

SAP HANA on IBM Power Systems Architectural Summary

Dino Quintero

Damon Bull

Vinicius Cosmo Cardoso

Cleiton Freire

Eric Kass



`d' Analytics

Power Systems





IBM Redbooks

SAP HANA on IBM Power Systems Architectural Summary

April 2020

Note: Before using this information and the	ne product it sup	pports, read the info	rmation in "Notices"	" on page v.
Note: Before using this information and the	ne product it sur	oports, read the info	rmation in "Notices'	" on page v.
First Edition (April 2020)				
This edition applies to: ► SUSE Linux Enterprise Server V12 SP4 ► SAP HANA V2.0 SPS04	4			

© Copyright International Business Machines Corporation 2020. All rights reserved.

Note to U.S. Government Users Restricted Rights -- Use, duplication or disclosure restricted by GSA ADP Schedule Contract with IBM Corp.

Contents

Notices Trademarks v
Preface
Chapter 1. Solution approach with SAP HANA on IBM Power Systems1.1 Solution approach with SAP HANA on IBM Power Systems21.2 SAP HANA architecture21.3 SAP HANA sizing considerations and advantages on Power Systems31.4 SAP HANA Multi-Database Containers considerations41.4.1 Moving or copying an SAP HANA tenant51.5 Scale-up and scale-out on Power Systems servers51.5.1 Scale-up benefits and challenges of Power Systems servers61.6 Backup and recovery71.6.1 SAP HANA overview71.6.2 Basic backup tools91.7 High availability and disaster recovery121.7.1 SUSE Linux Enterprise High Availability Extension141.7.2 SAP HANA HADR options on Power Systems servers15
Chapter 2. Hardware advantages of IBM Power Systems servers192.1 IBM POWER9 processor202.2 Memory architecture212.3 PCIe212.4 Virtual Persistent Memory storage212.5 Central electronics complex expansion architecture212.6 Virtualization222.6.1 Dedicated processor partitions222.6.2 Shared processor partitions222.7 Capacity on Demand222.8 Live Partition Mobility222.9 Additional resources23
Related publications 25 IBM Redbooks 25 Online resources 25 Help from IBM

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.

Redbooks (logo) 🔗 ®	IBM®	POWER8®
AIX®	IBM Cloud™	POWER9™
DB2®	Parallel Sysplex®	PowerVM®
GDPS®	POWER®	Redbooks®

The following terms are trademarks of other companies:

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

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

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.

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

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

Preface

This IBM® Redpaper publication delivers SAP HANA architectural concepts for successful implementation on IBM Power Systems servers.

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. Moreover, 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 Knowledge Center, and it aligns with educational materials that are provided by IBM Systems.

Authors

This paper was produced in close collaboration with the IBM SAP International Competence Center (ISICC) in Walldorf, SAP Headquarters in Germany and IBM Redbooks®.



Dino Quintero is an IT Management Consultant and IBM Level 3 Senior Certified IT Specialist with IBM Redbooks in Poughkeepsie, New York. He has 24 years of experience with Power Systems technologies and solutions. Dino shares his technical computing passion and expertise by leading teams developing technical content in the areas of enterprise continuous availability, enterprise systems management, high-performance computing, cloud computing, artificial intelligence (AI) (including machine and deep learning), and cognitive solutions. He is a Certified Open Group Distinguished IT Specialist. Dino holds a Master of Computing Information Systems degree and a Bachelor of Science degree in Computer Science from Marist College.

Damon Bull is a Senior SAP Software Performance Engineer who has worked with different SAP versions from 2.2F onward. His focus is on SAP performance, and he has conducted and published over 50 SAP benchmarks over the last 25 years. Damon has experience with IBM DB2®, Oracle, SQL Server, and SAP HANA databases (DBs), and SAP Sales and Distribution (SAP SD), SAP Business Warehouse (SAP BW), SAP Transaction Banking (SAP TRBK), SAP Financial Accounting (SAP FI), Assemble-to-Order (ATO), and others running on IBM AIX®, Linux, and Windows servers. Damon holds a Bachelor of Science in Computer and Information Sciences from the University of Oregon.

