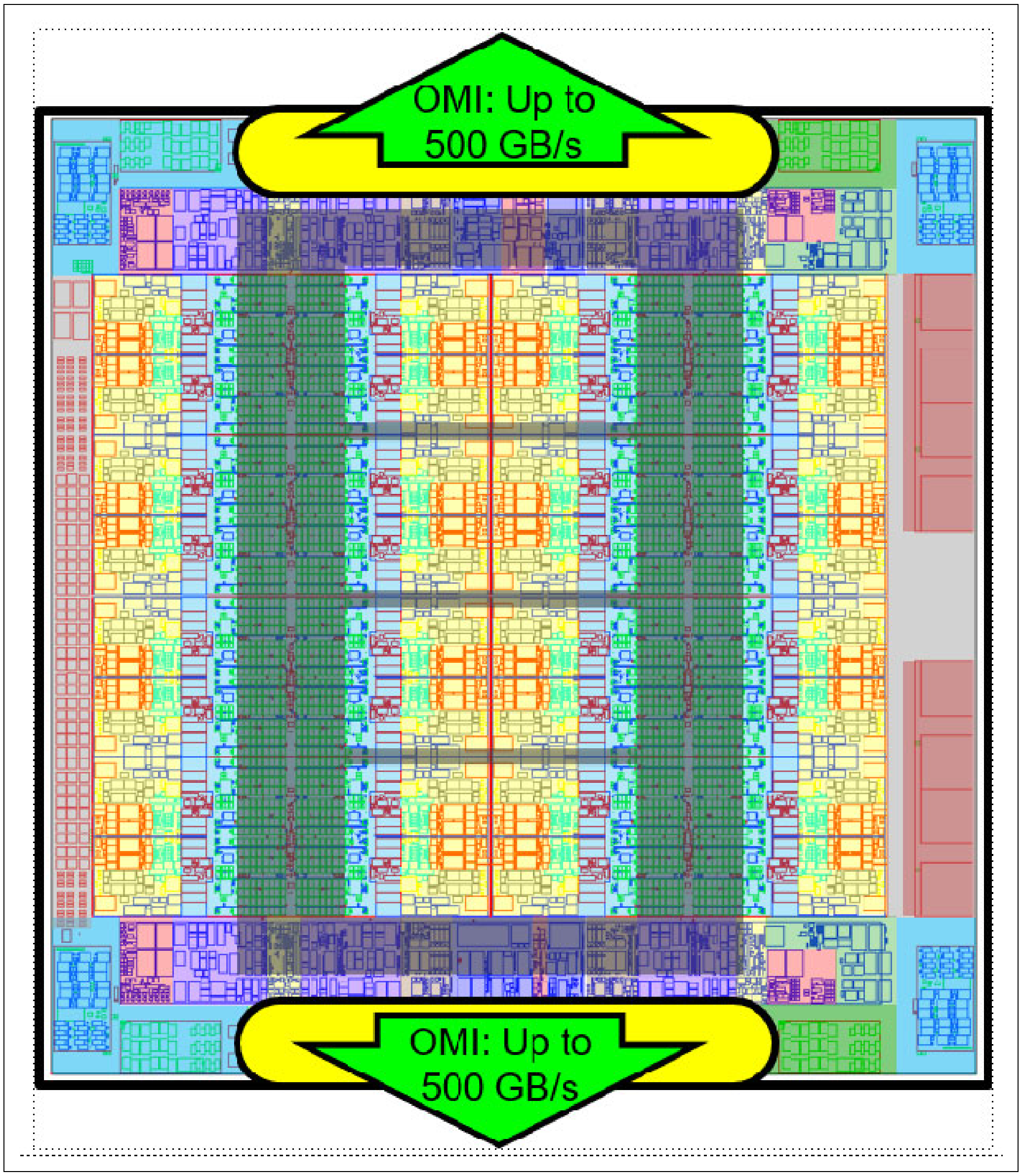


Figure 3-4 Power10 OMI

PowerAXON interface

The second type of interface, the PowerAXON (Figure 3-5 on page 65), provides a single physical interface that can be configured in several ways:

- As a first-tier or second-tier, glueless, SMP-link interface so that a maximum of 16 Power10 processors can be combined into a large, robustly scalable, single-system image
- As an OpenCAPI interface to attach the cache-coherent and I/O-coherent computational accelerators, load and store addressable host memory devices, low-latency network controllers, and intelligent storage controllers into a Power10-based system
- As a host-to-host integrated memory clustering interconnect, which enables multiple Power10 systems to directly use memory throughout the cluster.

i

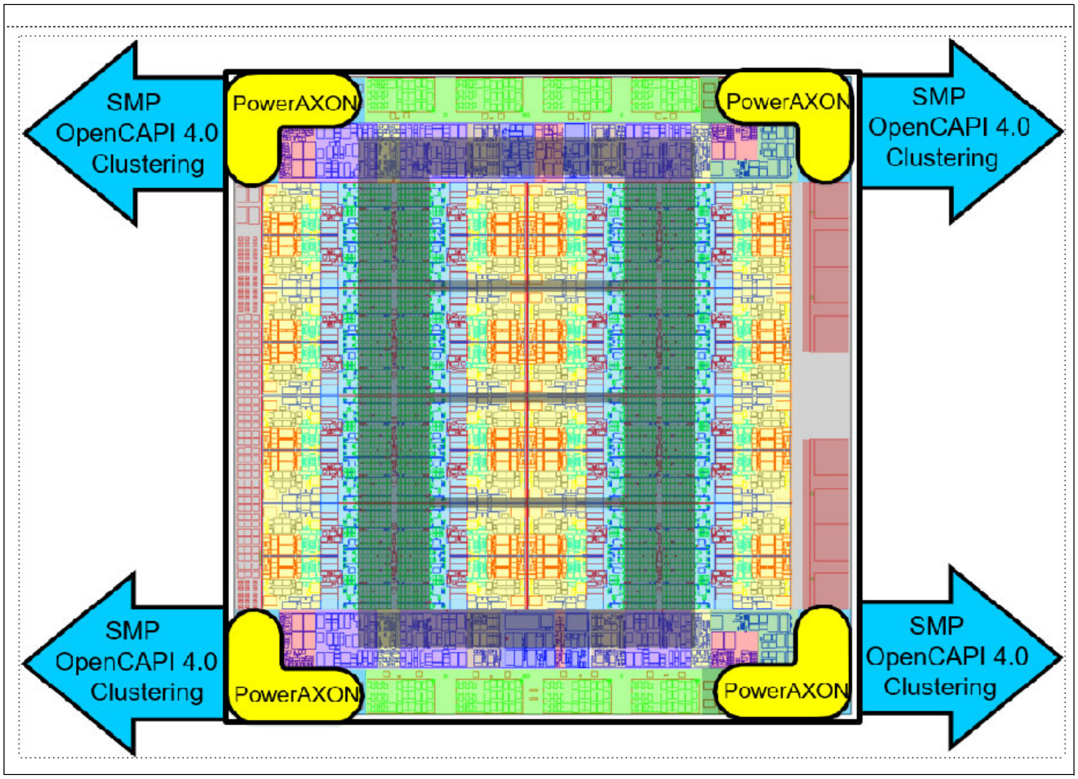


Figure 3-5 Power AXON Interface

3.2.2 Single-chip and dual-chip modules

The Power10 processor chip has been designed and optimized as a building block that can be used to cover a wide variety of single-chip module (single die) and dual-chip module (dual die) per socket configurations. The high maximum frequency at maximum voltage and robust core and cache micro architecture are foundational to single-chip module (SCM) scenarios. The strong focus on energy efficiency and SMP link-stranding technology are foundational to dual-chip module (DCM) scenarios. SCM options are best suited for cost, core and thread strength, and bandwidth-per-computation optimization. The DCM options favor balanced thread and socket performance and extreme socket performance optimization. Figure 3-6 on page 66 describes the Power10 building blocks.

