Redbooks
ibm.com/redbooks

IBM Power Systems Planning and Monitoring Best Practices for SAP Applications

Dino Quintero
Sahitya K. Jain
Ralf Kokowski
Katharina Probst
Marc-Stephan Tauchert

Analytics

Power Systems

IBM

Redpaper

IBM Redbooks

IBM Power Systems Planning and Monitoring Best Practices for SAP Applications

October 2020

REDP-5580-00

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

First Edition (October 2020)

This edition applies to:
- SUSE Linux Enterprise Server V11, V12, and V15
- Red Hat Enterprise Linux V6, V7, and V8
- SAP NetWeaver 7.5
- Performance monitor (nmon) V16g

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.

Redbooks (logo)  ®	POWER®	PowerPC®
AIX®	Power Architecture®	PowerVM®
DB2®	POWER8®	PureFlex®
IBM®	POWER9™	Redbooks®

The following terms are trademarks of other companies:

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

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

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

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

Preface

This IBM® Redpaper publication is part of a series of documents that supports clients and architects to plan, implement, and operate SAP workloads on IBM Power Systems servers. The focus of this publication is to plan and implement best practices for SAP workloads, and monitor the landscape for health.

The target audiences of this book are architects, IT specialists, and systems administrators deploying SAP workloads who often spend much time and effort managing, provisioning, and monitoring SAP software systems and landscapes on IBM Power Systems servers.

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 IBM 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 (HPC), cloud computing, artificial intelligence including machine and deep learning, and cognitive solutions. He also is a Certified Open Group Distinguished IT Specialist. Dino is formerly from the province of Chiriqui in Panama. Dino holds a Master of Computing Information Systems degree and a Bachelor of Science degree in Computer Science from Marist College.

Sahitya K. Jain is an Advisory Software Engineer who works for SAP platform support with IBM System Labs. He has over 12 years of experience in working with Power Systems servers. He has worked on functional verification testing for Virtual I/O Server (VIOS) and IBM AIX®. In his current role, he supports Power Systems customers running SAP applications, such as NetWeaver or SAP HANA. Sahitya holds a Bachelor of Engineering (Computer Science) degree from Visvesvaraya Technological University, Belagavi, India.

Ralf Kokowski is a Senior Software Developer and Engineer working for IBM Systems Labs in Germany. He has 10 years of experience with the IBM POWER® platform that uses AIX and Linux for SAP applications. In earlier years, he worked in different Software Engineering roles for IBM AIX environments. In his current role, he is responsible for IBM Power platform support for SAP applications, as NetWeaver and HANA. Ralf holds a German degree of Diplom Physiker (comparable to Master of Science in Physics) from Ludwig-Maximilians-University Munich, and a German degree of Diplom Informatiker (comparable to Master of Computer Science) from Friedrich-Alexander-University Erlangen-Nuremberg.

Katharina Probst leads the IBM Development team inside the SAP Headquarters that is responsible for the Datacenter Readiness and Eco System enablement of SAP HANA on Power Systems servers. She has 15 years of background in SAP on Power Systems servers in the area of storage, business continuity, disaster recovery (DR), AIX, and Linux.

Marc-Stephan Tauchert is a System Specialist at IBM Germany. In his 25+ years of experience in designing and implementing SAP Solutions on IBM Power Systems, he has covered various topics and roles. As a Technical Sales Consultant and SAP Solution Architect, he supports customer situations regarding the connectivity of SAP Applications and Infrastructure, SAP sizing, and hybrid solutions. He is an expert in SAP Database and Application performance, including SAP HANA. He is member of the IBM HANA on Power Systems Development Team in Sankt-Leon-Rot.

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

Wade Wallace
IBM Redbooks, Austin Center

Walter Orb and Tanja Scheller
IBM Germany

Ricardo Dobelin Barros
IBM Brazil

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

Planning and implementation best practices for SAP workloads

This chapter describes the general aspects to plan, implement, and operate SAP workloads on IBM Power Systems servers.

This chapter includes the following sections:

- Introduction
- Planning and implementation best practices for SAP workloads
- High-level SAP installation selection in the Product Availability Matrix
- SAP applications configurations
- Linux installation
- Linux customization with SAP base patterns
- Linux tuning for SAP workloads
- Hardware aspects

1.1 Introduction

This publication is part of a series of technical documents supporting clients in planning, implementing, and operating SAP workloads on Power Systems servers. The focus of this publication is to plan and implement best practices for SAP workloads, and monitor the landscape for health.

1.2 Planning and implementation best practices for SAP workloads

This chapter outlines the typical process of planning and implementing SAP software solutions that are based on the underlying software and hardware layers for various business needs. This approach is a top-down one, from higher to lower layers of the application stack.

The conceptual starting point is selecting an SAP product to fulfill a business need. This paper assumes that this decision was already made.

Section 1.3, “High-level SAP installation selection in the Product Availability Matrix” on page 3 introduces the procedure about how to select the correct SAP software components for an Power System server-based installation. This task can conceptually be thought of as selecting a platform, in a generalized sense, for the SAP product. The term *generalized platform* here defines the operating system, the database, and possibly other infrastructural components, which together build the components on which the SAP applications run.

Section 1.4, “SAP applications configurations” on page 9 briefly describes the SAP parameterizations for the selected software components, including the important aspects of sizing. SAP parameterization and sizing details are obtained from SAP documentation (Quick Sizer, SAP Notes, and knowledge base articles), including the calculation of resource demands that is required from the operating system and hardware (which is not in the scope of this document). When assistance with sizing is required, IBM Techline can be contacted.

The Linux parameterization for SAP application requirements are described in 1.6, “Linux customization with SAP base patterns” on page 13 and 1.7, “Linux tuning for SAP workloads” on page 15.

Finally, the best practices aspects of the lowest layer, the hardware, are described in this chapter.

1.3 High-level SAP installation selection in the Product Availability Matrix

After the decision for an SAP product is made based on your business needs, for example, enterprise resource planning or developing customized reports, the correct technical components must be selected to implement it.

From the perspective of a top-down approach that starts at the application layer, some lower layers can be identified, which together build a fully configured application stack:

- SAP kernel
- Database
- Operating system
- Server and extra hardware

A tool that is provided by SAP for this task is called the Product Availability Matrix (PAM), and it is available at SAP Maintenance.[1]

The PAM specifies the possible supported combinations of SAP products and required technical components for many use cases. Figure 1-1 provides a high-level architectural overview about how the PAM outlines the components of an SAP solution stack.

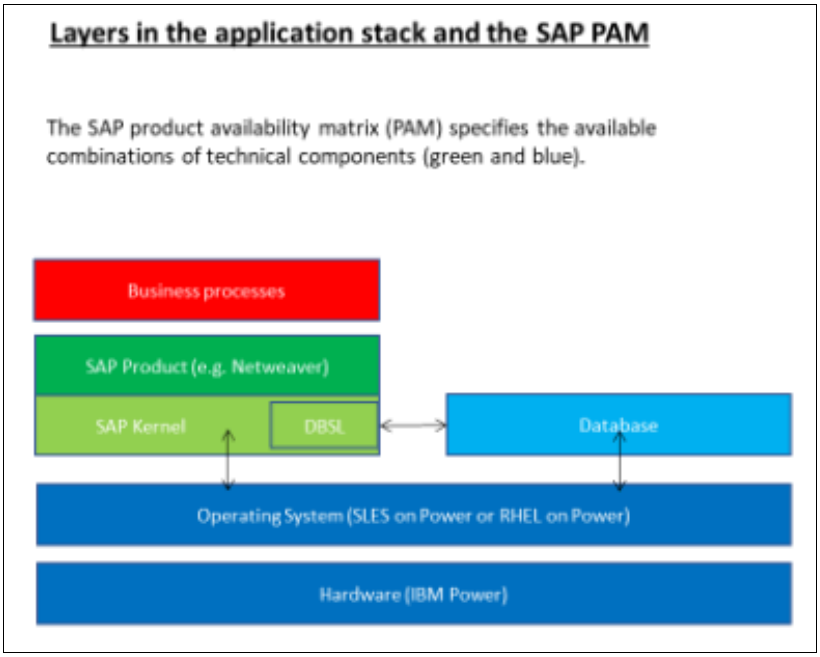

Figure 1-1 Schematic structure of the SAP Product Availability Matrix

In a traditional sense, choosing a platform is thought of as choosing the operating system and hardware, or even just the database component. SAP from the PAM perspective defines the term *platform* as the operating system, hardware, and the SAP Business Application Stack. Figure 1-1 shows the platform components, which include the SAP Kernel (light green), the database (blue), and the operating system and hardware (dark blue).

[1] This website is accessible through the SAP intranet or through the internet with a valid SAP user ID.

In the following steps, an example navigation through the PAM for the appropriate components of SAP NetWeaver V7.5 is described:

1. Access the SAP PAM by going to the PAM website.
2. Using the search tool in the upper right corner of the main PAM page, enter “NetWeaver 7.5”. From the resulting search results that are shown in Figure 1-2, click the first search result.

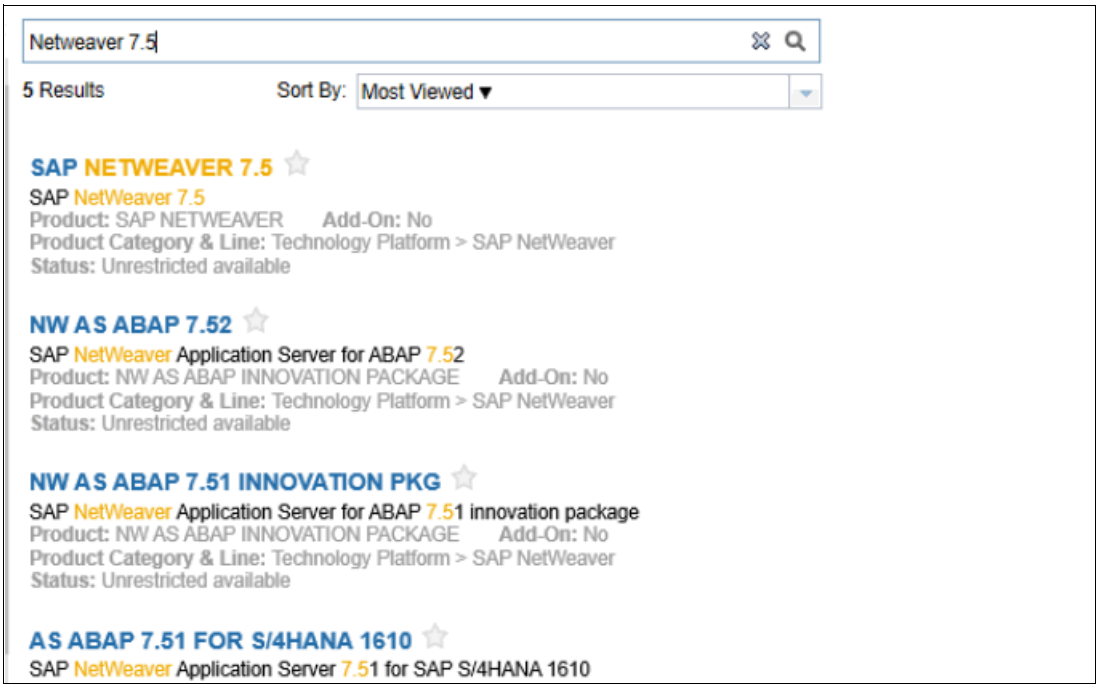

Figure 1-2 Searching for the NetWeaver entry point

Figure 1-3 shows the main page for the NetWeaver V7.5 product.