Vinicius Cosmo Cardoso is an IBM Certified Expert IT Specialist and Senior SAP Basis Administrator working for IBM Brazil. He has 12 years of experience in SAP, working on complex projects like new implementations, upgrades, and platform migrations. He is an SAP certified professional for SAP NetWeaver, SAP OS/DB Migration, and SAP HANA. Vinicius holds a Master of IT Solutions Architecture from 2018 and a Bachelor of Computer Science from 2008. Throughout these 12 years, he has worked for several Brazilian and global clients in several industry areas.

Cleiton Freire is an IBM Cloud™ Solution Leader for the Europe, Latin America, and Canada cluster. Cleiton has almost 20 years in the Information Technology industry and throughout the past 14 years working in IBM. He has served as a subject matter expert on SAP on IBM Power platform migrations, installations, and administration for large customers in Brazil, the United States, and Europe. For the past 2 years, Cleiton has helped customers find the best tech solution to address their business problems. Cleiton is a certified Technology Consultant for the SAP HANA platform and holds a Bachelor of Science degree from the Faculty of Technology in Sao Paulo, Brazil.

Eric Kass is an IBM Developer Architect in Worldwide SAP Technical Enablement, which is part of an IBM Systems group that is based in Boeblingen, Germany. Eric started with IBM in 1997 in operating system (OS) network and communication development, where he wrote device control and network routing code. In 2000, he spent a year with SAP to write the DB layer so that SAP can communicate between Windows and IBM i. He became a successful IBM Developer and architect working at the SAP development lab, where he designs and codes DB drivers; SAP kernel components; and SAP and DB high availability (HA) engines, network drivers, and security protocols. Eric has over 20 patents in SAP-related fields, and he supports worldwide SAP customers on IBM i and AS400 platforms.

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

Wade Wallace IBM Redbooks, Austin Center

Katharina Probst, Walter Orb, Tanja Scheller IBM Germany

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

Solution approach with SAP HANA on IBM Power Systems

This chapter provides IBM POWER® architectural details and about how SAP HANA takes advantages of the technology for its deployment on Power Systems servers.

This chapter covers the following topics:

- Solution approach with SAP HANA on IBM Power Systems
- SAP HANA sizing considerations and advantages on Power Systems
- SAP HANA Multi-Database Containers considerations
- Scale-up and scale-out on Power Systems servers
- Backup and recovery
- ► High availability and disaster recovery

1.1 Solution approach with SAP HANA on IBM Power Systems

This section provides information about SAP HANA technical characteristics and architectural information for implementing it on Power Systems servers.

1.2 SAP HANA architecture

The SAP HANA database (DB), as with most high-performance DB applications, is developed in C and runs on SUSE Enterprise Linux on Power Systems server. SAP HANA is composed of several components.

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

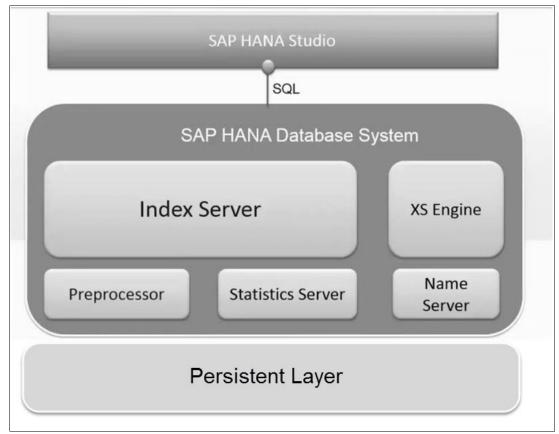


Figure 1-1 SAP HANA: Architecture overview

Here are the descriptions of these components:

Persistent Layer Even though SAP HANA is an in-memory DB, it has a mechanism to ensure that after a system restart, planned or unplanned, it is possible

return to the most recent state. On SAP HANA, this important task is

accomplished on the Persistent Layer.

Preprocessor 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 tenant DBs.

Name server This component runs on the system DB (exclusively), and it is the

directory of SAP HANA. Information about all running components can be found here, including distributed components in multi-host servers.