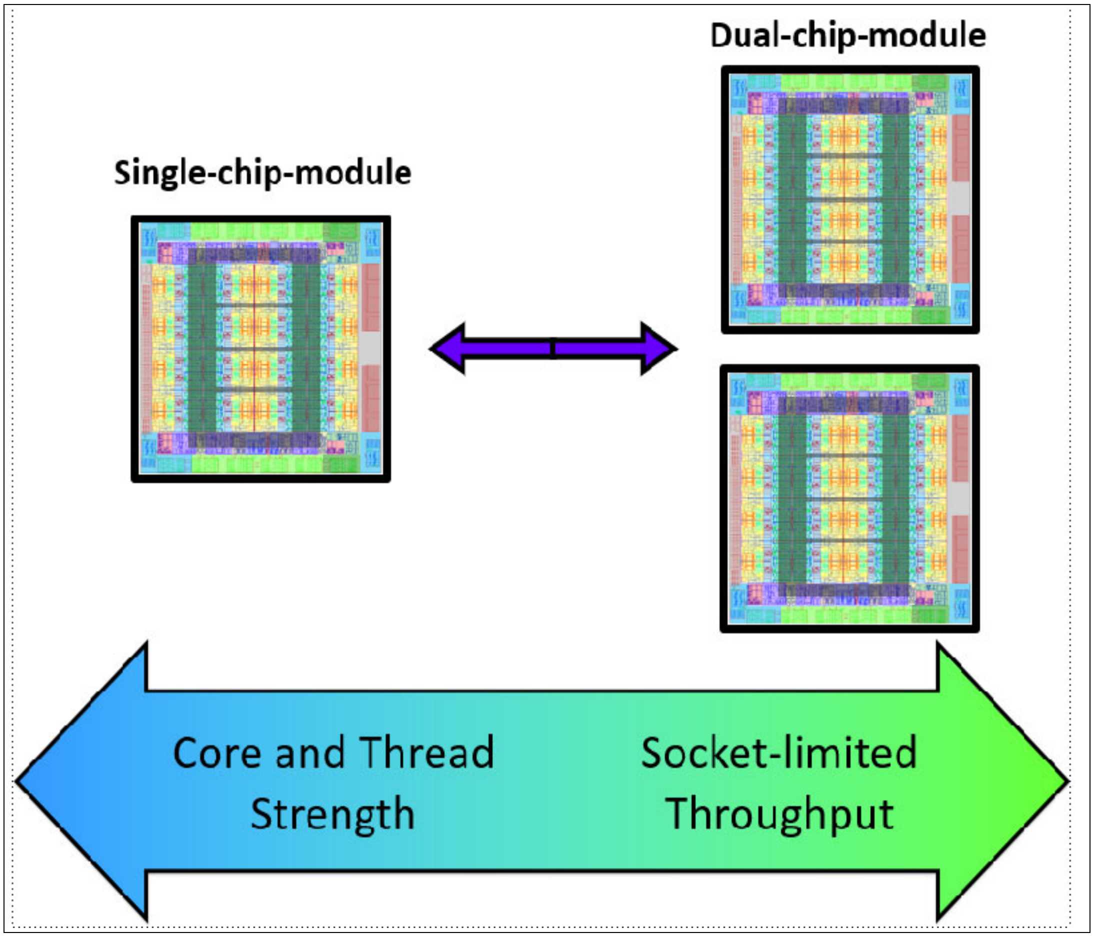


Figure 3-6 Power10_BuildingBlocks

3.2.3 Processor core

Similar to the Power9 core, the Power10 processor core elements consist of modular building blocks, which enable two variants of the processor core:

1. A smaller, 4-threaded version, which is targeted at the bare-metal Linux market and some segments of the KVM Linux market
2. A larger, 8-threaded version, which is targeted at the PowerVM market and some segments of the KVM Linux market

Power10 targeted environments

Unless otherwise specified, references to the Power10 processor core in this document describes the SMT4 variant. The Power10 processor offers superior performance, reliability, energy efficiency, and cost for a broad range of on-premises and cloud-focused offerings. The targeted application segments include the following environments:

- AIX and IBM i environments that are built upon PowerVM. Systems ranging from 1–16 processor chips, if you use the SMT8 variant of the processor core. This includes the established and emerging workloads that demand mission-critical reliability, availability, and serviceability (RAS), security, and scale.
- Enterprise Linux environments that run on systems ranging from 1–16 processor chips, if you use the SMT8 variant of the processor core, and targeting mission-critical, large-scale

Linux workloads such as in-memory databases, which demand high-memory capacity, robust scaling, and high reliability.

- Cognitive and high-performance computing environments that run on systems ranging from 1–4 sockets and typically taking advantage of the SMT4 variant of the processor core. The processor core provides superior single-instruction, multiple-data (SIMD) performance across a wide variety of data types. The targeted workloads include both emerging AI training and inferencing workloads and established and emerging HPC workloads. The chip also supports high-bandwidth, off-chip accelerator attach capabilities, a robust PCIe I/O subsystem, and high-bandwidth OMI memory to address these segments.
- Hyper scale data center deployments and OpenPOWER partner-driven deployments. Systems focus on one-socket and two-socket form factors using the SMT4 variant of the processor core. The Power10 processor is engineered with a strong focus on the energy efficiency and socket performance that is combined with differentiated acceleration, memory, and I/O capabilities.

Architectural focus areas

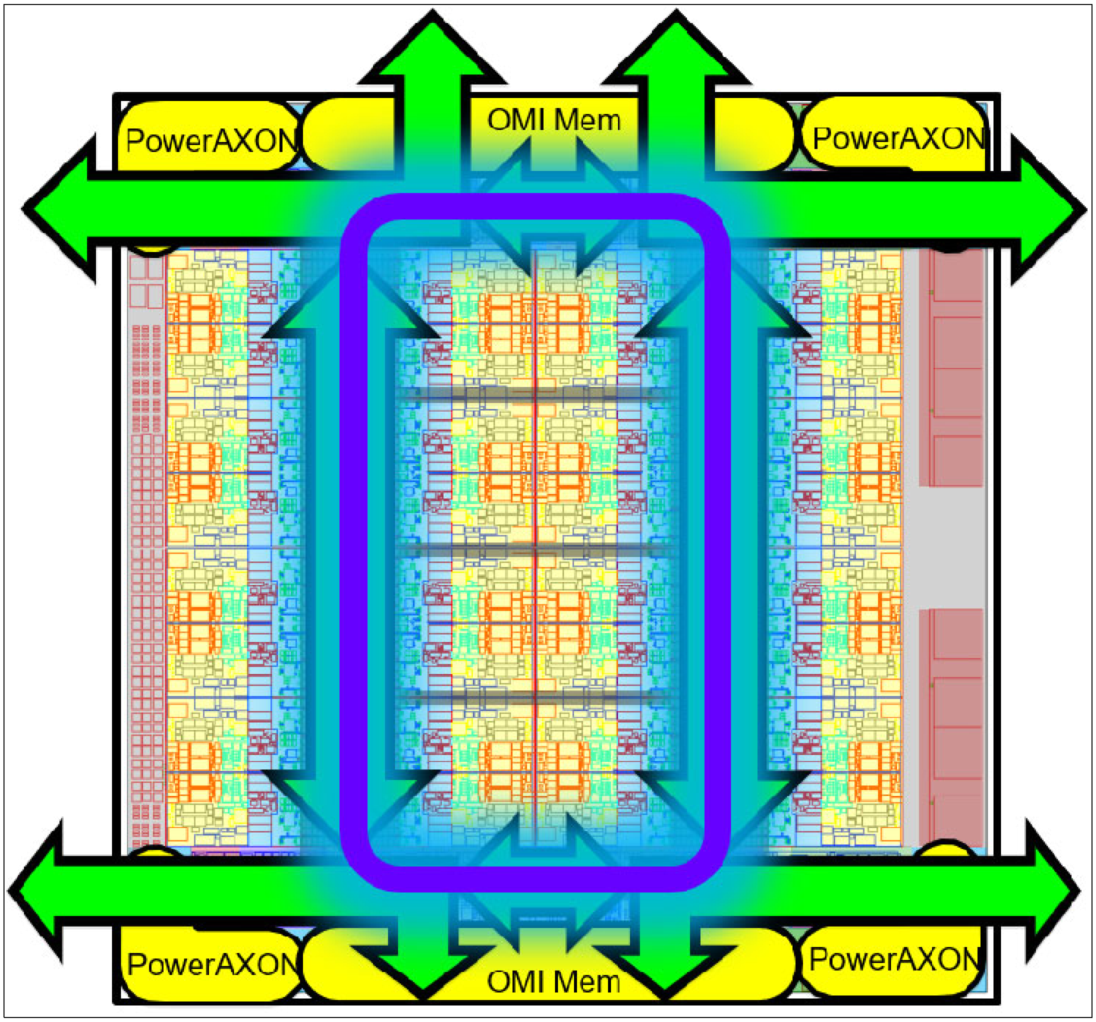
Given the previously stated environment objectives, five major architectural focus areas characterize the Power10 design point.