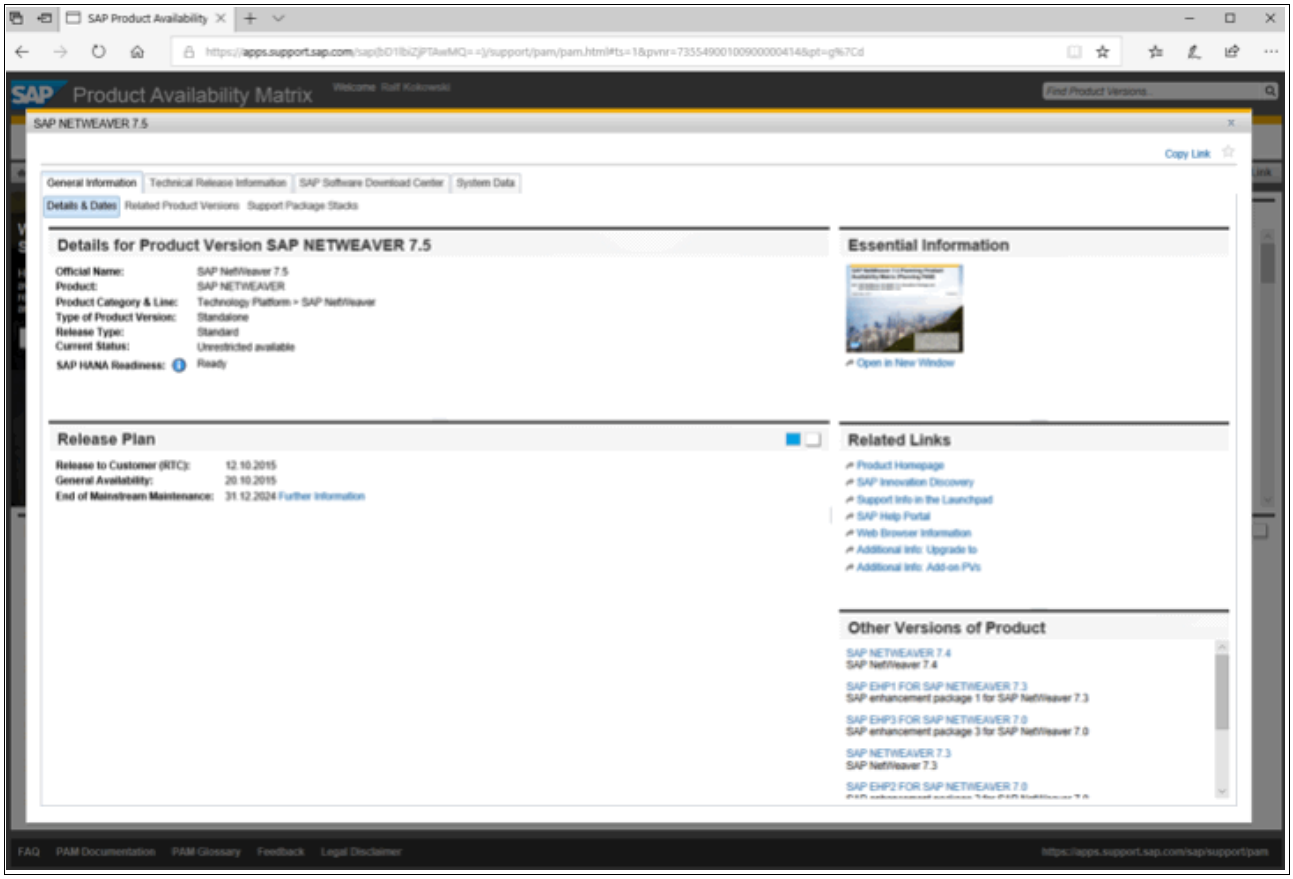

Figure 1-3 SAP Product Availability Matrix: Main page details of SAP NetWeaver

3. To see the possible choices for the generalized platform for NetWeaver 7.5, select the **Technical Release Information** tab, as shown in Figure 1-4. Then, to see more details about this platform, select an option from the Product Instance column.

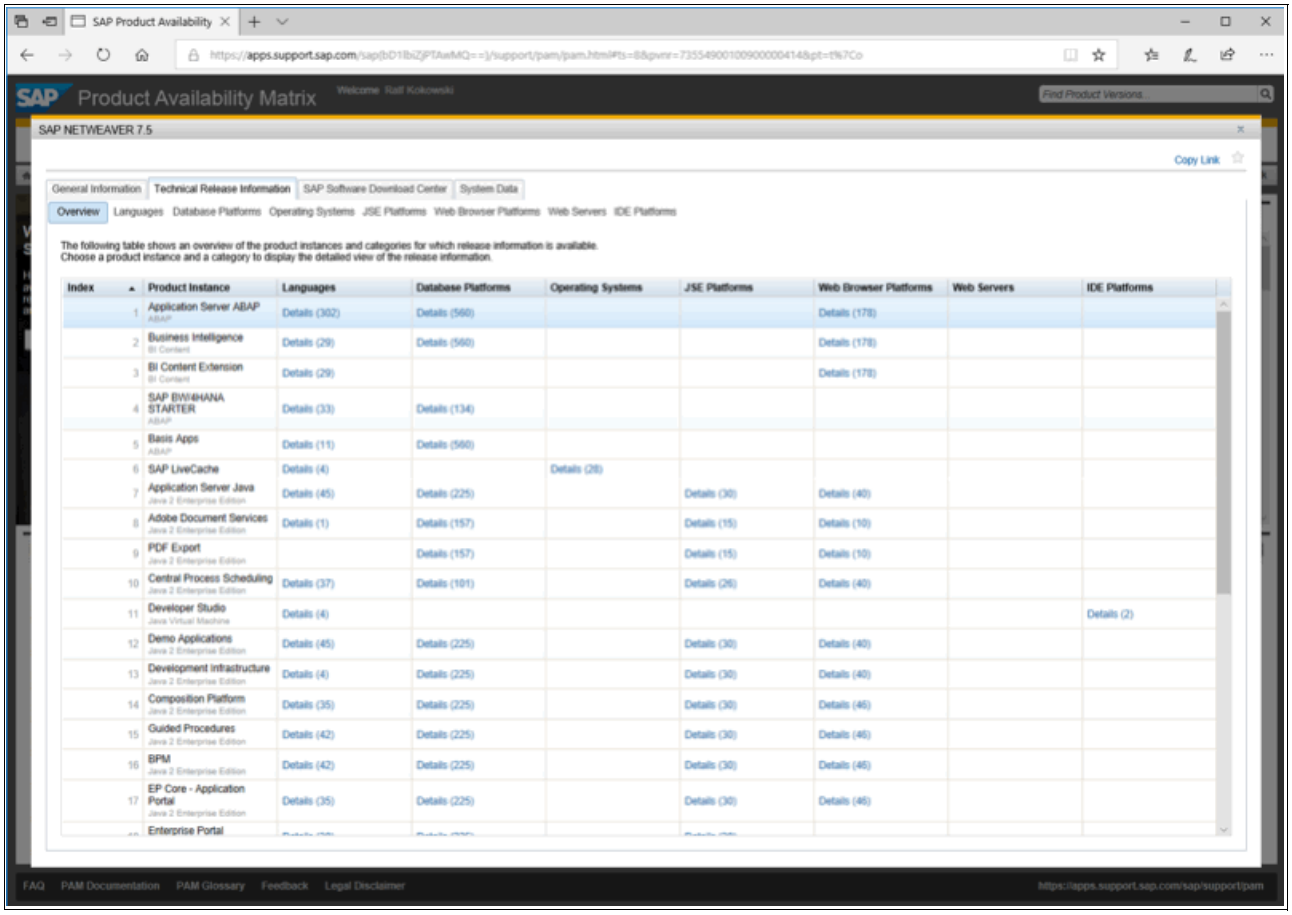

Figure 1-4 SAP Product Availability Matrix: Refining the product selection

4. Select the Application Server ABAP, as shown in Figure 1-4. To see the details of the database platforms for this solution, click **Details** under the Database Platforms column.

   Other columns, such as Languages and Web Browsers, can be interpreted as more dimensions of the overall selection matrix. These cases are not described in this publication.

   The resulting details, as shown in Figure 1-5 on page 7, have the full combination of compatible combinations of the SAP Product Instance: SAP kernel version, database option and versions, and operating system options and versions.

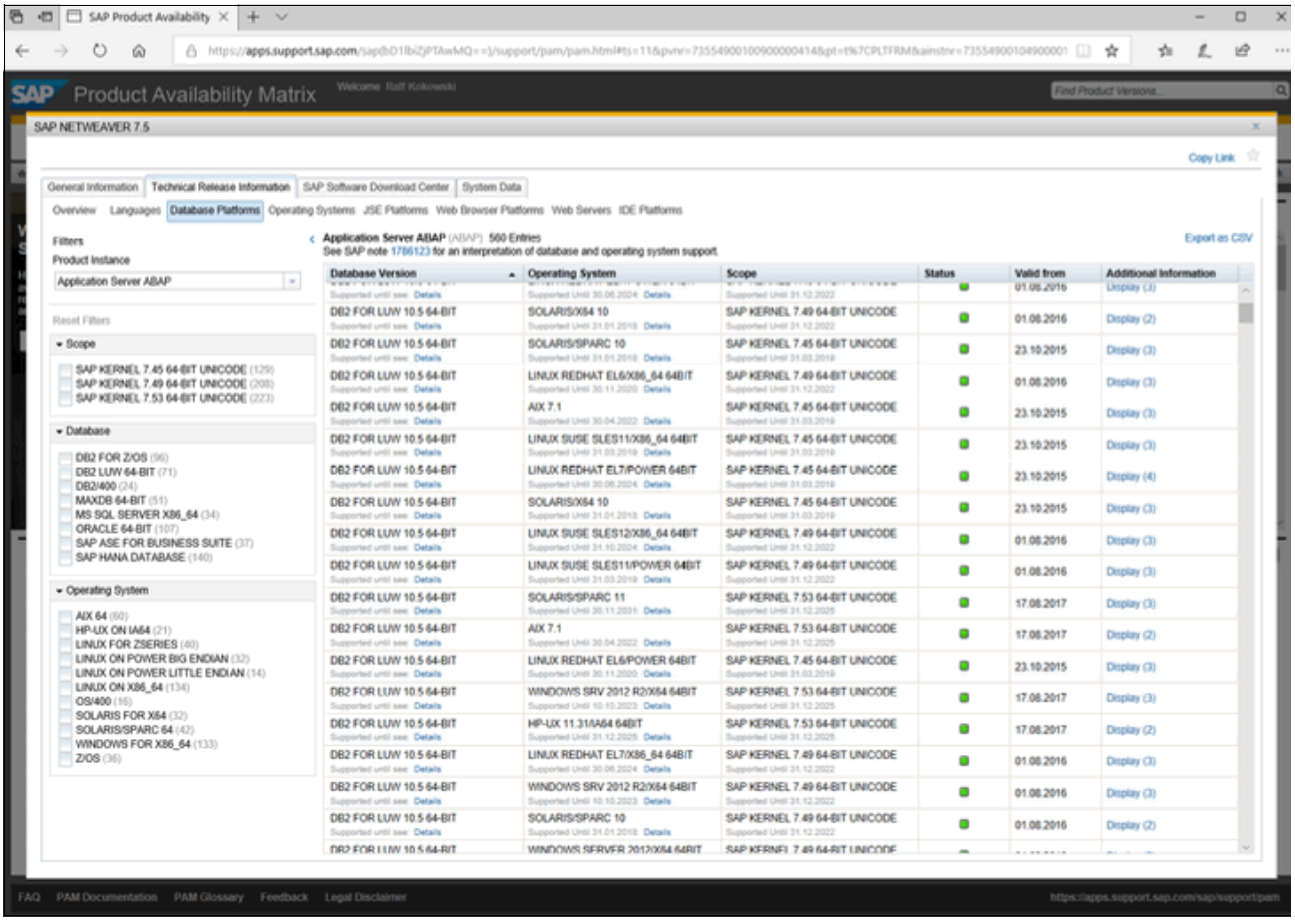

Figure 1-5 SAP Product Availability Matrix: Selecting the database component

The operating system choice also implies a hardware selection. Some more details are denoted in the options description as **LINUX on POWER BIG ENDIAN**, and others are contained implicitly as **AIX 64**, which both run on the Power platform.

From a conceptual view, the SAP products that you see in the PAM are built from two layers: the lower SAP kernel layer and the higher SAP business layer. The SAP kernel layer consists of a set of executable files and tools that at run time provides application server functions. The business layer's objects and entities then run conceptually in this application server environment. This modular approach can later be used by exchanging an SAP kernel with a newer one if you want.

The column at the left of Figure 1-6 allows for further filtering of the results by SAP kernel version, database, and operating system. For example, choosing **SAP Kernel 7.53 64-bit UNICODE** in the Scope section, then **IBM DB2® LUW 64-bit** in the Database section, and then **LINUX ON POWER BIG ENDIAN** in the Operating system section refines the search results to a smaller set of compatible options, as shown in Figure 1-6.

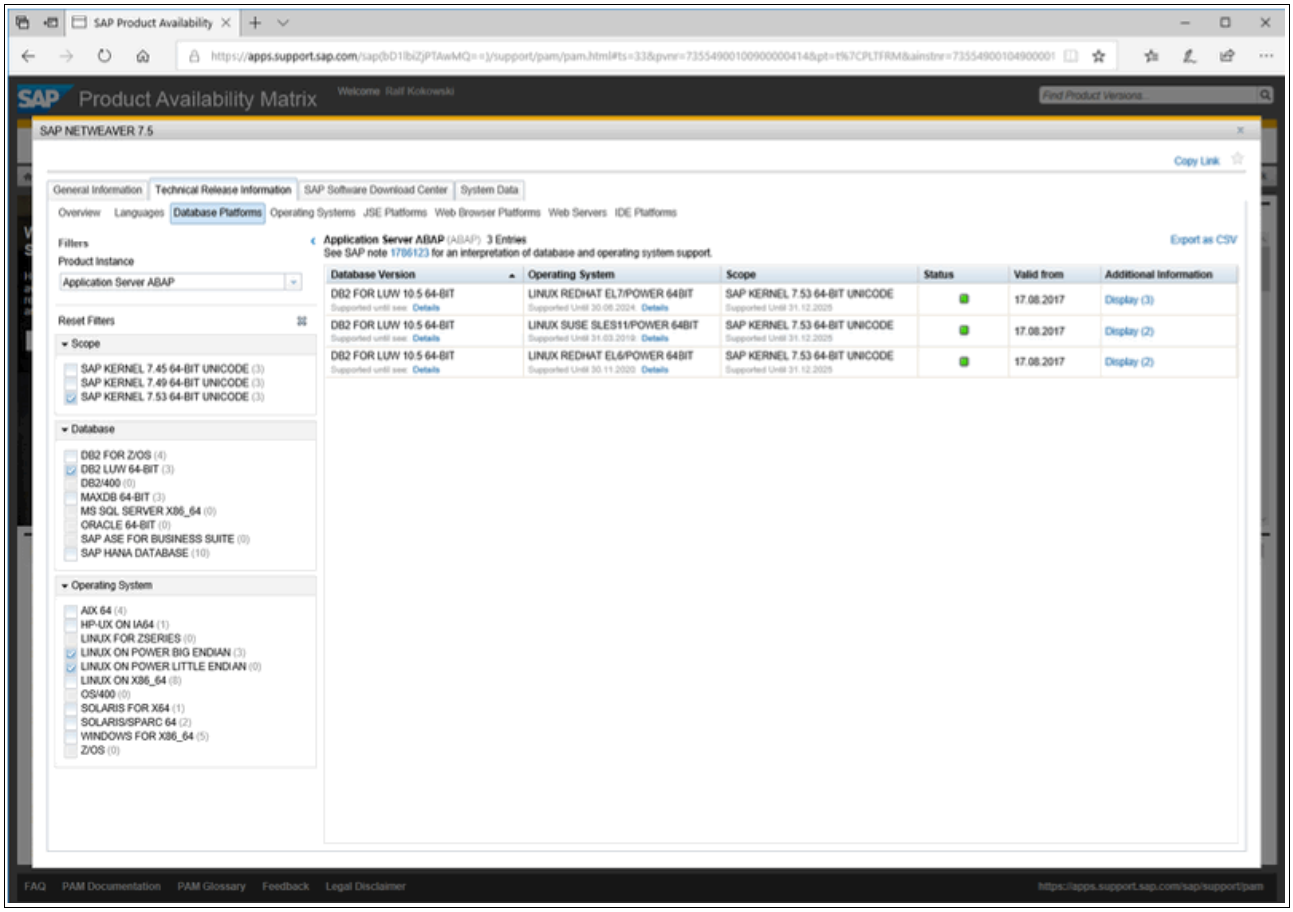

Figure 1-6 Product Availability Matrix: Further refinement of the SAP platform stack options