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 tenant DB.

XS Engine An optional component that you can use to connect to an SAP HANA

DB to fetch data by using HTTP.

1.3 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. Due to 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: CPU, memory, storage size, I/O, and network bandwidth. The process is a combined effort of customers and hardware vendors (including Power Systems) with SAP.

Figure 1-2 shows an overview flow of the sizing process.

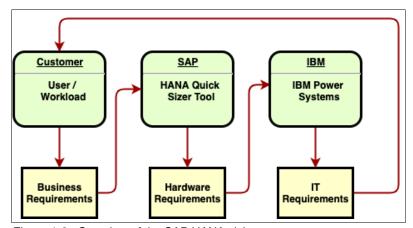


Figure 1-2 Overview of the SAP HANA sizing process

An SAP HANA on Power Systems implementation is based on the Tailored Datacenter Integration (TDI) model, which allows customers to cut 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 up to 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.

1.4 SAP HANA Multi-Database Containers considerations

With the introduction of Multi-Database Containers (MDC), which was made available with SAP HANA 1.0 SPS09, SAP enabled SAP HANA with DB multitenancy, where a single SAP HANA system with all its underlying infrastructure can be shared among several DB tenants. Converted or newly installed systems (with the latest SAP HANA version) have at least the system tenant, which synchronizes the systemwide landscape information monitoring, connectivity, and DB configuration across the system, and the main tenant is used for a particular application configuration.

With the default implementation, it is possible to have up to 20 tenants in a single SAP HANA system, but that limit can also be relaxed (assuming enough resources are available) by increasing the number of ports for more tenants in the configuration.

Figure 1-3 illustrates a typical MDC scenario with SAP HANA on Power Systems.

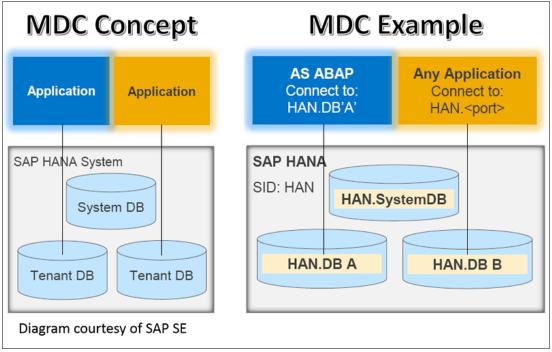


Figure 1-3 MDC configuration

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 benefits of using multitenancy with Power Systems are summarized as follows:

- Integrated system administration and monitoring
- Optimized systems resources utilization
- ► 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, and instead of performing the same activity in multiple systems, it is done only once. Another benefit is the reutilization 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, where 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, and others). However, the reality is that memory, vCPU, and storage resources can be saved by using the combination approach because impact of the OS and DB can be shared. In addition, tailored sizing allows a much more demand-driven combination.

1.4.1 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 add flexibility to system management because you can adjust the SAP HANA tenant placement according to the current demand.

1.5 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 classical challenge to IT business due to the degree of uncertainty to predict growth rates of computational power requirements (processing, memory, storage, and bandwidth). The infrastructure scalability to handle the application demands is addressed through two main vectors: scale-up (vertical) and scale-out (horizontal).

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 multiples nodes (hosts). The main purpose of this approach is to overcome the limitations of a single server, but you can also use this approach to gain performance or to meet a specific application requirement. Several considerations must be accounted for this approach:

- Verify 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.
- Application-specific demands 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.

Power Systems servers support both scale-up and scale-out modes for SAP HANA systems. Scale-up mode addresses the demand of different workloads depending on the individual characteristics that are related to node synchronization, for example, if the use case is an Online Transactional Processing (OLTP) system or Online Analytical Processing (OLAP) system. For SAP Business Warehouse workloads on SAP HANA, the use case supports up to 16 nodes in one server, and for SAP Business Suite workloads on SAP HANA, it supports up to four nodes, as shown in Figure 1-4.