System data bandwidth

The Power10 processor has intrinsic, general-purpose, computational capabilities. The Power10 chip can also be viewed as a high-bandwidth switch (Figure 3-7 on page 69). The Power10 uses its highly-optimized, 32 GTps signaling technology to connect PowerAXON-attached computational elements, such as compute accelerators or other Power10 processors, to OMI-attached memory devices. The connections can be made in a latency-optimized manner at unprecedented bandwidth and scale. The entire processor die layout is organized to enable a racetrack on-chip coherence and data ring topology, with PowerAXON, OMI, and PCIe-protocol infrastructure located adjacent to 32 GTps and PCIe signaling technology on the chip edge. The optimized PowerAXON-to-racetrack-to-OMI connectivity supports a maximum of 1 TBps end-to-end switching bandwidth in-and-out of the processor. The optimized connectivity makes the Power10 chip an ideal hub for constructing heterogeneous, specialized data and compute architectures.

Cognitive-infused general-purpose compute architecture

Despite optimizations for connecting to off-chip computation-acceleration, the Power10 chip is strongly focused on providing industry-leading SIMD run capabilities in its general-purpose, Power instruction set architecture (ISA), processor cores. The ISA is enhanced to support the reduced-precision data types and matrix-math operations that are essential to high-performance inferencing and learning kernels. However, the ISA enhancement also focuses on higher-precision data types essential to existing analytic and high-performance-computing applications. In addition to this, increased investments in robust execution prediction, address translation caching, and instruction/data caching provide strong value for workloads. The range of workloads includes emerging scripting frameworks, data analytics, transaction processing, and data-intensive enterprise applications and provides robust virtualization for consolidated enterprise environments and container density for multi-tenancy cloud environments.



Energy-efficiency

Every aspect of the Power10 processor is engineered with a strong focus on energy efficiency. This substantial improvement, compared to prior POWER processors, combined with the modular building block approach, which enables optimized dual-chip module sockets for one-socket, two-socket, and four-socket, high-density data center and cloud deployments. These optimizations also bring value to higher-power sockets that are found in enterprise applications by enabling a larger number of high clock-speed, thread-strength optimized cores per socket, and eliminate the requirement to modulate clock speed based on workload intensity.

Thread-strength

There is an inherent tradeoff that is involved in biasing a single offering either toward thread-strength or socket-throughput. However, the Power10 chip's modular SCM and DCM building block approach enables the DCM to focus on throughput, freeing silicon area for strong micro-architecture investments in thread strength. Instruction-per-cycle (IPC) improvements are provided by the following features, which are enabled by providing more silicon area per core:

- ► More robust SIMD architecture
- ► Execution prediction
- ► Translation, instruction, and data caching
- ► Internal bandwidth
- ► Execution resource provisioning

Higher clock speeds are possible because of the strong focus on energy efficiency and the advances in the 7 nm semiconductor technology that is used by the Power10 chip. The combination of silicon area, IPC improvements, and higher clock speeds result in substantial thread strength improvement compared to prior POWER processors.

End-to-end security

The Power10 processor contains numerous hardware elements, which when combined with corresponding firmware and system software elements, form a high-performance set of end-to-end security solutions. The security solutions include the following elements:

- ► Pro-active, side-channel attack detection and avoidance mechanisms
- ► Hardware-enforced, host-processor secure-memory isolation for secure container enable,
- ► Secure memory isolation for secure-attached accelerator support
- ► Two forms of full main memory encryption that are targeted at both volatile and non-volatile memory technologies
- ► Cryptographic accelerators
- ► Application-enabled protections
- ► Secure-trusted-boot enable

Figure 3-7 Power10 Bandwidth

These elements enable Power10-based systems to provide mission-critical security not only for on-premises and private cloud environments but also for public and hybrid cloud environments.

3.3 IBM Power10 Packages

This section describes the different configurations that are used in the IBM Power10 server family.

3.3.1 Power10 Single-Chip Module (68.5 x 77.5 mm)

The Power10 Single-Chip Module (SCM) has the following general features:

- ► Body size: 68.5 mm × 77.5 mm
- ► Interconnect technology. Hybrid Land grid array socket
- ► 1.5 mm interstitial LGA pitch with minimum pitch of 1.06 mm and 4445 pins
- ► 8-2-8 organic package construction

Bus Features

Power10 includes the following bus features:

- 72 A-bus/X-bus/OpenCAPI lanes to the Nearstack connectors on top of the module:
  - Each Optical physical layer 1 (OP1), OP2 runs as a 2 × 9 SMP bus at 32 Gbps
  - Each OP4, OP6 runs as one of the following two modes:
    - 2 × 9 SMP at 32 Gbps
    - 2 × 8 OpenCAPI at 32 Gbps
- 72 A-bus/X-bus/OpenCAPI lanes to the bottom of the module:
  - Each OP0, OP3, OP5, OP7 runs as one of the following two modes:
    - 2 × 9 SMP at 32 Gbps
    - 2 × 8 OpenCAPI at 32 Gbps
- 128 OpenCAPI memory interface (OMI) lanes to the bottom of the module:
  - Each OMI port is composed of 2 × 8 lanes
  - Total 16 × 8 lanes for OMI[0:7]
  - Runs at 32 Gbps
- 32 PCIe lanes to the bottom of the module:
  - Each E0, E1 runs as one of the following modes:
    - 1 × 16 Gen4 at 16 Gbps
    - 2 × 8 Gen4 at 16 Gbps
    - 1 × 8 Gen4, 2 × 4 Gen4 at 16 Gbps
    - 1 × 8 Gen5 at 32 Gbps, 1 × 8 Gen4 at 16 Gbps
    - 1 × 8 Gen5 at 32 Gbps, 2 × 4 Gen4 at 16 Gbps

Figure 3-8 shows a logical diagram of the Power10 SCM.

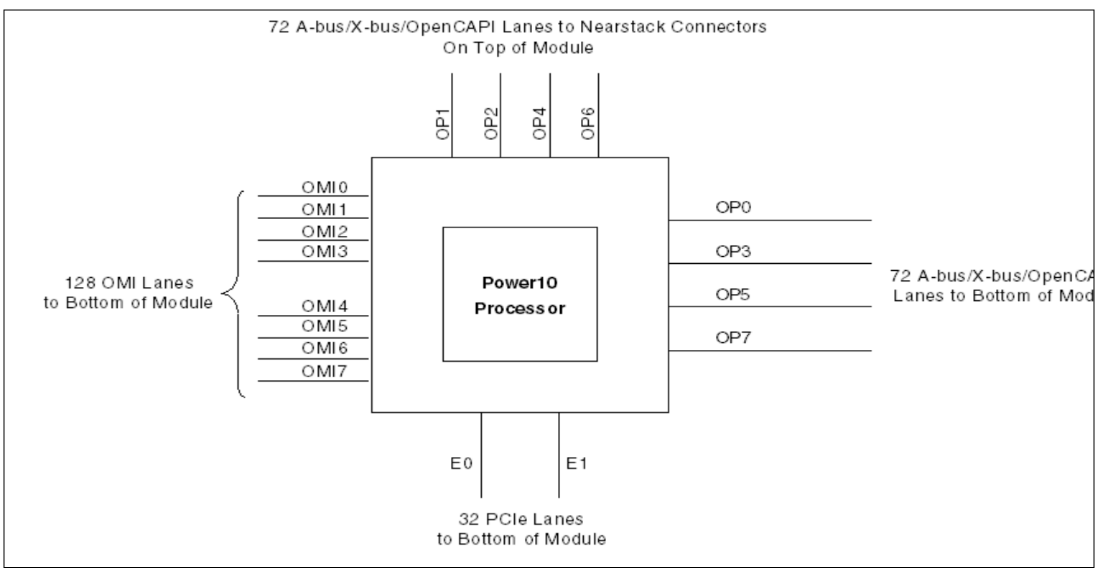


Figure 3-8 Power10 Single-Chip Module (Logical Diagram)

Figure 3-9 on page 71 shows a physical diagram of the Power10 SCM.