Not all combinations for the selectable items (SAP kernel, operating system, and database) exist. The PAM restricts the options in the left sections to only those combinations that are available. Components that are excluded are not selectable because they are not validated or available. For example, any option for **LINUX ON POWER LITTLE ENDIAN** is not selectable with **DB2 LUW 64-BIT**, and **LINUX ON POWER BIG ENDIAN** options exist (for example, Red Hat Enterprise Linux V6 and V7, and SUSE Linux Enterprise Server V11).

5. With this final list of compatible options, choose one as the final combination of components that make up the compatible stack for the chosen SAP software solution.

Some extra infrastructural decisions are made outside of the PAM selection process, such as storage technology, backup solutions, using new or existing hardware, and others.

1.4 SAP applications configurations

After a valid configuration of the SAP solution stack is determined, the next step is to understand the resource requirements of the underlying operating system and hardware.

The sizing of the chosen configuration requires a methodology that starts with a top-down approach:

- The specific business processes and the associated requirements must be fully assessed, defined, and outlined.
- From the result of these assessments, the associated application and database resources can be determined, such as the number of concurrent users, processing of parallel business processes, and batch processing requirements, among other resources.
- Then, these specifications can be mapped into SAP application parameters, such as the amount of extended memory for the required business objects (BOBJs) and the number of parallel work processes that are needed to meet the need for the business processes and concurrent users.
- Then, the result of the analysis at the application layer is applied to an appropriate sizing of the database configuration for memory and table space requirements.
- With the resource demands for application and database understood, the hardware and operating system requirements can be determined, which include the appropriate sizing of the hardware compute, memory, and storage resource requirements.

Note: To help in the sizing of each of these layers, SAP and IBM can provide subject matter experts (SMEs) and sizing tools.

For more information, see SAP Benchmark Sizing or contact IBM Techline to get started.

After the initial deployment of the software and hardware solution, as it becomes more mature and integrated into the business, a best practices approach requires a reevaluation of the solution's resource requirements to evaluate whether they are still meeting the business processes as they change and evolve or whether the amount of resources can be reduced. These changes impact the processing capability and capacity of the existing system, and can require a review of the original sizing assumptions and parameters.

The outcome of these periodical reviews can recommend a reconfiguration of the application, database, operating system, and hardware resource requirements. Either more hardware resources (for example, more CPU capacity, more memory, and other resources) are needed, or after business process optimizations the logical partition (LPAR) size can be reduced. New demands can include a need for high availability (HA), a disaster recovery (DR) solution, implementing new solution monitoring and management tools, backups, and others.

There are many solutions that can extend the functions of every SAP installation. Choosing to deploy new components to enhance the existing system requires validation of the interoperability of the component to the installed software stack. These valid combinations can be reviewed in the SAP PAM, SAP solution documentation, by SAP SMEs, or by the vendor of a product (for example, snapshot backup or security software).

1.5 Linux installation

Linux evolved over the years to become an enterprise-level solution with a vast infrastructure.

As an open source project, the Linux operating system is independent of distribution vendors. With a significant development and support community, many supported hardware platforms, and an ever-increasing array of software solutions, the Linux platform provides many features from still not native to enterprise-ready. IBM provides optimal support for Linux on Power Systems servers.

Choosing the correct components and solutions in the Linux infrastructure for example, compatibility, features, functions, support, and lifecycle, requires careful consideration. Most of this task is covered by SAP limiting the choice of supported Linux versions.

1.5.1 The Linux matrix

When you must choose a solution from a collection of many options, it is helpful to build a matrix of these decision points to allow for better visualization and comparison.

For SAP solutions running on Power Systems servers, the following sections illustrate some of the main dimensions and attributes for running Linux:

- Hardware dimension
- Linux distribution dimension
- Linux operating system customization dimension
- Endianness variants

1.5.2 Hardware dimension

Because each hardware platform has different capabilities, the platforms' hardware features can be leveraged by enabling it in the kernel. For the Power platform, some of these features include the following items:

- Running Linux on IBM PowerVM® or KVM Hypervisor or bare metal.
  - In SAP Landscapes, the typical deployment on Power Systems servers is to host many instances of the operating system in virtual machines by installing a hypervisor layer that the Linux operating system runs on to benefit from workload consolidation, and use the Power Systems features that are available on only the E-class models on PowerVM. For a more detailed description of this topic, see *IBM Power Systems Virtualization Operation Management for SAP Applications*, REDP-5579.
  - The small Power Systems servers can run the Linux operating systems where the operating system is installed directly on the hardware (bare metal). Only one version of the Linux operating system can run on the whole machine in this configuration. These servers are typically not used for SAP applications and are often deployed in high-performance computing (HPC) environments.
- Simultaneous Multi-Threading (SMT).
- Live Partition Mobility (LPM).
- An LPAR running on PowerVM can be configured to allow for live migration of the LPAR to another machine to address planned downtime reduction.

- Virtualization or hardware-related command-line tools, such as the **`lparstat`** command in the shell environment, or as the resource dump of the hypervisor, exist to help manage the virtualized environment.
- Reliability, availability, and serviceability (RAS) features, such as firmware-assisted kernel dump versus the usual kernel dump, help facilitate problem determination.

For more information about these and other features and capabilities of Linux on Power Systems servers, see IBM Knowledge Center.

1.5.3 Linux distribution dimension

For SAP solutions, there are two Linux distributions that are available: Red Hat and SUSE Both run on Power Systems servers and provide a special edition for SAP that provides an advanced support model, product lifecycles and features such as automatic tuning, predefined base patterns to ease the installation, and Cluster Manager support. These features are here described below:

- SUSE distributes SUSE Linux Enterprise Server:
  - As of 2019, the supported versions are SUSE Linux Enterprise Server V11, SUSE Linux Enterprise Server V12, and SUSE Linux Enterprise Server V15.
  - Each version further subdivides into a number of service packs, for example, SUSE Linux Enterprise Server V12 SP3.
  - Each service pack can support a different Linux kernel version, which in turn can define specific software solution options that can run on that kernel version. For example, SUSE Linux Enterprise Server V12 SP1 is based on Linux kernel V3.1, and SUSE Linux Enterprise Server V12 SP3 is based on Linux kernel V4.4.

  For more information about SUSE Linux Enterprise Server, see SUSE Linux Enterprise Server.
- Red Hat distributes Red Hat Enterprise Linux:
  - As of 2019, the supported versions are Red Hat Enterprise Linux V6, Red Hat Enterprise Linux V7, and Red Hat Enterprise Linux V8.
  - Each version is further subdivided into subversions, for example, Red Hat Enterprise Linux V7.6.

  For more information about Red Hat Enterprise Linux, see Red Hat Enterprise Linux.
- Both companies include a specific *SAP Edition* that supports SAP tools. The release cycles for it are synchronized to the SAP release cycle.
- Each Linux distribution has unique tools that manage the operating environment:
  - Package distribution methods, such as RPM Package Manager (RPM), Advanced Packaging Tool (APT), and tar.gz.
  - Red Hat and SUSE both use RPM.
- Installation and package update tools:
  - Red Hat provides the gpk-application GUI tool (X Window System), or the Yellowdog Update Manager (YUM) command-line tool.
  - SUSE provides the Yet Another Setup Tool (YaST) graphical tool (X Window System or cursors) and zypper command-line tool.
- Analysis tools, including **`support config`** for SUSE Linux Enterprise Server and **`sosreport`** for Red Hat Enterprise Linux. Also, SAP specific tools are provided by the distributors.

1.5.4 Linux operating system customization dimension

Both SUSE and Red Hat provide a customized version of their enterprise operating system that provides utilities, settings, and features that help with the administration of the system by using SAP specific tools and predefined settings:

- In addition to the standard SUSE Linux Enterprise Server distribution, SUSE provides SUSE Linux Enterprise Server for SAP Applications.

  For more information, see *Installation Quick Start SUSE Linux Enterprise Server for SAP Applications 15 SP1*.

- In addition to the standard Red Hat Enterprise Linux distribution, Red Hat provides the Red Hat Enterprise Linux for SAP Solutions version.

  For more information, see Overview of the Red Hat Enterprise Linux for SAP Applications subscription.

- It is also possible to choose patterns of SAP specific tools from the standard distributions by using the installer utility. These patterns are specific to an SAP product, like NetWeaver or HANA, and automatically installs extra packages.

- There are other operating system tools that are available to further refine the configuration of the operating system for the SAP solution. For example, the **`sapconf`**, **`saptune`** and **`tuned`** commands are available to optimize the operating system for SAP workloads. These commands are described in 1.7, “Linux tuning for SAP workloads” on page 15.

Note: Both vendors provide advanced support and product lifecycles.

1.5.5 Endianness variants

The term *endianness* describes the byte order of basic data types in the operating system. There are two different *endianness* for Power Systems servers: Big Endian and Little Endian. Data in the operating system is represented as bytes, and the different endianness store the data in different order. For example:

- The Big Endian data format stores the data starting with the most significant bit first and the least significant bit last.
- The Little Endian data format stores the data in the opposite order, with the least significant bit first and the most significant bit last.
- IBM POWER8® and POWER9™ processors support both types of endianness, and can simultaneously run LPARs with different endianness.

Note: When starting a Linux deployment, Little Endian is recommended if it is an option.

SUSE Linux Enterprise Server V11 and earlier use the Big Endian format, and SUSE Linux Enterprise Server V12 and later releases use the Little Endian data format.

Red Hat Enterprise Linux V7 and earlier use the Big Endian format, with Red Hat Enterprise Linux V7.3 to V7.6 also supporting Little Endian. Red Hat Enterprise Linux V8 and later releases use the Little Endian data format only.

The choice of the Linux distribution version impacts the endianness and the software solution products that are available to run on that endianness, for example, NetWeaver and HANA.

1.6 Linux customization with SAP base patterns

Linux operating system installation media can be obtained with predefined customization packages.

Note: Most Linux on Power for SAP installations are based on SUSE Linux Enterprise Server. However, Red Hat delivers the same functions and integration, and is equally supported by SAP and IBM.

For more information about the required software packages for Red Hat Enterprise Linux, see Overview of the Red Hat Enterprise Linux for SAP Applications subscription.

In collaboration with SAP and IBM, the Linux distributors made customizations to their operating systems that increase SAP applications performance. These customizations are *base patterns*, and they are customized for many SAP application configurations.

For example, SUSE developed SAP specific patterns. With these SAP patterns, implementing SAP NetWeaver or SAP HANA based on SUSE Linux Enterprise Server or SUSE Linux Enterprise Server for SAP Applications is easier.

The SAP base patterns from different SAP solutions on Linux are available as RPM packages. These packages include all the necessary files and programs that are needed to enhance the configuration, implementation, and management of an SAP solution. Each package has documentation that is included with the components of the package. All the dependent operating system packages are installed with the base pattern RPM.

There are three SAP solutions that are supported by SUSE:

- SAP NetWeaver
- SAP HANA
- SAP NetWeaver + HANA

In each case, customization of the Linux operating system happens by installing one or more base patterns. For each case, here are the suggested base pattern packages to install:

- For SAP NetWeaver, install SAP Application Server Base Pattern (required for any SAP installation), and then install SAP NetWeaver Server Base Pattern.
- For SAP HANA, install SAP Application Server Base Pattern, and then install the SAP HANA Server Base Pattern.
- For SAP NetWeaver + HANA, install all three server base patterns: SAP Application Server Base Pattern, SAP NetWeaver Server Base Pattern, and SAP HANA Server Base Pattern.