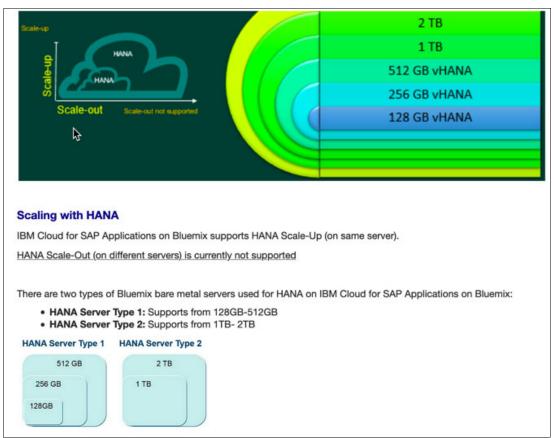


Figure 1-4 Scale-up node limits

Power Systems serves are designed to help you maintain performance while growing with efficiency and maintaining performance under required throughput growth for a longer period. For SAP HANA systems, both IBM and SAP recommends that you use the scale-up approach if your hardware systems can deliver the required sizing and performance (for more information, see SAP Note 2408419.

Tip: As a best practice, use scale-up configurations if they are economically justifiable, that is, account for operational costs and drawbacks.

1.5.1 Scale-up benefits and challenges of Power Systems servers

SAP documentation states that the scale-out configuration for SAP HANA, especially for OLTP workloads, has challenges that must be carefully assessed before you implement the OLTP model.

1.6 Backup and recovery

Although backups are a crucial component of any company business continuity plans, sometimes organizations do not have a valid backup, which causes serious operational problems and various business losses.

When it comes to SAP HANA on Power Systems, a good backup strategy that is well implemented and easy to operate ensures that a backup is available when necessary.

This section provides information about how to get started with the backup strategy, tools, and implementation that are required to implement SAP HANA on Power Systems for backup, which includes information about the hypervisor (PowerVM) and the SAP HANA DB tenant. In this use case, we use SAP HANA v2.0 SPS04 running on an IBM POWER9™ processor-based server with SUSE Linux Enterprise Server 12 SP4 as the OS for the SAP applications.

Here are the topics that are covered:

- 1. You start with an overview of SAP HANA technology, the main components, and the available deployment scenarios. The backup options are influenced by the architectural decision of the SAP HANA implementation, so you must know the basics of the SAP HANA architecture that is available for Power Systems servers.
- 2. Then, you learn about the basic backup tools for the supported OS for SAP HANA on Power Systems. You learn about the commercial options and alternatives for OS snapshot, which the administrator uses to create exact copies of the underlying OS of an SAP HANA instance virtual machine (VM).
- 3. Then, you learn about the backup utilities that are available for SAP HANA on Power Systems, which include the standard backup tools that are available for SAP HANA, such as Backup Studio, Backup Cockpit, their options, the connectivity between Backint and third-party tools, an overview of the certified tools for SAP HANA on Power Systems, and best practices on the platform.
- 4. Finally, you learn about the backup and restore procedures for an SAP HANA on Power Systems installation. The focus here is the SAP HANA DB. You learn about best practices for backup and backup types, see a comparison among some different backup strategies, explore sizing considerations, and review some backup examples by backup type.

1.6.1 SAP HANA overview

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 compress it. 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.¹

SAP HANA is an excellent tool for real-time analytics that does not require extra data warehouse systems and regular extract, load, and transform (ETL) processing. Data can be consolidated and recalculated for specific reports inside the OLTP system that is powered by SAP HANA.

¹ Source: https://www.sap.com/products/hana/features/in-memory-database.html

For backup and recovery strategies, you must have fundamental knowledge about the SAP HANA technical deployment options and how they affect these strategies. The architectural implementation options are:

Single-Host System The single host system is the most simple

and straightforward deployment scenario, where you have all components running on a single SAP HANA system (whether it is on-premises or in the cloud as a VM,

for example).

SAP HANA Multiple Host System A multiple-host or distributed SAP HANA

system is a system that is installed on more than one host. Components are spread

across multiple hosts.

SAP HANA MDC (Multitenant) Multitenant is a software principle where a

single instance of the software is used by

several instances simultaneously.

Multiple SAP HANA Systems on One Host Multiple Components on One System

(MCOS) is where several SAP HANA

systems run on the same host.