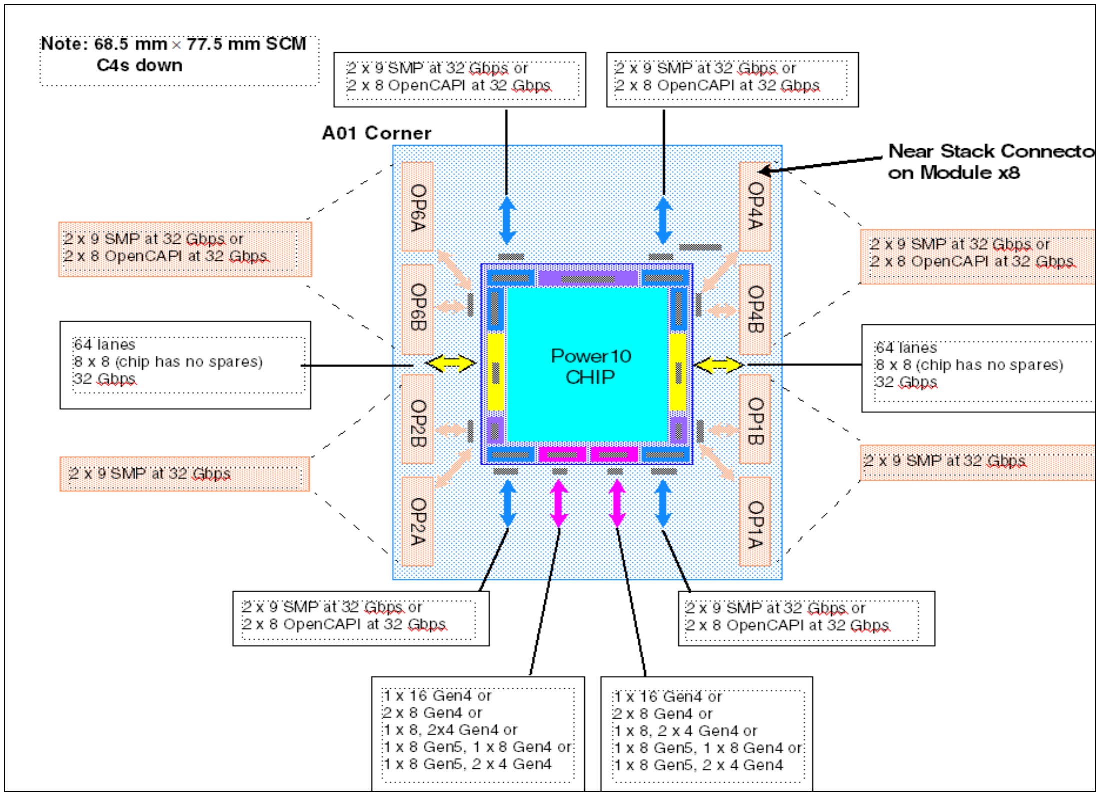


Figure 3-9 Power10 Single-Chip Module (Physical Diagram)

3.3.2 Power10 Dual-Chip Module (74.5 x 85.75 mm)

The Power10 Dual-chip module (DCM) has the following general features:

- Body size: 74.5 mm × 85.75 mm
- Interconnect technology: Hybrid LGA socket
- 1.5 mm interstitial LGA pitch with a minimum pitch of 1.06 mm and 5387 pins
- 8-2-8 organic package construction

Bus Features

The Power10 DCM has the following bus features:

- 36 X-bus lanes for chip-to-chip connection:
  - 2 × 9 OP2 of chip 0 connects to 2 × 9 OP1 of chip 1
  - 2 × 9 OP6 of chip 0 connects to 2 × 9 OP4 of chip 1
  - Run as an SMP bus at 32 Gbps
- 216 A-bus/X-bus/OpenCAPI lanes to the bottom of the module:
  - OP0, OP1, OP3, OP4, OP5, OP7 from chip 0
  - OP0, OP2, OP3, OP5, OP6, OP7 from chip 1
  - Each OP1, OP2 runs as a 2 × 9 SMP bus at 32 Gbps
  - Each OP0, OP3, OP4, OP5, OP6, OP7 runs one of the following two modes:
    - 2 × 9 SMP at 32 Gbps
    - 2 × 8 OpenCAPI at 32 Gbps
- 128 OMI lanes to the bottom of the module:
  - OMI [0:3] from chip 0
  - OMI [4:7] from chip 1

– Each OMI port is composed of 2 × 8 lanes
– Total 16 × 8 lanes for OMI[0:7]
– Runs at 32 Gbps

- 64 PCIe lanes to the bottom of the module
  – E0, E1 from chip 0
  – E0, E1 from chip 1
  – Each E0, E1 can run as one of the following modes:
    • 1 × 16 Gen4 at 16 Gbps
    • 2 × 8 Gen4 at 16 Gbps
    • 1 × 8 Gen4, 2 × 4 Gen4 at 16 Gbps
    • 1 × 8 Gen5 at 32 Gbps, 1 × 8 Gen4 at 16 Gbps
    • 1 × 8 Gen5 at 32 Gbps, 2 × 4 Gen4 at 16 Gbps

Figure 3-10 shows a logical diagram of the Power10 DCM.

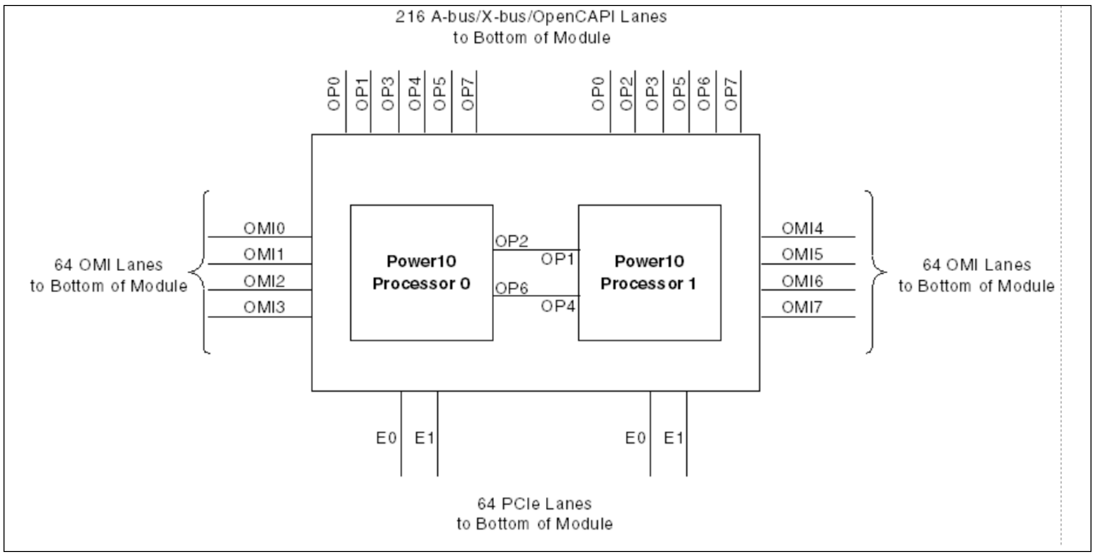


Figure 3-10 Power10 Dual-Chip Module (Logical Diagram)

Figure 3-11 on page 73 shows a physical diagram of the Power10 DCM.