Note: This is a generalized approach that can change. For more information about patterns, see your SAP and operating system product installation guides.

1.6.1 SAP Application Server Base Pattern

The base pattern (`patterns-sles-sap_server`) for SAP Application Server Base contains only the package Kernel Parameter Initialization for SAP Systems (`sapconf`). This base pattern can help automatically apply the base tuning and be adjusted to specific needs, and it ensures compliance of the tuning and consistency in the landscape over time.

1.6.2 SAP NetWeaver Server Base Pattern

The base pattern (`patterns-sap-nw`) for SAP NetWeaver Server Base includes the following required operating system packages:

- Virus Scan Adapter (VSA) for ClamAV
- Utility for setting up Linux Unified Key Setup (LUKS)-based disk encryption
- Installation wizard for SAP applications
- Symlink to Firefox
- Integration of SAP Network Interface Router into systemd
- SUSE Linux Enterprise Server for SAP white papers
- Dynamic adaptive system tuning daemon
- Collection of basic system utilities
- Helper daemon to ensure uniqueness of time-based universally unique identifiers (UUIDs)
- YaST Extension for SAP SUSE Connect Program and SAP High Availability

1.6.3 SAP HANA Server Base Pattern

The base pattern (`patterns-sap-hana`) for SAP HANA Server Base includes the following required operating system packages:

- System firewall for SAP HANA database setup
- Resource agents to control the HANA database in system replication setup
- YaST2 Auto Installation Modules
- GNU Command-Line Calculator
- Point-to-Point Bandwidth Measurement Tool
- File System Benchmark
- Vector Graphics Library with Cross-Device Output Support
- VSA for ClamAV
- Utility for setting up LUKS-based disk encryption
- Tool for Automating Interactive Programs
- Graph Visualization Tools
- GTK+ toolkit library (version 2)
- IBM Java Version V7 Release 1 Runtime Environment with JDBC/ODBC bridge driver.
- MIT Kerberos5 implementation - client programs
- Library for C compiler run time, International Components for Unicode, JPEG Support, PNG support, and SSH Support
- Support Utilities for NFS
- Network Time Protocol daemon
- Non-Uniform Memory Access (NUMA) Policy Control
- Installation wizard for SAP applications
- Symlink to Firefox
- Integration of SAP Network Interface Router into systemd
- SUSE Linux Enterprise Server for SAP white papers
- Hardware health monitoring for Linux
- Run some commands as root

- **sar** and **iostat** commands for Linux
- The C Shell
- Utilities for managing the XFS file system
- Open source remote desktop protocol (RDP) server
- YaST Extension for SAP SUSE Connect Program, HANA Firewall, and SAP High Availability

1.7 Linux tuning for SAP workloads

The SAP base patterns on SUSE Linux Enterprise Server ensure that all the necessary operating system packages are installed. For valid, correct, and pre-optimized system tuning, there are different solution approaches between Red Hat and SUSE:

- On Red Hat Enterprise Linux V6.5 and later, the `tuned` service package is available, which contains some specific tuned profiles for SAP solutions.
- On SUSE Linux Enterprise Server and SUSE Linux Enterprise Server for SAP, the package `saptunev2` optimizes SAP solutions depending on the SUSE Linux Enterprise Server version.

1.7.1 The tuned daemon

The `tuned` daemon is an optional Linux operating system daemon that monitors connected devices and tunes system settings for CPU tuning according to a selected profile that is adjusted according to key SAP Notes and other application-specific best practices. The daemon builds the foundation for the automation of Red Hat Enterprise Linux and SUSE Linux Enterprise Server.

The daemon helps to tune the base operating system configuration by monitoring the system and adjusting operating system parameters for changing workload characteristics. The daemon is in an independent RPM package, and it is installed by using the RPM command-line tool (**rpm -i tuned**) or by the operating system package manager zypper or YaST on SUSE Linux Enterprise Server and YUM on Red Hat Enterprise Linux.

SUSE Linux Enterprise Server provides the extra packages `sapconf` and `saptune`, which provide extra input to tuning that is related to the running of SAP workloads.

Since Red Hat Enterprise Linux V6.5, Red Hat provides the RPM packages `tuned-profiles-sap` and `sapconf`, which contain specific tuned profiles for SAP solutions.

The `tuned` daemon can be used for system tuning of the following listed system applications and tools:

- `sysctl`
- `sysfs`
- Various configurations (usually in `/etc`)
- Various tools (**ethtool**, **hdparm**, **taskset**, and others)
- Boot parameters (**elevator**, **nohz**, **isolcpus**, and others)
- Services and systemd units
- `hotplug` (due events processing)

The static tuning mainly consists of predefined `sysctl` and `sysfs` settings and one-time activation of several configuration tools, such as **`ethtool`**. The `tuned` daemon also monitors the use of system components and tunes system settings dynamically based on monitored information. The base package contains several profiles for different workloads.

The `tuned` daemon is based on a plug-in architecture. The options are centralized in factory-based or user-based profiles. The `tuned` package includes rollback and hot-plug support. Figure 1-7 shows the design, and the usage of the configuration profiles and their placement in the system environment.

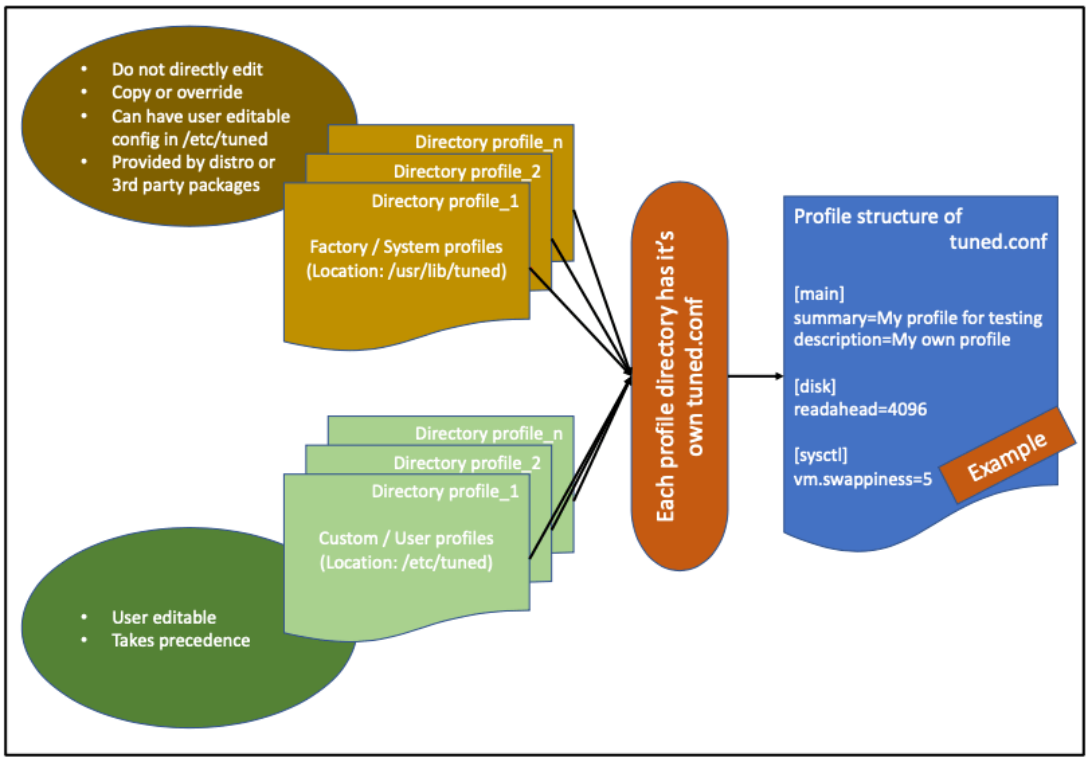

Figure 1-7 Configuration profile

The active profile is defined in `/etc/tuned/active_profile`.

When `saptune` is installed (see 1.7.2, “SUSE Linux Enterprise Server V11 sapconf” on page 17), new profiles are installed, including:

- `sap-hana`
- `sap-netweaver`
- `saptune`

To change the current profile, use the **`tuned-adm`** utility by passing to it the new profile name. The tuning is dynamically applied based on the settings that are defined in the new profile.

1.7.2 SUSE Linux Enterprise Server V11 sapconf

The `sapconf` package is the first tool that is provided by SUSE to tune operating system parameters for SAP solutions in a simple and automated way. The `sapconf` package is available in the base SUSE Linux Enterprise Server distribution and in the SUSE Linux Enterprise Server for SAP Applications distribution. The `sapconf` package works directly with the `tuned` daemon package. The following RPM packages are required:

- `tuned`
- `uuidd`
- `sysstat`

The `tuned` profiles that are provided by `sapconf` are `sap-ase`, `sap-bobj`, `sap-hana`, and `sap-netweaver`. When `sapconf` is installed, these profiles are placed in the Factory / System profile directory (see 1.7.1, “The tuned daemon” on page 15).

Note: Starting with SUSE Linux Enterprise Server V12 SP2, it is highly recommended to use only `saptune` and not `sapconf`.

Starting `sapconf` starts the `tuned` daemon and set ups the required `tuned` profile. The `tuned.conf` file from the `sapconf` profile includes parameter changes that are required for the specific SAP solution. All other tuning parameters are started by **`script.sh`**. The `sapconf` executable file controls the active profile, and can start and stop the service.

For more information about installing, configuring, updating, and the architecture for `sapconf`, see sapconf - A way to prepare a SLES system for SAP workload - Part 1.

1.7.3 The saptune package

With SUSE Linux Enterprise Server V12 SP2, `saptune` V2 is the recommended mechanism to use. `Saptune` is a program package that is offered by SUSE that is available only in SUSE Linux Enterprise Server for SAP Applications. This program package adapts parameter recommendations from SAP Notes or SUSE to specific SAP solutions. There are two versions of `saptune` available: The old Version 1 and the new Version 2. The `saptune` V2 package contains Version 1, which you can switch to by using the general configuration profile. This option allows customers that use Version 1 to use `saptune` with their old settings.

Note: It is mandatory to use `saptune` V2 starting with SUSE Linux Enterprise Server V12 SP2.

The new `saptune` V2 is based on parameter files that contain recommended parameter changes. Each parameter file refers to an SAP Note, vendor-specific tuning definitions, or SUSE recommendation articles. In the following sections, this parameter file is called a *note*. The structure of a note is based on sections. Each section has its own parameters. The section *[version]* is required. The available section, correct syntax, and parameters are documented in the manual page `saptune-note`.

To list all available notes, run the **saptune note list** command, as shown in Example 1-1. The order of enabled notes is 941735, 1771258, 1980196, 2578899, 2684254, 2382421, and 2534844.

Example 1-1 The saptune note list command

```
# saptune note list


All notes (+ denotes manually enabled notes, * denotes notes enabled by solutions,
- denotes notes enabled by solutions but reverted manually later, O denotes
override file exists for note):

        1410736         TCP/IP: setting keepalive interval

                        Version 6 from 13.01.2020

        1680803         Sybase - SAP Adaptive Server Enterprise

                        Version 24 from 20.11.2017

*       1771258         Linux: User and system resource limits

                        Version 5 from 18.06.2018

        1805750         SYB: Usage of HugePages on Linux Systems with Sybase ASE

                        Version 6 from 14.11.2017

*       1980196         Setting Linux Kernel Parameter /proc/sys/vm/max_map_count
on SAP HANA Systems

                        Version 7 from 18.10.2017

        2161991         VMware vSphere configuration guidelines

                        Version 26 from 02.12.2019

*       2382421         Optimizing the Network Configuration on HANA- and OS-Level

                        Version 36 from 16.01.2020

*       2534844         Indexserver Crash During Startup due to Insufficient Shared
Memory Segment

                        Version 12 from 15.11.2017

*       2578899         SUSE LINUX Enterprise Server 15: Installation notes

                        Version 20 from 29.11.2019

*       2684254         SAP HANA DB: Recommended OS settings for SLES 15 / SLES for
SAP Applications 15

                        Version 5 from 03.01.2019
```

```
*       941735          SAP memory management system for 64-bit Linux systems
                        Version 11 from 04.05.2018
         SAP_BOBJ       operating system tuning for SAP Business OBJects (BOBJ)
                        Version 0 from 10.03.2017
The current order of enabled notes is: 941735 1771258 1980196 2578899 2684254
2382421 2534844
```

The `saptune` tool can create, verify, apply, simulate, customize, or revert an SAP Note. Multiple notes are consolidated in a solution. All available solutions by SAP are saved in the configuration file `/usr/share/saptune/solutions`. To get an overview of which kind of solutions are available, which notes are included in a solution, and which of the solutions are active, run the **`saptune solution list`** command, as shown in Figure 1-8. Notes that are active are marked in green.

```
linux:~ # saptune solution list

All solutions (* denotes enabled solution, O denotes override file exists for
solution):
        BOBJ              - 941735 1771258 1984787 SAP_BOBJ
        HANA              - 941735 1771258 1980196 1984787 2205917 2382421 2534844
        MAXDB             - 941735 1771258 1984787
 *      NETWEAVER         - 941735 1771258 1984787
        NETWEAVER+HANA    - 941735 1771258 1980196 1984787 2205917 2382421 2534844
        S4HANA-APP+DB     - 941735 1771258 1980196 1984787 2205917 2382421 2534844
        S4HANA-APPSERVER  - 941735 1771258 1984787
        S4HANA-DBSERVER   - 941735 1771258 1980196 1984787 2205917 2382421 2534844
        SAP-ASE           - 941735 1410736 1680803 1771258 1984787
linux:~ #
```

Figure 1-8 Output of a saptune solution list

Operations like apply, verify, simulate, and revert for the installed solutions can be run by running **`saptune`**.