For this publication, use the Single-Host System architecture scenario with the Single Application on One SAP HANA System (SCOS) deployment option because they encompass the fundamental components of every SAP HANA landscape architecture and keep it simple for explanation purposes. Whenever necessary, there are notes that contain relevant information about other combinations.

On a SCOS system, the backup approach is similar to standard DBs where you have a single set of components that are included in the backup.

Note: Although we are using Single-Host Systems with SCOS as the basis for operations and strategy for backup for the use case in this paper, there are specific sections that describe best practices and implementation tactics for distributed systems (multi-host), nearline storage (NLS), Multitenant Database (standard for SAP HANA 2.0 SPS01 onwards), and MCOS.

While performing a basic initial backup of a single SAP HANA system, SAP HANA supports the following functions:

- ► Full backups for both data backups and snapshots.
- ► Delta backups for incremental and differential backups.
- ► Redo log backups.
- Backups and recovery by using third-party tools.
- Integrity checks for backups.
- ► Backup lifecycle management.
- ► Recovery to a point-in-time.
- Recover a specific data backup or data snapshot.
- ▶ DB copies by using backup and recovery.

1.6.2 Basic backup tools

Basic backup tools and recovery utilities are the backup utilities that you find in the standard SUSE Enterprise Linux for SAP Applications on IBM Power Systems distribution. These tools and utilities include command-line tools like **tar**, **cpio**, and **rsync**, which are part of most Linux distributions, but there are tools that are specific to SUSE Linux Enterprise Server, such as Snapper.

Most of these backup tools have been part of UNIX and later Linux OSs for a long time, and some of them include features that are not aggregated into commercial backup tools. Therefore, these tools are a good fit, inexpensive, and highly customizable solutions for doing the backups. For now, these tools can be used until a more robust commercial backup is available.

Note: It is never a good idea to run standard backup or synchronization tools against running DBs such SAP HANA because the resultant backup file can be corrupted.

This section describes the benefits and drawbacks of some of the utilities that are available on the SUSE distribution for Power Systems servers. Among them, **cpio** has the best functions, but it is not as easy to use as **tar**, which is the most commonly used backup tool in the batch.

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 come 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, time stamps, 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, which eliminates many constraints that represent in regular tar, such as file name length and preservation of time stamps.

The utility is the most used archive 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 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 even full backups.

Table 1-1 shows the **cpio** benefits and drawbacks.

Table 1-1 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
Preserves hard links.	meaningful differences between releases.
File length limitation.	
Timestamps are preserved by default.	

The **cpio** utility can be a powerful backup tool. The main feature that sets it apart is its ability to accept the list of files to be backed up from standard input (although 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 SUSE 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 **0utFile** 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, and can even copy data from stdin to stdout while modifying it while it is in route. Again, 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 pax (portable archive exchange) 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 actively parses files that are being restored. If a file blocks aligned and sized areas that are NULL populated, pax does not allocate physical space for those file system blocks. The file size in bytes remains the same, but the actual space that is taken within the file system is for the non-NULL areas.

A typical pax command creates a disk backup, as shown in Figure 1-5.

Figure 1-5 Using the pax command for 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 also can 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 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 while focusing 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.

1.7 High availability and disaster recovery

Although business continuity plans are a critical requirement for successful businesses, it is still common to find customers that 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. Then, the consultants show that every improvement increases costs, so 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 utilization based on its consumption. As expected, the utilization is higher during business hours, but with remote access, utilization often occurs outside of business hours too. Assessing the number of users per hour can help to identify the criticality of the system then.

Other considerations that needs 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 this information must be collected to determine the systems' true availability.

Availability is measured by the SLA, and the systems utilization data that is mapped is used to determine the systems' allowed outages per month, which then can be used to determine the relevant SLA, as shown in Table 1-2.