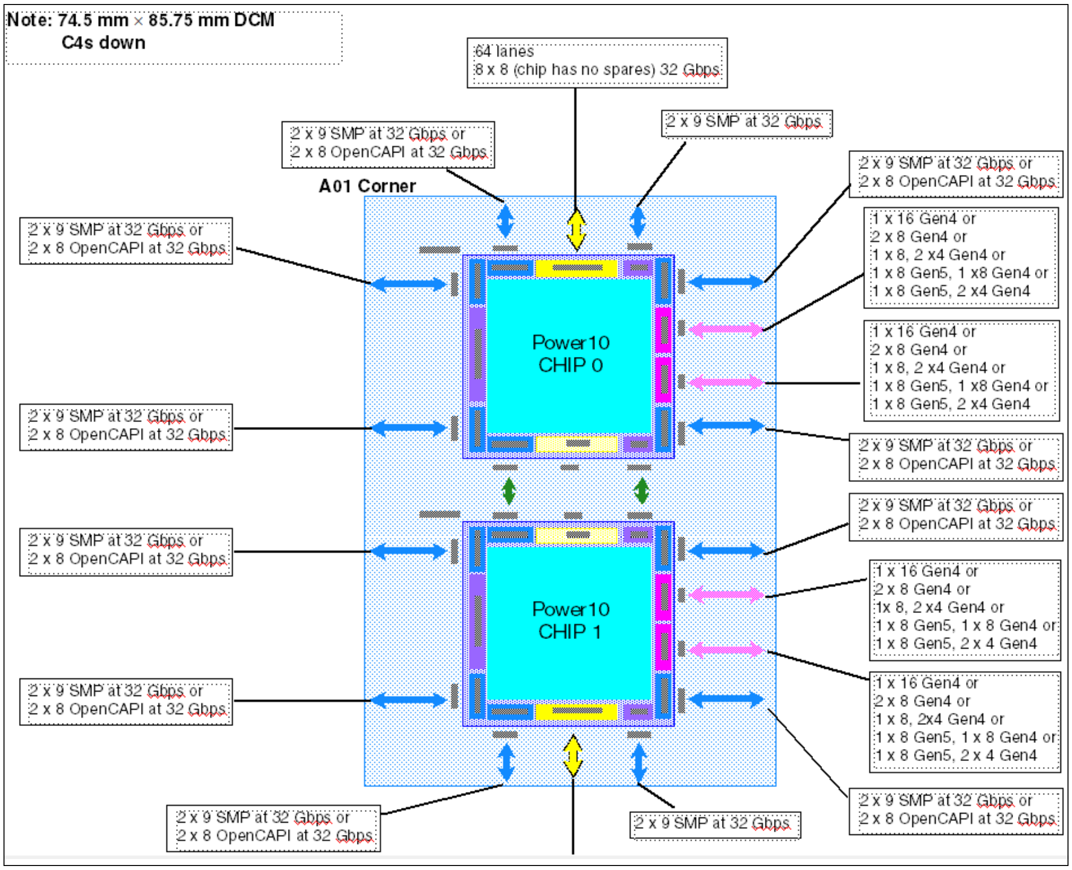


Figure 3-11 Power10 Dual-Chip Module (Physical Diagram)

3.4 Memory architecture

The Power10 memory controller unit (MCU) provides the system memory interface between the on-chip SMP interconnect fabric and the OpenCAPI memory interface (OMI) links. These OMI links can be attached to external memory buffer chips that conform to the OpenCAPI 3.1 specification, which refers to Memory Interface Class only. The buffer chips directly connect to industry-standard memory DIMM interfaces or other memory media, such as storage-class memory. The MCU acts as a subordinate unit only. It is not a source of commands to the SMP fabric. There are logically eight, essentially-independent MCU channels on the chip that interface to a total of 16 high-speed OMI links. Each memory channel supports two OMI links, which are also called subchannels. Physically, the MCUs are grouped into four instances of an extended memory OMI (EMO) unit. Each EMO unit contains two MCU channels. The MCUs process the following operations:

- 128-byte read requests and 64-byte or 128-byte write requests from processor cores, caches, and I/O host bridges
- Partial-line writes of 1–128 bytes
- Atomic memory operations (AMOs)

The MCU also handles address-only operations for purpose of address protection and acts as the lowest-point of coherency (LPC).

The eight MCUs on the chip can be configured into one or more address interleave groups. Within each group, the address space is divided into portions, such that each sequential portion is handled by a different MCU in a round-robin fashion. The maximum memory addressing per Power10 chip is 128 TB.

Within a single MCU channel, the two OMI subchannels are always address-interleaved on a 128-byte boundary, assuming both subchannels are populated with memory.

The MCU resides in a single clock domain. That is, the memory controller asynchronous (MCA) domain runs asynchronously to the SMP interconnect fabric, but synchronously to the OMI interface.

3.4.1 Basic Configuration and Grouping

Each MCU channel includes a set of Base Address Registers (BARs), which are programmed by boot firmware at Initial program load (IPL) time. How they are configured depends on which OMI ports are populated with memory and the sizes of the memory behind each port. The following BAR registers are supported:

- MCFGP: Primary memory BAR
- MCFGPR-CFG: Configuration memory space BAR
- MCFGPR-MMIO: MMIO memory space BAR

The MCU architecture allows for 1, 2, 3, 4, 6, or 8 MCUs to be grouped together for address interleaving. As a group, the MCU channels then hash a contiguous address space among themselves to more efficiently distribute the memory workload. Figure 3-12 shows the memory setup.

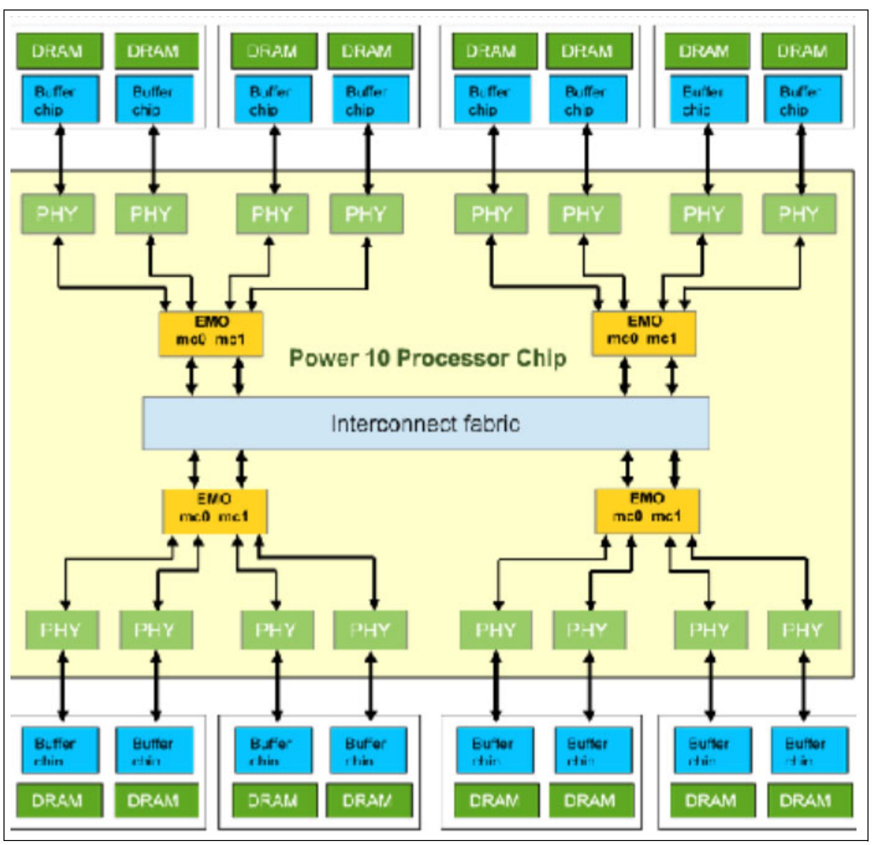


Figure 3-12 Power10 System Memory High-Level Diagram

For 2, 3, 4, 6, or 8 MCUs to be grouped, the total amount of memory that is defined by each MCU's primary memory configuration facilities must be the same. The DIMM configurations and sizes that make up the total amount of memory can be different, but the total memory size plugged behind each MCU/buffer chip in the group must be the same.

The total amount of physical memory behind an MCU can be less than the memory size specified by that MCU's primary memory configuration register. However, if the MCU decodes an address that falls within its programmed address range but does not decode to a valid physical dynamic random access memory (DRAM) address, an address error combined response is generated.

Each MCU receives a 56-bit address for each snoop operation from the SMP fabric, and forwards a 43-bit (8 TB) logical address to the buffer chip.

3.4.2 Atomic Memory Operations

Each MCU supports the processing of a set of atomic memory operations (AMOs). AMOs are read-modify-write (RMW) type operations and have store data that is associated with them. The MCU implements an Arithmetic logic unit (ALU), which operates on the store and memory fetch data. The intent is to provide high throughput of multiple AMOs that target the same address. There are two major features that enable this high level of throughput:

1. Multiple AMOs that target the same address can be queued in the command list queue such that they are not serialized because of retries on the SMP fabric interface. Each AMO command must be performed atomically, and the AMO commands are run in the order in which there are snooped on the SMP fabric.
2. The 64-entry × 128-byte RMW buffer is managed as a cache. When one AMO finishes, its data is maintained in the RMW buffer so that a subsequent AMO to the same address receives its data directly from the buffer instead re-fetching the data from memory. This caching capability is also used for partial writes, which provides for the gathering of multiple partial writes before writing back to memory, and for MDI bit updates. Each command list entry that represents data in the RMW buffer has three cache states: idle, valid, and clean. The command-list logic controls the deallocation of RMW buffer entries in response to snoop traffic to clear data in the RMW buffer.

AMOs are only supported in 4-byte and 8-byte sizes. Both big-endian and little-endian modes are supported. Table 3-2 lists the 23 supported AMO types.