To activate and view the overall status of `saptune`, there are several options available. To view the status of `saptune`, run **`saptune daemon status`**, as shown in Figure 1-9.

```
linux:~ # saptune daemon status
Daemon (tuned.service) is running.
The system has been tuned for the following solutions and notes:
        NETWEAVER
linux:~ #
```

Figure 1-9 Command saptune daemon status

The **`saptune daemon start`** command starts `saptune`. If no solution or note is selected, there is a warning to apply one of the available notes and solutions. The **`saptune revert all`** command removes all parameter changes and resets the operating system to the defaults.

The `saptune` tool is enabled by the `tuned` daemon at operating system start. Figure 1-10 shows the file and directory structure of the `saptune` environment, including the file structure of the `tuned` daemon.

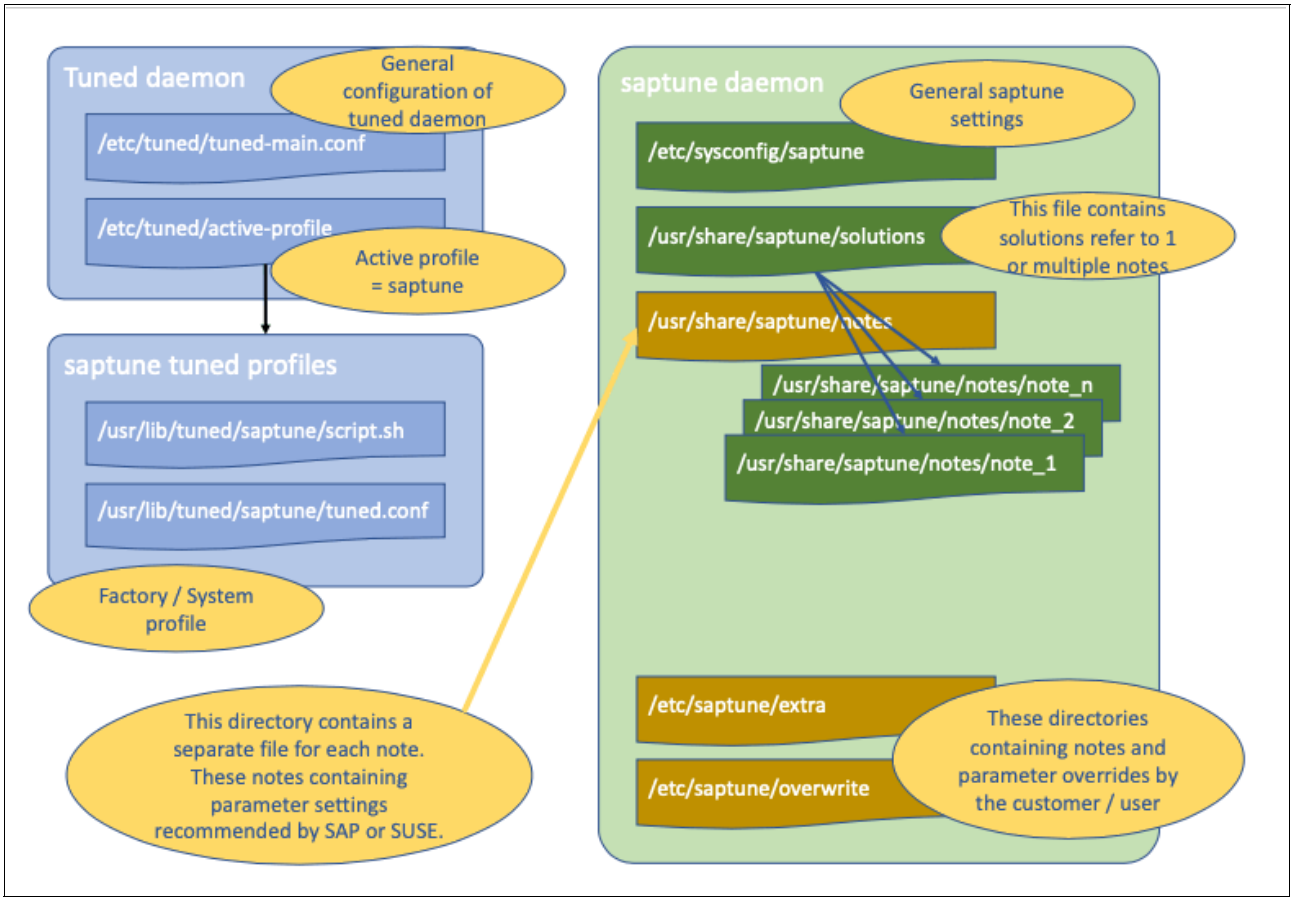

Figure 1-10 Directory structure of saptune

The `saptune` daemon handles the system settings, which are defined in the note files without using the `tuned` daemon. The rules for how to configure the note files are defined in the man page for `saptune`.

1.8 Hardware aspects

When you plan a deployment of SAP applications, it is a best practice to plan your hardware initially to avoid scheduling unplanned downtimes to adjust configurations. Use smaller tools to determine whether the setup meets the expected performance to catch configuration mistakes before an application is in production. This section helps you to understand some of the important hardware aspects that you can encounter while planning deployments of SAP HANA or any other SAP application on Power Systems servers.

1.8.1 NUMA

NUMA is used in IBM Power Architecture® platforms. In a system that has NUMA characteristics, each processor has local memory that is available, but it can access memory that is assigned to other processors. The memory access time is faster for local memory. A NUMA node is a collection of processors and memory that is mutually close. Memory access times within a node are faster than outside of a node.

Power Architecture maps memory by locality (core, chip, dual chip module/ socket, node, and others). The memory affinity process locality can impact performance on any system.

With IBM POWER9 processors, the interconnect bandwidth on each node and across multiple nodes is improved by 4x compared to IBM POWER8 processors. This situation results in the improvement of throughput and reduction of latency for SAP applications running on IBM POWER9 processor-based systems compared to their IBM POWER8 predecessors.

Viewing the NUMA topology from inside the operating system

Looking at the NUMA topology from inside a Linux operating system requires knowledge about what information is updated because the LPAR placement can change after a Live Partition Mobility (LPM) operation or similar tasks. The command that is used is updated only after a restart, so it might display outdated information. For the current information, you must do a memory dump on the HMC, as described in “Viewing the LPAR placement from the HMC” on page 23.

Log in to the LPAR and ensure that the **`numactl --hardware`** command has a symmetric output. Confirm that every NUMA node with a core holds memory too. Nodes without cores might not hold memory too, which is often the case for NUMA node 0 not having cores or memory assigned. In such a case, SAP HANA internally maps all NUMA nodes and works only with nodes with both CPU and memory. Therefore, it has no impact on HANA.

Background on empty NUMA node0

NUMA node 0 historically comes from a bare metal installation where it was part of a sequence of NUMA nodes starting with 0. Later in first virtualized environments, this designation was kept as a constant to represent an anchor point. Today, an empty NUMA node 0 is a legacy item. In current implementations, troubleshooting an empty node0 is irrelevant because it has no functions and no impact because HANA is fully aware of this concept.

Example 1-2 shows that NUMA node 0 has no cores or memory that is associated with it. SAP HANA detected it and listed a number of NUMA nodes with logical CPU, allowed memory, and both CPUs and Memory as 1, ignoring NUMA node 0.

Example 1-2 Empty NUMA node representation

```
hd1adm@LINUXLPAR:/usr/sap/HD1/HDB00> numactl --hardware
available: 2 nodes (0,6)
node 0 cpus:
node 0 size: 0 MB
node 0 free: 0 MB
node 6 cpus: 0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15
node 6 size: 245201 MB
node 6 free: 125343 MB
node distances:
node   0   6
  0:  10  40
```

```
  6:  40  10

hd1adm@LINUXLPAR:/usr/sap/HD1/HDB00> hdbcons "mm numa -t"
## Start command at: 2020-03-20 11:57:37.796
************************************************************
               Configuration of NUMA topology
************************************************************
     Is NUMA-awareness enabled? (0|1) :    1
           Valid NUMA Topology? (0|1) :    1
                 Number of NUMA Nodes :    2
Number of NUMA Nodes with logical cpus :    1
Number of NUMA Nodes with allowed memory:    1
Number of NUMA Nodes with both cpus and memory:    1
              Number of Logical cpus  :   16
                    Cpu-only node IDs : NONE
                    Mem-only node IDs : NONE
************************************************************
[OK]
## Finish command at: 2020-03-20 11:57:37.796 command took: 112.000 usec
```

Similarly, you can also check other **hdbcons** options such as **"jexec info"**, which lists an active NUMA node as 1. In this case, as shown in Example 1-3, the command ignores a NUMA node with no memory and no CPU.

Example 1-3 The hdbcons command

```
hd1adm@LINUXLPAR:/usr/sap/HD1/HDB00> hdbcons "jexec info"
## Start command at: 2020-03-20 11:58:34.289
Using 2 numa nodes
SMT level: 8 using 2 physical cores
numa_features: 1, config: -1
bind_workers: 1, config: -1
max_concurrency: 16 (cfg=, dyn=17)
max_concurrency_hint: 17 (cfg=0)
min_concurrency_hint: 4 (cfg=0)
concurrency_policy: 2 (cfg=0)
max_concurrency_min_pct: 30
max_concurrency_hint_min_pct: 50
stealing_policy: 11 (cfg=0)
0 statement limiters
System info:
2 possible NUMA nodes, 40 possible cores, 1 active NUMA nodes, 16 active logical
cores
Using global restriction to a subset of cores: [11111111 11111111 00000000
00000000 00000000 ]
Numa node [0], Socket ID [0]: usable cores=0, available memory=0 KB
  has 1 neighbors: 1
  max_concurrency: 1, dyn=1
Numa node [1], Socket ID [6]: usable cores=16, available memory=251086144 KB
  has 1 neighbors: 0
  max_concurrency: 16, dyn=16
current memory usage, operative: 14560, background: 652912
[OK]
## Finish command at: 2020-03-20 11:58:34.289 command took: 17.000 usec
```

Figure 1-11 shows that the LPAR has an uneven NUMA layout. NUMA node 0 has cores but no memory. NUMA node 6 holds all the memory but no cores, which degrades the performance of SAP HANA (more on POWER8 than POWER9 processors).

In this type of setup, when you run your HANA workload, the processor and memory are not on the same NUMA node, which leads to more memory fetches and higher latency. This situation can happen when your Power Systems server is already running multiple LPARs, and newly created LPARs can access only the remaining resources. The hypervisor tries to allocate resources, but fails to provide an ideal setup due to resource constraints. It is always a best practice to get the best allocation, but in a few cases it is not possible.

For better performance, place the memory and processor cores on the same NUMA node. In multi-cores setup, you have multiple NUMA nodes with some memory and some cores, where:

- Access to memory that is in the same node (local memory) is direct with low latency.
- Access to memory that is in another node is achieved through the interlink bus with a higher latency.

```
lsh30021:~ # numactl --hardware


available: 2 nodes (0,6)
node 0 cpus: 0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29
30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60
61 62 63
node 0 size: 0 MB
node 0 free: 0 MB
node 6 cpus:
node 6 size: 32768 MB
node 6 free: 31649 MB
node distances:
node   0   6
  0:  10  40
  6:  40  10
```

Figure 1-11 Worst case NUMA distribution

Viewing the LPAR placement from the HMC

This option always delivers the current information. To create a resource dump, complete the following steps:

1. Open an SSH to the HMC console and go to the command-line login.
2. On the HMC, check which resource dumps are already in the dump directory by running the following command:

   ```
   hscroot@:~> lsdump -h
   dump_type=resource,name=RSCDUMP.109130D.09000001.20140624
   065824,size=2490992,source_size=0
   dump_type=resource,name=RSCDUMP.109130D.0B000001.20141114
   162422,size=16128,source_size=0
   ```

3. Run the following **startdump** command:

   ```
   startdump -m -t resource -r 'hvlpconfigdata -affinity -domain'
   ```

   Note: Creating the dump can take a few seconds.

4. Check whether the dump was created by running the **ls –ltr /dump** or **lsdump** command:

   ```
   hscroot@:~> lsdump -h dump_type=resource,name=RSCDUMP.,size=12784,source_size=0
   dump_type=resource,name=RSCDUMP.,source_size=0
   dump_type=resource,name=RSCDUMP.,source_size=0
   ```

   This dump contains binary and human readable data.

5. Identify the dump by running the following command:

   ```
   cat /dump/RSCDUMP.<my dump ID> | more
   ```

   This command displays a set of data. In the table that is shown, verify that none of the LPAR IDs belonging to a HANA LPAR are spanning the drawers, or download the table to another machine by using the following command:

   ```
   scp hscroot@:/dump/RSCDUMP.<my dump ID>.
   ```

6. After you download and check the dump file, delete it on the HMC by running the following command:

   ```
   hscroot@:~> rmdump -f RSCDUMP.<my Dump ID>
   ```

Reading the HMC resource dump

The resource dump contains binary data and human readable tables. Figure 1-12 shows the key items to look at, especially LPAR 6 (green) and LPAR 8 (red).