Table 1-2 Typical service-level agreements

Maximum outage (minutes)	SLA percentage (%)
864 (14.4 hours)	98
600 (10 hours)	98.6
432 (7.2 hours)	99
216 (3.6 hours)	99.5
131 (2.18 hours)	99.7
42	99.9
7	99.99

Now, there are some other requirements that 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 also play a role here. 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

Several 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, you should have 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 1-6):

- Recovery Period Objective (RPO): The maximum amount of time that a system can be down.
- ► Recovery Time Objective (RTO): The maximum period that is required to recover the system.

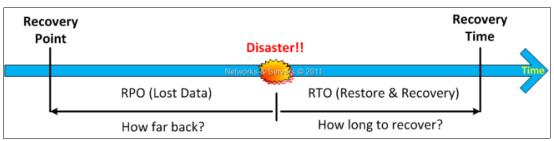


Figure 1-6 Disaster recovery: RPO and RTO

After the availability requirements are clear 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, like CPUs, storage devices, input/output devices, firmware, applications, power, and cooling.

1.7.1 SUSE Linux Enterprise High Availability Extension

SAP HANA is supported only on Power Systems servers that run SUSE Linux Enterprise Server, so the recommended HA solution is SUSE Linux Enterprise High Availability Extension, which is a featured 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 a 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 SAP HANA HADR options on Power Systems servers

SAP HANA has a toolset 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 to the standby server. Multiple standby hosts can be added to increase the SLA goals of the system, which addresses full host issues by providing a hardware ready node, as shown in Figure 1-7.

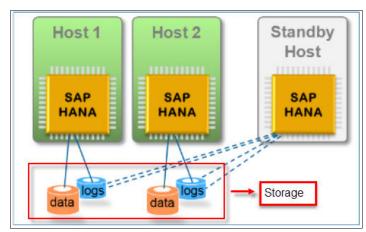


Figure 1-7 SAP HANA: Host-Auto Failover

Because the data is preinstalled in the standby system, both RPO and RTO on this scenario tend to towards 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 1-8.

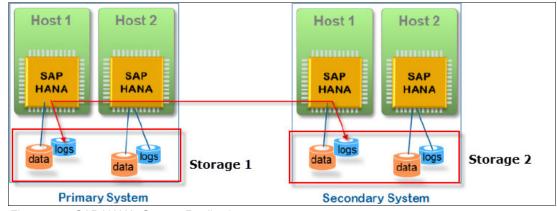


Figure 1-8 SAP HANA: System Replication

This scenario is similar to classical shadow DB solutions. The automation of the failover is done by using SUSE Linux Enterprise High Availability Extension on IBM Power Systems. Additionally, this scenario is used both 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 HANA, and commit transactions only after the replication is complete (synchronous replication). The distance between the primary and secondary sites is limited up to 100 kilometers (one to a few hops, with less than ~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 using the Storage Replication solution, SAP system management can be automated with the SUSE Linux Enterprise High Availability Extension, as shown in Figure 1-9.

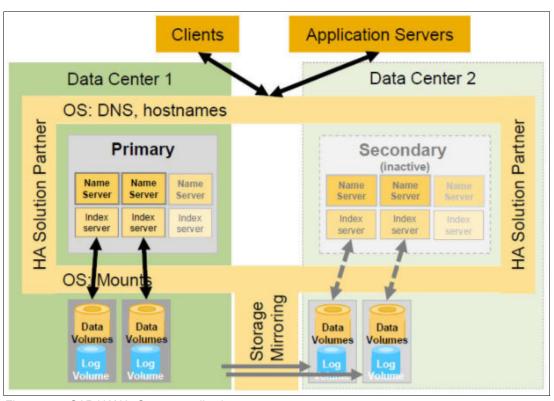


Figure 1-9 SAP HANA: Storage replication

Geographically Dispersed Resiliency (GDR)

For over 10 years, the IBM Geographically Dispersed IBM Parallel Sysplex® (IBM GDPS®) family of DR solutions have supported VM restart-based DR. GDR is the same technology that is available on Power Systems servers, as shown in Figure 1-10 on page 17.

GDR is a simplified way to manage DR. It has the following features:

- The VM restart technology has no OS or middleware dependencies.
- Automated DR management.
- Easier deployment for DR, unlike clustering or middleware replication technologies.
- Eliminating hardware and software resources at the backup site to save money.