Table 3-2 Supported AMO types

Store Add Store XOR Store OR Store AND Fetch and ADD Fetch and XOR Fetch and OR Fetch and AND Compare and swap equal Compare and swap not equal Swap unconditional Fetch and increment bounded	Fetch and increment equal Fetch and decrement bounded Store twin Store maximum unsigned Store maximum signed Store minimum unsigned Store minimum signed Fetch and maximum unsigned Fetch and maximum signed Fetch and minimum unsigned Fetch and minimum signed

All of the Fetch operations return the unmodified memory data to the SMP fabric, except for the following operations:

- Fetch and increment bounded
- Fetch and increment equal
- Fetch and decrement bounded

These three operations return either the unmodified memory data or the minimum unsigned-integer value based on the result of a compare of two adjacent granules of memory data.

From an MCU data flow and sequencing perspective, AMO operations are always 64 bytes, and partial writes and MDI updates are always 128 bytes.

3.4.3 Selective Memory Mirroring

To improve memory RAS for large systems, the memory controller can be configured for symmetric memory mirroring (SMM). In this configuration, memory subchannels are grouped into mirrored pairs. Selective memory mirroring means that separate mirrored and non-mirrored BAR registers enable memory access to be targeted to either mirrored or non-mirrored space. Although mirroring in all MCFG configurations is supported, it is only supported if both subchannels for all MCUs within an interleave group are populated with the same memory of equal capacity.

Memory mirroring takes effect when an incoming address lands in the mirrored address region as defined by the mirrored BAR and mirrored-size register MCFGPM.

Write operations are issued to both subchannels of a mirrored pair. Read operations to a mirrored address space attempt to access memory first from the primary subchannel (subchannel A for real address bit 56 = 0 (RA[56] = 0), and sub-channel B for RA(56) = 1). If read data cannot be delivered from the primary subchannel because of a read data error or other subchannel error conditions, then the read is retried on the secondary subchannel (subchannel B for RA(56) = 0 and subchannel A for RA(56) = 1), to get the correct read data.

The two registers MCFGP and MCFGPM BAR provide a non-mirrored and mirrored view of physical memory. Software must allocate mirrored and non-mirrored regions of memory such that these regions do not overlap in physical memory space.

3.4.4 Whole Memory Encryption

The memory controller supports Advanced Encryption Standard (AES) encryption and decryption of all traffic to system memory. Encryption is enabled through configuration bits accessible to firmware. Accesses to OMI configuration and MMIO spaces are never encrypted because they are not part of the system memory media. Other than that, if encryption is enabled, then all traffic to system memory is encrypted. Selective encryption of specific system memory regions is not supported.

3.4.5 Modes of Operation

If used, persistent DIMM technology keeps the data stored inside the DIMMs even if the power is turned off. If a memory card is replaced, the data stored in the DIMM would leave the data center in the clear. Therefore, an AES block cipher with strong encryption in the form of an AES XTS mode is supported as part of the Power10 design to protect the data.

In addition to the XTS mode, the AES Counter (CTR) mode is also supported. AES CTR is a low-latency AES bock cipher and is intended for memory encryption of volatile memory only. The goal is to protect against physical attacks.

Because memory latency in AES-XTS mode is higher than in AES-CTR mode, the encryption of XTS is much stronger. This stronger encryption of XTS is required for persistent memory because an attacker has arbitrary access to the persistent DIMM data when the card leaves the Data Center for repair or replacement. With the volatile DIMMs, the data will be gone after the power is removed. Therefore, AES CTR is a choice that makes it more difficult to physically gain data through the memory card interfaces with a reduced latency add.

The MCU crypto engine is designed as fully pipelined. Fetch and store bandwidths are not compromised by either of the encryption modes. Meta bits are not encrypted. All encryption relates to the data only.

3.4.6 Encryption Keys

Both the XTS and CTR modes use 128-bit encryption keys. The host boot (HB) firmware sets up the crypto mode to be used based on the system configuration. The MCU implementation supports that each OMI subchannel pair has its own keys. For CTR mode, such a key per OMI pair is recommended, but host boot can choose to either have a single system-wide key or a key per OMI subchannel pair.

Not all systems in all markets have this feature enabled. An eFuse on the Power10 chip can be failed to prevent the enabling of encryption. After the eFuse fails, the entire memory encryption function cannot be enabled for that chip.

3.5 PCI Express Controller

The PCIe Express controller (PEC) provides a PCIe root-complex port to connect to an adapter, slot, or as a link to a PCIe switch. It acts as a PCIe host bridge (PHB) from the internal, coherent SMP interconnect, which is also known as the processor bus, to the PCIe I/O. A PEC can support 16 lanes of Gen4 PCIe, which can be configured as 1–3 PCIe stacks, or it can support eight lanes of Gen5 PCIe plus eight lanes of Gen4 PCIe, which can be bifurcated.

3.5.1 Overview

The PEC is composed of six major building blocks:

- Processor bus common queue (PBCQ)
- Processor bus to ASIC interface bus (AIB) interface (PBAIB), which is an asynchronous boundary crossing
- Express transaction unit (ETU)
- PCIe ASIC building block (PCIASIC)
- Physical coding sublayer (PCS)
- Physical media access (PMA)

Figure 3-13 on page 78 shows an overview of the major blocks and defined interfaces.

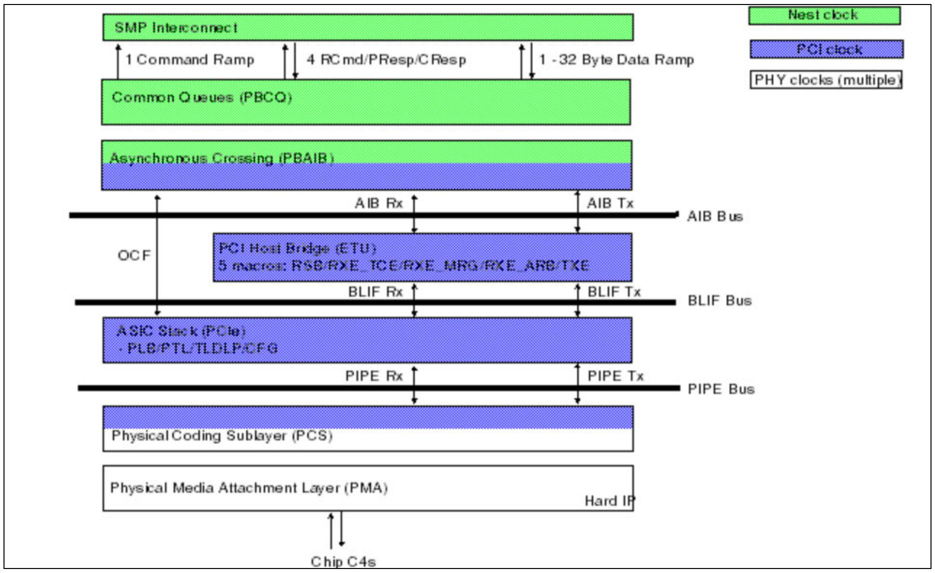


Figure 3-13 PCIe Major Blocks

3.5.2 Processor Bus Common Queues

The processor bus common queue (PBCQ) logic is responsible for managing the transactions on the coherent processor/cache fabric, the SMP interconnect.

The PBCQ includes the following key features:

- Inbound Direct memory attach (DMA) capability
  - Supports 96 DMA read transactions on the SMP interconnect. DMA read transactions are sourced from non-posted read transactions from the PCIe.
  - Supports 48 DMA write transactions on the SMP interconnect. DMA write transactions are sourced from posted write transactions from the PCIe.
  - PCIe 4-byte and 8-byte atomic operations.
- Outbound MMIO capability
  - Two Base Address Registers (BARs) for external MMIO address ranges
  - 32 MMIO stores
  - 128 MMIO loads
- The ability to share resources with more than one PHB stack

3.5.3 Express Transaction Unit

The ETU is responsible for address translation, interrupt management, and error isolation.

The ETU includes the following key features for SMT8 (SMT4):

- 512 KB (256 KB) partitionable endpoints
- 1 KB (512 KB) 4-way set-associative translation cache
- 4,000 (2,000) MSI interrupts supported
- Eight LSI interrupts supported

3.5.4 PCIe ASIC Intellectual Property

The PCIe ASIC building block is composed of the packet buffer layer (PBL), the packet transaction layer (PTL), the transaction and data link layer (TLDLP), and the PCIe Configuration Register core (CFG). These blocks implement the PCIe transaction and data link layers.