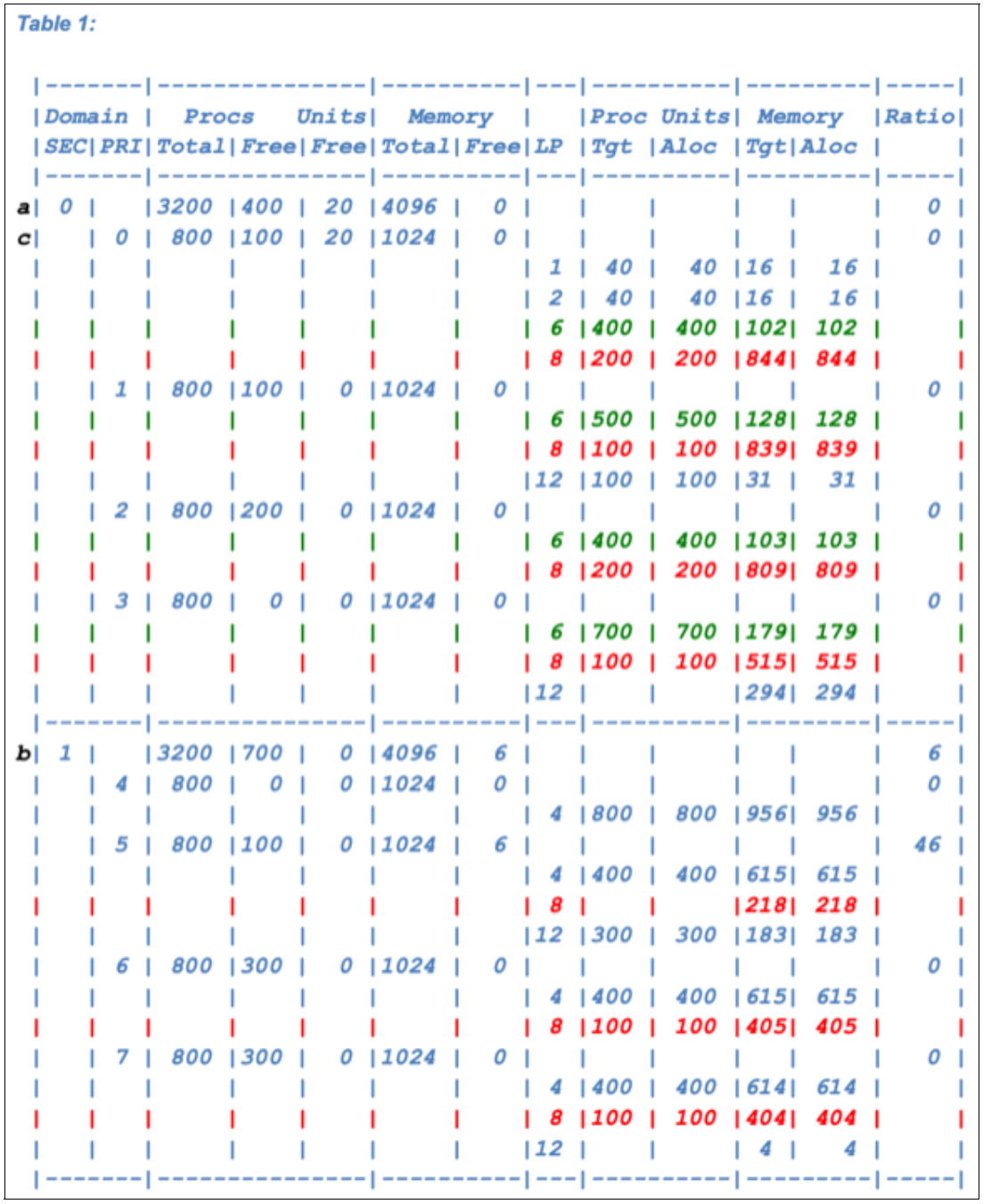

Table 1:

	Domain		Procs		Units	Memory			Proc Units		Memory		Ratio
	SEC	PRI	Total	Free	Free	Total	Free	LP	Tgt	Aloc	Tgt	Aloc	
a	0		3200	400	20	4096	0						0
c		0	800	100	20	1024	0						0
								1	40	40	16	16	
								2	40	40	16	16	
								6	400	400	102	102	
								8	200	200	844	844	
		1	800	100	0	1024	0						0
								6	500	500	128	128	
								8	100	100	839	839	
								12	100	100	31	31	
		2	800	200	0	1024	0						0
								6	400	400	103	103	
								8	200	200	809	809	
		3	800	0	0	1024	0						0
								6	700	700	179	179	
								8	100	100	515	515	
								12			294	294	
b	1		3200	700	0	4096	6						6
		4	800	0	0	1024	0						0
								4	800	800	956	956	
		5	800	100	0	1024	6						46
								4	400	400	615	615	
								8			218	218	
								12	300	300	183	183	
		6	800	300	0	1024	0						0
								4	400	400	615	615	
								8	100	100	405	405	
		7	800	300	0	1024	0						0
								4	400	400	614	614	
								8	100	100	404	404	
								12			4	4	

Figure 1-12 Resource dump

To understand the table, review the following list:

- Each enclosure is symmetrically equipped:
  - The *Domain* column describes the physical topology.
  - *SEC* describes the enclosures (this server has two enclosures).
  - *PRI* describes the sockets (this server has eight sockets with 4 in each enclosure).
  - The rows a) and b) describe the configuration of each enclosure:
    - Each enclosure has 32 cores (Column Procs/Total - 3200 Units) and 1 TB of memory. The memory is listed in 256 MB Blocks → Column Memory Total 4096 Blocks = 1 TB memory.
    - Similarly, the data for each socket, for example, row c): 8 cores, 256 GB of memory (1024 blocks).
- LPAR placement:
  - LPAR ID 6 - green:

    In Figure 1-12 on page 25 (Table1), the LPAR is spread across four sockets in enclosure 0 (see green lines). The LP column describes the LPAR ID (in this case, 6). The reason that this LPAR is divided across the four sockets is that it does not fit into one socket (it uses 20 cores, which are more than what one socket contains). Therefore, the hypervisor keeps the LPAR in one enclosure and distributes it across all sockets in this enclosure. This is the best solution for this case.

    If the LPAR is smaller than eight cores and 256 GB of memory and there is a socket that can provide those resources, the hypervisor places the LPAR into one socket.
  - LPAR ID 8 - red:

    This LPAR was created as the last LPAR and has an inconvenient size (eight cores and ~1 TB of memory). When the LPAR was created, the hypervisor used the rest of the available resources on the machine, and as a result the LPAR was scattered across seven cores and two enclosures, and in one case got memory from one socket but no cores from this socket (socket 5).

    The relevant lines are marked red in Figure 1-12 on page 25 (Table 1).

    Therefore, the LPAR has not shown good performance characteristic, although there is local, near and far memory access. This LPAR can heavily benefit from a better placement on the machine.

Analyzing and optimizing NUMA placement

In the newer versions of the firmware, the hypervisor team provided methods to analyze and fix memory placement issues on the SSH shell of the HMC. This process is not apparent to the applications and can be run while every LPAR is running if there is available memory on the machine and this feature is enabled. To list the servers, run the following command:

```
lssyscfg -r sys -F name
```

To analyze the current situation, run the **lsmemopt** command, as shown in Figure 1-13.

```
hscroot@<ip-hmc>:~> lsmemopt -m <Power Server Name> -r lpar -o currscore

lpar_name=<name of LPAR>,lpar_id=1,curr_lpar_score=100
lpar_name=<name of LPAR>,lpar_id=2,curr_lpar_score=100
lpar_name=<name of LPAR>,lpar_id=3,curr_lpar_score=none
lpar_name<name of LPAR>,lpar_id=4,curr_lpar_score=100
lpar_name=<name of LPAR>,lpar_id=5,curr_lpar_score=none
lpar_name=<name of optimal LPAR>,lpar_id=6,curr_lpar_score=100
lpar_name=<name of improvable LPAR>,lpar_id=8,curr_lpar_score=74
lpar_name=<name of LPAR>,lpar_id=12,curr_lpar_score=32
lpar_name=<name of LPAR>,lpar_id=31,curr_lpar_score=none
```

Figure 1-13 The lsmemopt command

The command does not show the exact placement of memory and cores. Instead, it does a rating, where 100 is the best and 0 is the worst. It rated LPAR 6 with a 100, which means the placement cannot be improved according to the rules of the hypervisor. Figure 1-13 shows that LPAR 8 does not have a perfect rating.

Running the command as shown in Figure 1-13 with the option **-o calcscore** shows to what degree the Dynamic Platform Optimizer (DPO) can optimize the LPAR based on the current situation.

To use DPO to optimize the LPARs, run the following command:

```
optmem -m <Power Server Name> -o start -t affinity -p <name(s) of improvable
LPAR(s) >
```

All other LPARs are candidates to be changed to achieve the best placement. If an LPAR is not touched, another option can be specified (see the man pages for **optmem**).

Although the command is running in the background, you can check the optimization status by running the following command:

```
lsmemopt -m <Power Server Name>
```

Analyzing and fixing the memory placement manually

If DPO does not optimize the memory placement to the required extent or the feature is not enabled, there are also some options to fix the issues manually.

When the hypervisor creates an LPAR, it assigns the cores (entitlement) and memory (DIMMs) to the LPAR, and this assignment is fixed, even when the LPAR is later restarted.

The idea behind the solution is to remove the fixed assignment of the resources and wait for the hypervisor to choose a better placement when the LPAR is activated again.

This process works only if you changed the profile of the LPAR to better fit the available resources, or if other LPARs that are using the needed resources changed.

To get the memory assignment for the LPAR, run the **lshwres** command, as shown in Figure 1-14.

```
hscroot@<hmc>:~> lshwres -m <Power Server Name> -r mem  --level lpar -F
lpar_name,lpar_id,curr_mem

<LPAR name>,12,131072
<LPAR name>,31,0
<LPAR name>,6,131072
<LPAR name>,5,0
<LPAR name to be fixed>,4,716800
<LPAR name>,3,0
<LPAR name>,2,4096
<LPAR name>,1,4096
<LPAR name>,8,1032704
```

Figure 1-14 The lshwres command

To remove the fixed memory assignment, run the **chhwres** command on the HMC:

```
chhwres -m <Power Server Name> -r mem -o r –id 4 -q 716800
```

This command removes the memory assignment of LPAR (ID = 4). The **-q** option is the quantity to remove, which is the **curr_mem** from the previous list.

To fix the memory placement, complete the following steps:

1. Shut down every LPAR that must be changed.
2. Get the correct memory settings by running the **lshwres** command (see Figure 1-14).
3. Remove the memory assignment by running the **chhwres** command.
4. Start the LPARs in the correct priority order and wait for the hypervisor to perform a better placement.
5. Check the new placement again with the resource dump.

Best practices for LPAR placement

There are best practices to apply when creating an LPAR on a Power Systems server:

- Check the `lpar_curr_score` value by running the **lsmemopt** command to verify how the LPAR is placed regarding CPU and memory.
- When running the **optmem** command, this task must be coordinated with the application owner because some applications such as HANA do not adapt automatically.
- On new servers, always create or import the most relevant LPARs first to give them the best placement.
- On some servers, the **optmem** command does not exist. In this case, the LPAR shutdowns and smart restarts can be used to give higher priority to the important LPARs.

For more information about SAP HANA best practices on LPAR placement, see the *SAP HANA on IBM Power Systems and IBM System Storage Advanced Operations Guide*.

1.8.2 I/O configuration and tuning

When a new SAP system is set up, the I/O configurations, storage, and network need attention. The performance of an I/O configuration is defined by its bandwidth (throughput when using large packages) and latency (when using many small packages). Both storage and network configuration performance are determined by these attributes. Users have the choice to use either virtual or dedicated storage and network adapters. To measure the bandwidth and latency of storage or network, there are multiple tools that are available from SAP, IBM, or other sources.

This section focuses mainly on the tools that measure the storage or network bandwidth of an LPAR.

Note: For more information about the aspects of storage and network I/O configuration including latency, see *IBM Power Systems Infrastructure I/O for SAP Applications*, REDP-5581.

Network bandwidth test

Network is always measured by the speed at which data can be transferred between two hosts over the Internet Protocol (IP). Each host is recognized with an IP address, and one host becomes a host and the other becomes a receiver. This setup is like a client/server architecture. Most of the network measurement tools give results in Mbps or Gbps.

Niping

`niping` is an SAP tool to measure the line speed of a network device. For more information about **`niping`**, how to use it, and a sample output, see the following resources:

- Using “niping” for Troubleshooting Network Issues
- SAP Note 500235 - Network Diagnosis with NIPING

iperf

`iperf` is another open source tool that measures network bandwidth. This tool easily checks network bandwidth between a client and a server. For more information about this tool, see the following resources:

- Changes between iPerf 2.0, iPerf 3.0 and iPerf 3.1
- What is iPerf / IPerf3?

When you run **`iperf`** against AIX, validate that you have the same **`iperf`** version. The results between iperf2 and iperf3 cannot be compared one to one.