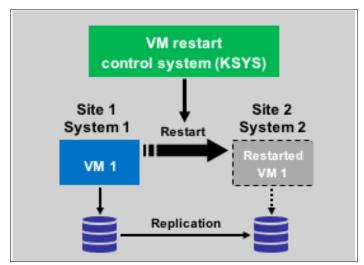


Figure 1-10 GDR VM move to a second site

GDR supports SAP HANA with DR failover of the environment (cold start). Along with storage replication, it can make up a complete DR solution, as shown in Figure 1-11.

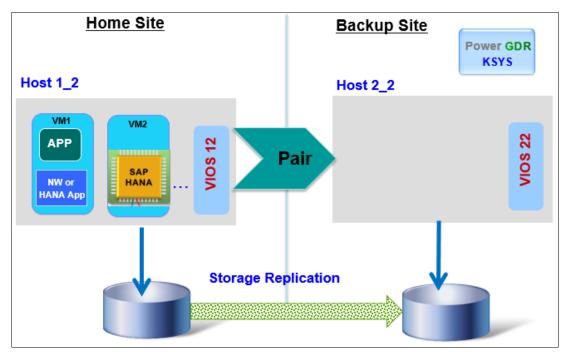


Figure 1-11 SAP HANA reference configuration: DR that uses GDR technology

Table 1-3 summarizes the pros and cons of each option.

Table 1-3 Scenario summaries

Scenario	Pros	Cons	
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. 	
GDR	 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. 	



Hardware advantages of IBM Power Systems servers

IBM POWER processor-based servers provide many technological capabilities for running SAP HANA on Power Systems. The POWER 9 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 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:

- ► IBM POWER9 processor
- Memory architecture
- ► PCle
- ► Virtual Persistent Memory storage
- Central electronics complex expansion architecture
- Virtualization
- Capacity on Demand
- Live Partition Mobility
- ► Additional resources

2.1 IBM POWER9 processor

The POWER9 processor cache hierarchy includes a 32 K L1 instruction and 64 K L1 data cache, a 512 K 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 PCle 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 like SAP HANA.

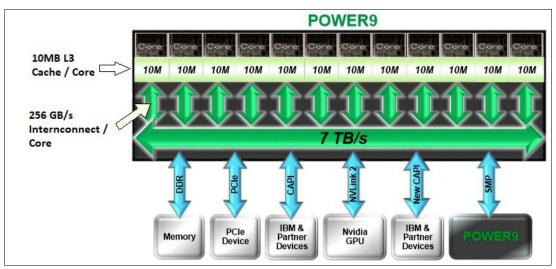


Figure 2-1 IBM POWER9 processor bus

The POWER9 processor family is fully binary compatible for applications that are optimized for earlier POWER processors, although providing 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 makes it easy 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 (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 4 TB per socket (16 x 128 GB), producing 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 PCle

POWER9 processor-based servers use the new PCIe Gen4 bus slot technology, which doubles transfer rates and throughput compared to the PCIe Gen3 design. For large-scale systems, PCIe Gen4 busses 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 increases the start times of the SAP HANA DB while keeping data consistency 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, and 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 new 25 GBps optical-style interconnect signal technology, which is four times faster than the 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 (processors, memory, I/O adapters, and others) 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 LPAAR 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 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

PowerVM can move LPARs from one physical machine to another one while the operating system (OS) and application are live by moving the memory contents and the persistent disk resources over a high-speed network connection to the receiving system. SAP HANA active (live) and inactive (shutdown) migrations are supported.

Migrations can be done between the same or different processor architectures. This compatibility feature allows LPARs to be easily moved from older processor-based systems, for example POWER8 to POWER9 systems. This feature is the easiest way for SAP HANA systems to upgrade hardware to more recent platforms while minimizing application downtime.

2.9 Additional resources

For more information about technical resources for the IBM POWER architecture, see "Related publications" on page 25.

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/35jOR1J

► SAP Note 2055470:

https://launchpad.support.sap.com/#/notes/2055470

SAP Note 2230704:

https://launchpad.support.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-5569-00 ISBN 0738458686

Printed in U.S.A.