3.5.5 Physical Coding Sublayer

The PCS manages the low-level networking protocol and signaling between the physical media and the higher-level link protocol layer across the PIPE interface. The 16 lanes of the PCS can be bifurcated into 2 × 8 lanes or trifurcated into 1 × 8 and 2 × 4 lanes.

3.5.6 Physical Media Access

The PMA provides the Serializer/Deserializer (SERDES) and analog protocols necessary to connect to the chip C4s. It also provides the Phase-locked loop (PLLs) that is used to drive the PCI clock grid.

3.5.7 Power10 PCIe Configurations

The Power10 chip has two PCIe controllers of 16 lanes each for a total of 32 lanes of PCIe Gen4 I/O. The two PECs can have the following configurations:

- Gen4: 1 × 16 lanes, 2 × 8 lanes (bifurcation), or 1 × 8 lane and 2 × 4 lanes (trifurcation)
- Gen5: 1 × 8 lanes plus 8 lanes of Gen4 PCIe, which operate as a 1 × 8 stack or 2 × 4 stacks

Each grouping of lanes is called a stack and each stack has dedicated ETU and PCIe blocks. Each set of 16 lanes has only one PBAIB and PBCQ pair to interface to the SMP interconnect. The resources of the PBCQ are shared between the stacks that it services. If not all stacks in a PEC are used, the stacks should be used left to right as shown in Figure 3-14 on page 80. For example, if only one stack is used in PEC0, use PHB0 first.

Within a stack grouping, the lanes can be swapped to facilitate board wiring.

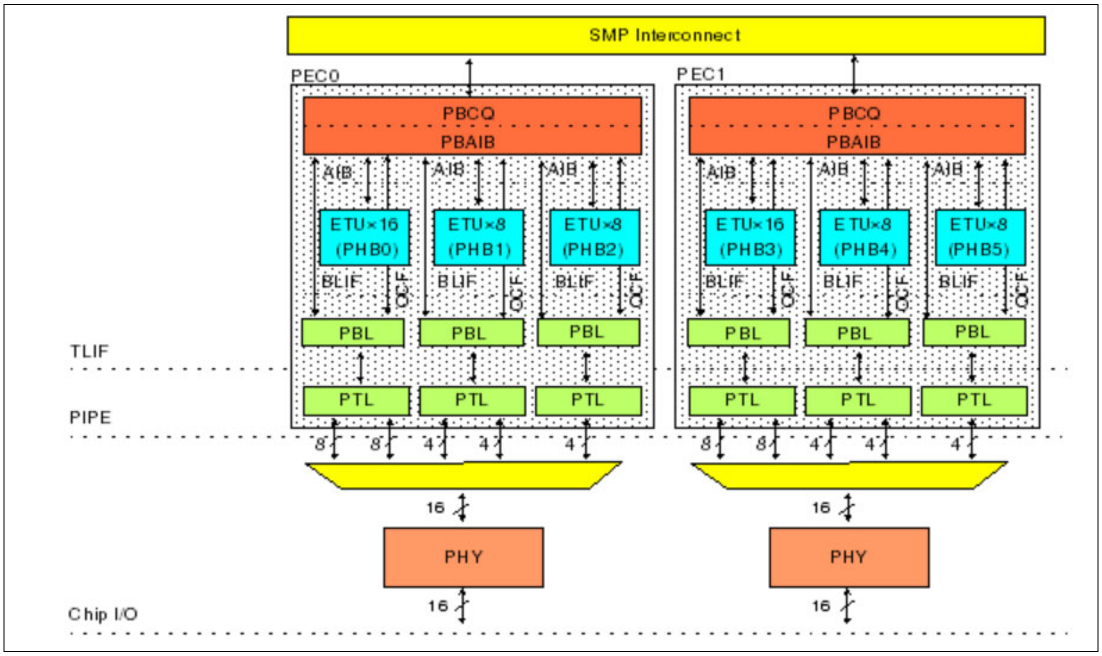


Figure 3-14 Power10 PCIe High-Level Diagram

3.5.8 Reliability, Availability, and Serviceability

Reliability, Availability, and Serviceability (RAS) includes the following features:

- End-to-end data protection from the processor bus Error correcting code (ECC) to the PCI packet Link cyclic redundancy check (LCRC) and end-to-end cyclic redundancy check (ECRC).
- Arrays have Single-error correction, double-error detection (SEC), and double-error detection (DED) ECC.
- Register files have parity and some have SEC/DED.
- Support all SMP interconnect parity/ECC.

Major control registers have parity protection.

Enhanced Error Handling (EEH)

If an error can be isolated to an endpoint, this endpoint is blocked from introducing new transactions until the error can be resolved.

Freeze Mode

A stack enters freeze mode when an error causes a reset of the stack. Freeze mode blocks all new transactions to and from a stack. Outstanding operations on the SMP interconnect run to completion and marks data as bad if required. Reset and initialization can be performed on the stack without a checkstop of the chip. A freeze on one stack does not affect the actions of another stack even if they share a PEC.

3.6 Partition Table Control Register

The Power10 core supports the Partition Table Control Register (PTCR) mostly as described in the *Power ISA (Version 3.1B)* for the purpose of accessing virtual memory either in virtual real mode (HPT) and guest real mode (nested Radix) or when translation is enabled for either HPT or Radix. The PTCR implements bits 12–51 for the Partition Table Base (PATB) field and bits 59–63 for the Partition Table Size (PATS) field. All other PTCR bits are reserved and return zero when read. However, the Power10 core ignores the value in the PATS field and only supports a 64 KB partition table size.

3.7 SMP Interconnect

The Power10 SMP interconnect is the underlying hardware used to create a scalable cache-coherent multiprocessor system. The Power10 SMP interconnect controller provides coherent and non-coherent memory access, I/O operations, interrupt communication, and system controller communication. The SMP interconnect provides all of the interfaces, buffering, and sequencing of command and data operations within the storage subsystem. The SMP interconnect is integrated on the Power10 chip with up to 16 SMT8 processor cores or 32 SMT4 processor cores and an on-chip memory subsystem. The Power10 chip has a maximum of eight SMP external links that can be used to connect to other Power10 chips.

The external SMP interconnect link is a split-transaction, multiplexed command and data bus that can support up to sixteen Power10 chips in a system. Aggregation of data links between the same source and destination chips is supported to increase data bandwidth.

Cache coherence is maintained by using a snooping protocol. Address broadcasts are sent to the snoopers, snoop responses are sent back in order to the initiating chip, and a combined snoop-response broadcast is sent back to all of the snoopers. Multiple levels of snoop filtering are supported to take advantage of the locality of data and processing threads. This approach reduces the amount of interlink bandwidth required, reduces the bandwidth required for system-wide command broadcasts, and maintains hardware enforced coherency using a single-snooping protocol. When the transaction cannot be completed coherently using chip scope, the coherency protocol forces the command to be re-issued to an increased scope of the system.

3.7.1 SMP Interconnect Features

General Features

Power10 provides enhance master command and data request arbitration.

Command requests are tagged and broadcast using a snooping protocol that enables high-speed cache-to-cache transfers.

Multiple command scopes are used to reduce the bus usage system wide. The SMP interconnect architecture uses cache states that indicate the following conditions:

- The last known location of a line (sent off chip)
- Information maintained in the system memory (memory domain indicator [MDI] bits)
- A coarse grained directory that indicates when a line has gone off the chip
- Combined response equations that indicate if the scope of the command is sufficient to complete the command or if a larger scope is necessary

The command snoop responses that are specified by the SMP interconnect implementation are used to create a combined response that is broadcast to maintain system cache state coherency. Combined responses are not tagged. Instead, the order of commands from a chip source, using a specific command-broadcast scope, is the same order that combined responses are issued from that source. The order is also affected by the snoop bus usage as well.

Power10-Specific Features

The Power10 SMP interconnect includes the following specific features:

- Command broadcast scopes (such as, snoop filtering):
  - Local Node Scope (LNS). Broadcast within a local chip with nodal scope. Node is defined as one chip.
  - Near Node Scope (NNS). Broadcast to a local chip and targeted chip in the local group.
  - Remote Node Scope (RNS). Broadcast to a local chip and targeted chip on a remote group.
  - Group Scope (GS). Broadcast to a local chip with access to the memory coherency directory (MCD).
  - Vectored Group Scope (VGS). Broadcast to a local chip and targeted remote chip.
- 1–8 sockets, 1-hop system configuration support
- 1–4 group, 2-hop system configuration support
- 4 × snoop bus support
- Memory controller fastpath support
- 256 Local master (LM) system-pump queue size (64 per snoop bus)
- 256 Group master (NM) group-pump queue size (64 per snoop bus)
- MCD with per-group bit vector. Service processor accessible Scan communications (SCOM) registers for configuration setup

On-Chip Features

The Power10 SMP interconnect includes the following on-chip features:

- Eight EQ core chiplets. Each chiplet contains two SMT8 cores or four SMT4 cores with a shared PBI chiplet interface (4 × data port).
- Four memory controller (MC) chiplets (2 × data port).
- Two endcap chiplets (1 × data port).
- Decentralized command-request arbitration.
- Dynamic command rate throttling.
- Translation look-aside buffer invalidate (TLBI) tokenizer.
- Decentralized data-request arbitration.
- 32-byte data arbitration size. Unit specifies total transfer size.