Storage I/O bandwidth test

A storage I/O bandwidth test or line speed test measures the throughput that a storage adapter gives when it attached directly to a system or LPAR. The test measures the following aspects:

- *Input/output operations per second* (IOPS) is a measurement of performance for hard disk drives (HDDs) and storage area networks. IOPS represents how quickly a storage device or medium can read and write commands every second.
- *Throughput* measures the data transfer rate to and from the storage media in megabytes per second (MBps). Although your bandwidth is the measurement of the total possible speed of data movement along the network, throughput can be affected by IOPS and packet size.

- *Latency* is how fast an I/O operation can be completed after submitted to a storage subsystem.

fsperf

fsperf is a tool to measure storage throughput and latency through a portion of the SAP HANA I/O layer. It is part of the SAP HANA HW Configuration Check Tool (HWCCT) bundle, which is downloadable from the SAP Service Market. For our purposes, the match to a specific HANA version is not important. Although clients can use it as intended by configuring a JSON file, it also can be run alone, which allows for more selective test modes and a faster return. For a sample call of **fsperf**, see Appendix A, “The fsperf tool” on page 43.

dd commands

The **dd** command is a command-line utility that is available on Linux. Its primary purpose is to convert and copy files according to the operands that are passed to the command. It is a lightweight command that can be used to copy files in any format. It also can be used to measure the line speed and throughput of a storage device.

To measure the throughput of a storage device, run the following **dd** command:

```
dd if=<input_file> of=<output file > bs=<block_size> count=<count> oflag=dsync
```

The command requires the following items, as shown in Figure 1-15:

- An input file, which is the test file with which to test the throughput.
- An output file, which is the location of the disk or storage device under test.
- `block_size`, which is the size of the block file to be written.
- `count`, which is how many blocks to be copied.
- `oflag=dysnc` to make the **dd** command use synchronized I/O for data.

```
> dd if=/dev/zero of=/hana/data/test bs=1G count=10  oflag=dsync
10+0 records in
10+0 records out
10737418240 bytes (11 GB, 10 GiB) copied, 7.85442 s, 1.4 GB/s
```

Figure 1-15 Output of the dd command

Figure 1-15 shows that a throughput of 1.4 GBps was measured, which around 9.6 Gbps. A block size of 1 GB was used. The **dd** command was called with `count=10`, meaning it writes 10 blocks of 1 GB each on output file `/hana/data/test`.

2

Monitoring

This chapter describes the SAP monitoring components for measuring the performance of the underlying operating system.

This chapter includes the following sections:

- SAP monitoring components
- Operating system-specific monitoring tools on Linux on Power
- Differences in SAP monitoring between hypervisors PowerVM and RHV-KVM

2.1 SAP monitoring components

This section gives a quick overview of the SAP monitoring components for measuring the performance of the underlying operating system and its components, like CPU, memory, network, and the I/O subsystem. Most SAP solutions communicate with the SAP Host Agent to gather performance matrixes from the operating system. The SAP Host Agent is the focal point for operating system monitoring, but also it is the focal point for managing and controlling SAP solutions. Depending on the SAP solution, there are several ways to present the operating system performance metrics. Figure 2-1 shows a snapshot of an SAP NetWeaver ABAP Server that uses transaction ST06.

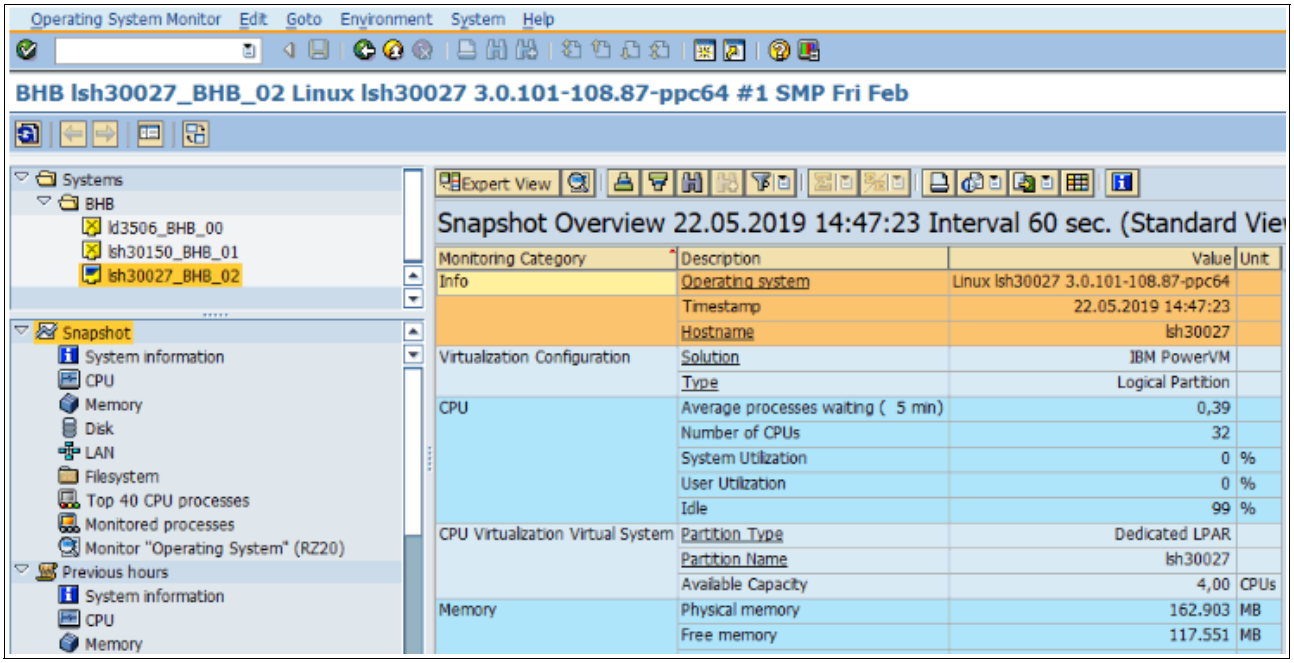

Figure 2-1 Snapshot: NetWeaver

The ST06 transaction displays the operating system statistics for all SAP instances and other monitored systems and hosts, and can be called from any SAP instance. ST06 uses information from SAPOSCOL, the SAP operating system collection agent that was adapted and maintained to reflect the different options for IBM Power Systems virtualization and sharing.

2.1.1 SAP Host Agent

The SAP Host Agent is a service agent that is used for several SAP lifecycle management tasks, such as:

- Operating system monitoring
- Database monitoring
- System instance control and management (discovery and inventory functions)
- Provisioning

The SAP Host Agent is automatically installed during the deployment of new SAP solutions, such as SAP NetWeaver or SAP HANA. However, it is possible to install the SAP Host Agent independently from any SAP solution. One option is to install it within the Virtual I/O Server (VIOS) to get an end-to-end monitoring view of the entire Power Systems server. The architectural picture (Figure 2-2) shows the components of the SAP Host Agent.

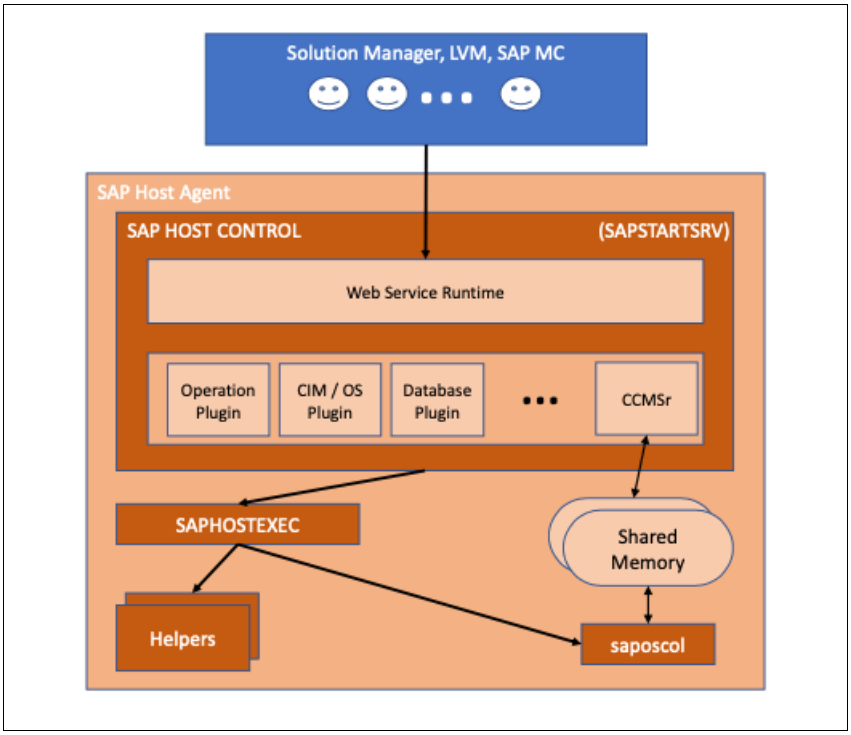

Figure 2-2 The SAP Host Agent

SAPOSCOL

SAPOSCOL is a stand-alone program that runs as a background process in the operating system. It runs independently of SAP instances exactly once per monitored host. SAPOSCOL collects metrics of operating system resources. The following measurements are included:

- Usage of virtual and physical memory
- CPU utilization
- Utilization of physical disks and file systems
- Resource usage of running processes

SAPOSCOL makes this data available to other applications and all SAP instances on a host through a shared memory segment.

Common Information Model client (SAPCIMC)

The Common Information Model (CIM) is an open standard that defines how managed elements in an IT environment are represented as a common set of objects and relationships among them. The Distributed Management Task Force (DMTF) maintains the CIM to allow consistent management of these managed elements independently of their manufacturer or provider. The SAP Host Agent contains a stand-alone CIM client (`sapcimc`). The following classes are provided by CIM and the operating system plug-in specifically for Linux on Power (as of 2019):

- `CPUVirtCIMhost`
- `CPUVirtCIMhost_HardwareId`
- `CPUVirtCIMhost_Model`

- CPUVirtCIMhost_PhysCPUIdle
- CPUVirtCIMhost_PoolId
- CPUVirtCIMhost_PoolUtilAuth
- CPUVirtCIMhost_Processor
- CPUVirtCIMvm
- CPUVirtCIMvm_AvailCapacity
- CPUVirtCIMvm_AvailCapConsumd
- CPUVirtCIMvm_CapaAvailAdd
- CPUVirtCIMvm_CapaConsumed
- CPUVirtCIMvm_CapaGuarant
- CPUVirtCIMvm_Capped
- CPUVirtCIMvm_EntitlemConsumd
- CPUVirtCIMvm_Entitlement
- CPUVirtCIMvm_OperatingFrequ
- CPUVirtCIMvm_PartitionName
- CPUVirtCIMvm_PartitionType
- CPUVirtCIMvm_SMTmode
- CPUVirtCIMvm_SMTthreads
- CPUVirtCIMvm_VirtualCPUs
- CPUVirtCIMvm_Weight

The CIM and operating system plug-in is used by the SAP Solution Manager, SAP Landscape Management, and other components.

Command-line client (saphostctrl)

The command-line tool **saphostctrl** is the interface to run different operations in the SAP application that is called *Webmethods*. A Webmethod can have different options, depending on its function. By running the **saphostctrl** command, it is possible to start, stop, manage, verify, observe, and control an SAP application instance on the same or different host system. On all hosts that are running an SAP instance, an SAP Host Agent must be active. The following command example shows how to extract capacity data from the local system:

```
linux:~ # saphostctrl -function ListOSMetrics -metype VirtualComputerSystem
Timestamp           Metric                    MeasuredElement                Value Unit
2019-05-24-16:21:49 VCS.AMSenable             VirtualComputerSystem              0 Count
2019-05-24-16:21:49 VCS.GuaranteedCPUCapacity VirtualComputerSystem          12.00 Count
2019-05-24-16:21:49 VCS.GuaranteedMemory      VirtualComputerSystem   10485760 KiloBytes
2019-05-24-16:21:49 VCS.MemoryMaximum         VirtualComputerSystem  536870912 KiloBytes

linux:~ #
```

2.2 Operating system-specific monitoring tools on Linux on Power

This section gives an overview of the Linux tools for monitoring IBM Power Systems servers. Monitoring tools are required for SAP HANA on Power or SAP Applications on Power because many of the SAP delivered programs use specific information from the Power Systems architecture. For example:

- Information about logical partition (LPAR) type (shared or dedicated)
- LPAR mode like capped or uncapped
- Entitlement capacity
- Capacity limits

Many of the examples in the following sections are based on SMT2 configured systems to reduce the example output. In SAP environments, it is a best practice to use SMT8.

2.2.1 IBM Performance Management for PowerLinux systems

Your Power Systems model (including IBM Power Blades, PowerLinux, IBM PureFlex®, IBM System i, and IBM POWER) can be one of your company's most valued assets. It is imperative to understand the utilization and usage growth of your system to aid in creating plans to reduce cost, improve service, and manage risk. The IBM Performance Management for Power Systems (PM for Power Systems) is designed to fulfill this task.

This powerful tool provides critical resource and performance information about the current and long-term utilization trends from a total system view to an individual LPAR view. It helps to enable capacity planning for your Power Systems servers running IBM i, AIX, and Linux (Red Hat, SUSE, Fedora, CentOS, and Ubuntu operating systems) that are based on IBM PowerVM (partitions only) or PowerKVM (host and guests).