Off-Chip External SMP Features

The Power10 SMP interconnect includes the following off-chip features:

- 8 × (A or X link), 2 × 9-bit 25 Gbps and 32 Gbps links (asynchronous clocking)
- 2.0 M length (module and board)
- 22 UI MAX bit-lane skew
- Aggregate data-link support
- Indirect data-link support

Power Management Features

The Power10 SMP interconnect supports variable nest clock frequency (0.9–2.1 GHz). It also supports EQ chiplet power-management, which includes dynamic lane reduction support (external SMP links).

3.7.2 RAS Features

The Power10 SMP interconnect includes the following RAS features:

- Cyclic redundancy check (CRC) link-level retry on external SMP links
- 100% Error correcting code (ECC) protection on internal data flow
- Hang recovery mechanism
- Trace array
- Performance monitor
- FIR error reporting
  - Protocol errors
  - Underflow/overflow checkers
  - Asynchronous drop/repeat checkers
  - Parity checkers on coherency register files
- Error injection for single-bit and double-bit errors on external SMP links

3.7.3 External Power10 Fabric

The Power10 off-chip SMP interconnect is a highly scalable, multi-tiered, fully-connected topology. The off-chip links use 18-bit high-speed differential links running 25 Gbps or 32 Gbps and can be configured in either a 1-hop or 2-hop configuration.

In the 1-hop configuration, the Power10 processor chip can fully connect up to seven other processor chips to create an eight-chip SMP system. Each chip is a group using a maximum of seven inter-group A-links for a maximum system of eight processor chips.

In the 2-hop configuration, the Power10 processor chip can fully connect up to three other processor chips to create a four-chip group. The intra-group links are designated as X-links. Each Power10 processor in a group connects to its corresponding processor chip in each other group. Three of the inter-group A-links are provided per chip supporting a total of four groups, each containing four processor chips. A full four-group system of four chips per group consists of a maximum system of 16 processor chips.

3.8 Virtualization

SAP HANA support for Power10 provides support for the highest number of production virtual partitions that can be concurrently run on any system, providing clients with flexibility and granularity to consolidate with SAP landscape and reduce the server footprint.

The following features of Power10 provide the ability for customers to efficiently and securely run those consolidated workloads:

- System Control Unit (SCU). Virtualization control functions through a pair of redundant Flexible Service Processor (FSP) devices.
- Network virtualization is an area with significant evolution and improvements, which benefit virtual and containerized environments. The following recent improvements were made for Linux networking features on Power E1080:
  - SR-IOV allows virtualization of network cards at the controller level without the need to create virtual Shared Ethernet Adapters in the VIOS partition. It is enhanced with a (vNIC) virtual Network Interface Controller, which allows data to be transferred directly from the partitions to or from the SR-IOV physical adapter without transiting through a VIOS partition.
  - Hybrid Network Virtualization (HNV) allows Linux partitions to use the efficiency and performance benefits of SR-IOV logical ports and participate in mobility operations. Examples of mobility operations include active and inactive Live Partition Mobility (LPM) and Simplified Remote Restart (SRR). HNV is enabled by selecting a new Migratable option when an SR-IOV logical port is configured.
  - NVMe over Fibre Channel fabrics (NVMe-OF) is now available on Power E1080 running Linux when the 2-port 32 Gbps adapter (#EN1A or #EN1B) is used and the selected IBM FlashSystem® high-end (FS9110) and mid-range (FS7200) model are used.
- The IBM Power Virtualization Center (Power VS) is an enterprise tool that is built-in with no extra cost to help clients save on operation management and maintenance.

3.9 Capacity on Demand

Running SAP HANA with Power Enterprise Pools 2.0 is the best choice for controlling cost and flexibility to meet change from customer demands and ability to scale large environments:

- All installed processors and memory on systems in a pool are activated and made available for immediate use when a pool is started. Processor and memory usage on each server are tracked by the minute and aggregated across the pool.
- The capacity in this model consists of Base Activations and Capacity Credits, which are shared across the pool without having to move them from server to server. The unpurchased capacity in the pool can be used on a pay-as-you-go basis.
- IBM Power Systems Private Cloud with Shared Utility Capacity solution is supported only on specific Power9 and Power10 processor-based systems. Power E1080 servers can co-exist with Power E980 systems in the same pool.
- A single Power Enterprise Pool 2.0 supports a maximum of 2000 VMs. A maximum of 1000 VMs is supported per HMC. At the time of this writing, maximum of 200 VMs are supported per Power E1080 server, which is planned to be increased to 750 VMs per Power E1080.

3.10 Micro Partition

One of the significant features of IBM Power10 for SAP HANA is the memory control tool SPT (System Planning Tool). It can estimate the amount of memory needed to keep hypervisor always active or the amount of memory needed to deploy the new server.

IBM Micro-Partitioning® technology can allocate fractions of processors to an LPAR. An LPAR that uses fractions of processors is also known as a shared processor partition or a micropartition. Micropartitions run over a set of processors that are called a shared processor pool (SPP). Virtual processors are used to enable the operating system to manage the fractions of processing power that are assigned to the LPAR.

From an operating system perspective, a virtual processor cannot be distinguished from a physical processor, unless the operating system is enhanced to determine the difference. Physical processors are abstracted into virtual processors that are available to partitions.

On the Power10 processor-based server, a partition can be defined with a processor capacity as small as 0.05 processing units. This number represents 0.05 of a physical core. Each physical core can be shared by up to 20 shared processor partitions, and the partition's entitlement can be incremented fractionally by as little as 0.05 of the processor. The shared processor partitions are dispatched and time-sliced on the physical processors under the control of the Power Hypervisor. The shared processor partitions are created and managed by the HMC.

The Power E1080 supports up to 240 cores in a single system and 1000 micropartitions. PowerVM supports a maximum of 1000.

> **Note:** Although the Power E1080 supports up to 1000 micropartitions, the real limit depends on application workload demands in use on the server.

3.11 Additional resources

For more information about technical for Power10, see *IBM Power E1080 Technical Overview and Introduction*, REDP-5649 chapter 2, page 49

Related publications

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

IBM Redbooks

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

- *IBM Power Systems Security for SAP Applications*, REDP-5578
- *IBM Power Systems Virtualization Operation Management for SAP Applications*, REDP-5579
- *SAP HANA on IBM Power Systems: High Availability and Disaster Recovery Implementation Updates*, SG24-8432
- *SAP Landscape Management 3.0 and IBM Power Systems Servers*, REDP-5568

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

ibm.com/redbooks

Online resources

These websites are also relevant as further information sources:

- Capacity on Demand:

  https://www.ibm.com/it-infrastructure/power/capabilities/capacity-on-demand
- Guide Finder for SAP NetWeaver and ABAP Platform:

  https://help.sap.com/viewer/nwguidefinder
- IBM POWER9 architecture:

  https://ibm.co/2YMptRt
- IBM POWER9 processor overview:

  https://ibm.co/2qQ3YCU
- IBM PowerVM Virtualization:

  https://www.ibm.com/us-en/marketplace/ibm-powervm
- Live Partition Mobility:

  https://ibm.co/35jORlJ
- SAP Note 2055470:

  https://me.sap.com/notes/2055470
- SAP Note 2230704:

  https://me.sap.com/notes/2230704/E

- SAP Support Portal:
  https://support.sap.com/en/index.html
- Software Logistics Tools:
  https://support.sap.com/en/tools/software-logistics-tools.html
- Virtual Persistent Memory:
  https://ibm.co/2rGztQe
- Welcome to the SAP Help Portal:
  https://help.sap.com

Help from IBM

IBM Support and downloads

ibm.com/support

IBM Global Services

ibm.com/services

REDP-5569-01

ISBN DocISBN

Printed in U.S.A.