Insight is delivered through the collected PM for Power Systems performance data to show which upgrade can be required to accommodate future growth or a server consolidation in a virtualized IT world, or to assess the impact on the environment regarding a Capacity on Demand (CoD) processor or new processor hardware technology. Ongoing interactive assessments are also provided with the historical performance data to make it easily to *revisit* the utilization and capacity environment from up to 24 months ago.

PM for Power Systems is available in both *no additional charge* and *nominal charge* options depending on the level of detail that is required.

PM for Power Systems is available with minimal initial setup. After initial setup, the remaining process is automated, which helps to relieve system administrators of the tedious and recurring tasks that are involved with systems management. As a result, PM for Power Systems provides secure internet access to the analysis graphs.

In summary, PM for Power Systems asset management and virtualization capabilities provide a comprehensive process to manage today’s environment while planning for tomorrow’s challenges and opportunities.

2.2.2 Utilities for PowerPC hardware

The RPM Package Manager (RPM) package `powerpc-utils` contains utilities that are required for the maintenance of IBM PowerPC® platforms that follow the Power Architecture Platform Reference (PAPR). This section provides detailed information about relevant tools for monitoring. If these tools are used to change operating system parameters, they must be run as root. The command-line tools **`amsstat`**, **`lparstat`**, and **`ppc64_cpu`** are explained in more detail in this section.

To download the correct `powerpc-utils` package for all supported hypervisors and Linux distributions, see Service and productivity tools.

Active Memory Expansion

Active Memory Expansion (AME) can reduce the amount of memory that is used for NetWeaver LPARs (in the double digit range). This tool consists of two main areas, which are planning and viewing:

- AME Planning and Advisory Tool (`amepat`)
- AME stats can be found by running the following performance monitoring commands:
  - **`topas`**: The main window.

`TMEM`	True memory
`CMEM`	Compressed memory, `co=compress out` and `ci=compresses in`
`EF[T/A]`	Expansion Factor Target and Actual

  - **`vmstat -c`**: See `co=compress out` and `ci=compresses in` columns.
  - **`lparstat -c`**: See `%xcpu=expansion` CPU use.

Active Memory Sharing statistics (amsstat)

Although Active Memory Sharing (AMS) is not used in the context of SAP, this section lists it for completeness. In contrast to AMS, AME is widely used.

The **`amsstat`** tool captures the memory statistics that are relevant in an active AMS environment. This tool can run once or set to run repeatedly with a specified timeout, in seconds, between instances of data collection.

For most SAP Solutions, it is a best practice to disable the AMS feature. Only for some restricted configurations of SAP NetWeaver should AMS be enabled.

SAP HANA does not support AMS.

To check whether AMS is enabled or disabled, run (as root) **`amsstat`** without any parameters:

```
linux:~ # amsstat
Active Memory Sharing is not enabled on this system.
linux:~ #
```

The command is supported by only the PowerVM hypervisor. On a Linux system running under RHV-KVM, the hypervisor prints the following information message:

```
linux:~ # amsstat
amstat: is not supported on the Power KVM pSeries Guest platform
linux:~ #
```

Logical partition statistics

This command provides the starting point for all analysis. The **`lparstat`** command provides a report of the current LPAR-related parameters and hypervisor information, and utilization statistics for the LPAR. A requirement for valid LPAR statistics is a working resource monitoring and control (RMC) connection to the HMC and Novalink. If this connection does not exist or is damaged, the output of **`lparstat`** is outdated and not representative.

The following output shows an example of **`lparstat`** with five iterations and a 2-second interval time. To reduce the size of the output, Simultaneous Multi-Threading (SMT) level 2 is chosen:

```
linux:~ # lparstat 2 5
System Configuration
type=Dedicated mode=Capped smt=2 lcpu=2 mem=261832640 kB cpus=0 ent=2.00
%user  %sys %wait    %idle    physc %entc lbusy  vcsw phint
----- ----- -----    -----    ----- ----- ----- ----- -----
09.41 18.95  0.00    71.64 1.047284 52.364200 28.36 4873155174 26028487
```

```
11.22 16.90  0.00    71.88 0.524164 26.208200 28.12 4873180426 26028487
09.31 19.25  0.00    71.45 1.048336 52.416800 28.55 4873207632 26028487
10.94 17.74  0.00    71.32 0.524170 26.208500 28.68 4873234255 26028487
10.48 18.06  0.00    71.46 0.524168 26.208400 28.54 4873261114 26028487
linux:~ #
```

To get detailed information about the LPAR configuration, use the **-i** option. The options for interval and count do not affect the output. Example 2-1 shows the configuration of a dedicated LPAR with two cores.

Example 2-1 Output of the lparstat command

```
linux:~ # lparstat -i
Node Name                                  : linux
Partition Name                             : linux-a1ffa9ef-00000120
Partition Number                           : 6
Type                                       : Dedicated
Mode                                       : Capped
Entitled Capacity                          : 2.00
Partition Group-ID                         : 32774
Online Virtual CPUs                        : 2
Maximum Virtual CPUs                       : 6
Minimum Virtual CPUs                       : 1
Online Memory                              : 261832640 kB
Minimum Memory                             : 10240
Minimum Capacity                           : 1.00
Maximum Capacity                           : 6.00
Capacity Increment                         : 1.00
Active Physical CPUs in system             : 48
Active CPUs in Pool                        : 0
Shared Physical CPUS in system             : 0
Maximum Capacity of Pool                   : 0.00
Entitled Capacity of Pool                  : 0
Unallocated Processor Capacity             : 0
Physical CPU Percentage                    : 100
Unallocated Weight                         : 0
Memory Mode                                : Dedicated
Total I/O Memory Entitlement               : 268435456000
Variable Memory Capacity Weight            : 0
Memory Pool ID                             : 65535
Unallocated Variable Memory Capacity Weight  : 0
Unallocated I/O Memory Entitlement         : 0
Memory Group ID of LPAR                    : 32774
Desired Variable Capacity Weight           : 0
linux:~ #
```

The command is supported by only the PowerVM hypervisor. On a Linux system running under RHV-KVM, the hypervisor displays the following message:

```
linux:~ # lparstat
lparstat: is not supported on the Power KVM pSeries Guest platform
linux:~ #
```

CPU characteristics of PowerPC systems

The **ppc64_cpu** command is used to display and modify CPU characteristics like SMT level, DSCR settings, run-mode, and the subcore settings on the Power platform.

The following command shows which logical CPUs are on which core. Each enabled thread has a * mark after the thread number. For systems running SAP HANA on Power, the **ppc64_cpu** utility can be used for verification. Normally, SAP HANA runs in SMT8 mode.

```
linux:~ # ppc64_cpu --info
Core   0:    0*    1*    2*    3*    4*    5*    6*    7*
Core   1:    8*    9*   10*   11*   12*   13*   14*   15*
linux:~ #
```

This command determines the CPU frequency for 10 seconds:

```
linux:~ # ppc64_cpu --frequency -t 10
Power Savings Mode: None
min:    4.359 GHz (cpu 56)
max:    4.359 GHz (cpu 2)
avg:    4.359 GHz
linux:~ #
```

This command lists examples of input switches for the SMT option:

```
linux:~ # ppc64_cpu --smt
SMT=8
linux:~ # ppc64_cpu --smt=off
linux:~ # ppc64_cpu --info
Core   0:    0*    1     2     3     4     5     6     7
Core   1:    8*    9    10    11    12    13    14    15
linux:~ # ppc64_cpu --smt=4
linux:~ # ppc64_cpu --info
Core   0:    0*    1*    2*    3*    4     5     6     7
Core   1:    8*    9*   10*   11*   12    13    14    15
linux:~ #
```

2.2.3 Performance monitor

The **nmon** (Nigel's monitor) program is a performance system monitoring tool for Linux and AIX operating systems. **nmon** is delivered as an RPM package. The source code and compiled versions for the different operating systems are available at Nmon for Linux.

The **nmon** tool can be used for dynamic real-time observation of the running operating system or in batch mode to gather performance data into a file based on a spreadsheet-compatible format. Since version16g, the PowerVM LPAR section is available. When running on the PowerVM hypervisor, **nmon** collects comparable stats on AIX and Linux.

Figure 2-3 on page 39 shows the statistics of a 2-core LPAR in dedicated mode running at SMT2 level.

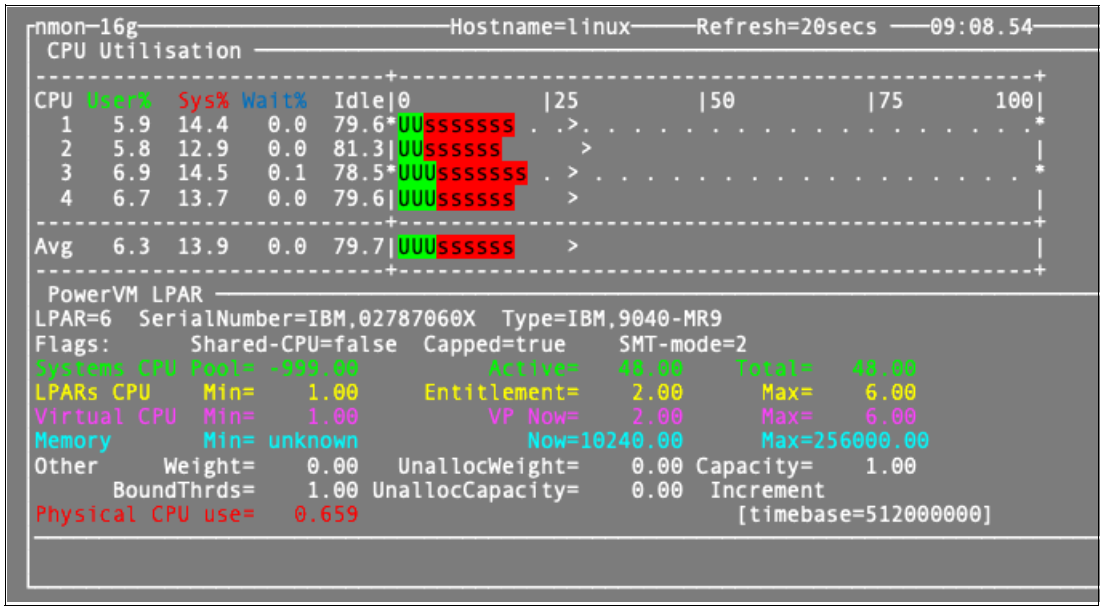

Figure 2-3 nmon performance monitor

Figure 2-4 shows the **nmon** output of a shared processor pool LPAR.

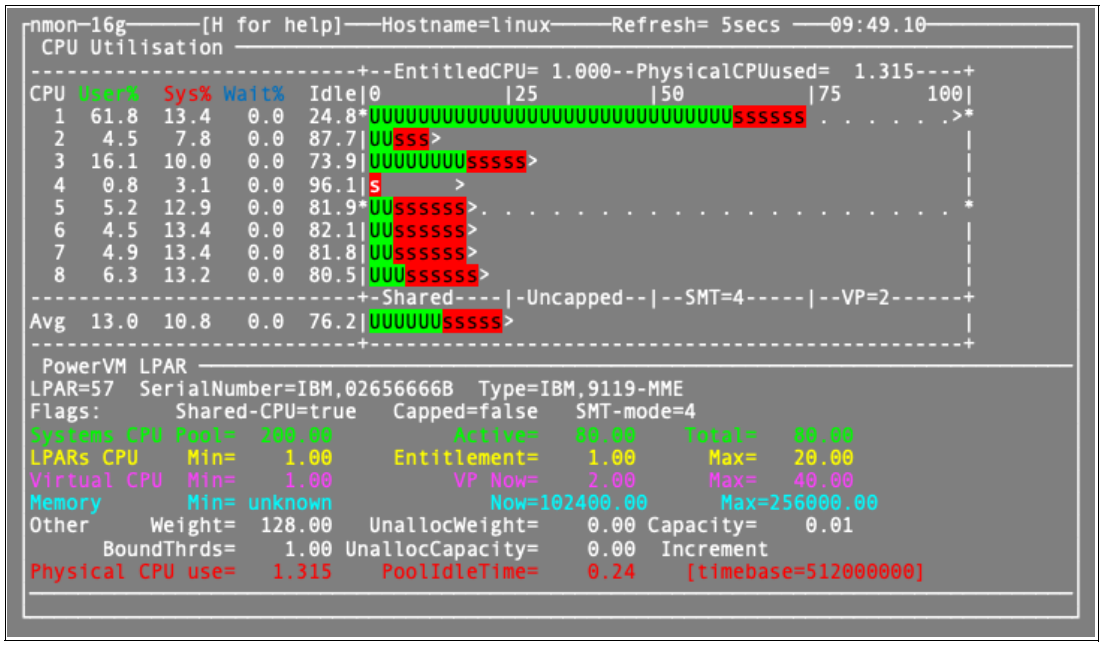

Figure 2-4 nmon performance monitor output: Shared processor pool LPAR

njmon for Linux and AIX is similar and valid as **nmon**. **njmon** saves data to JSON format for a new generation of online time-series databases and web-browser graphing.

2.3 Differences in SAP monitoring between hypervisors PowerVM and RHV-KVM

This section shows SAP monitoring differences between hypervisors PowerVM and RHV-KVM.

2.3.1 SAP monitoring for virtualization configuration and metrics

When using the SAP monitoring infrastructure (for example, with transaction ST06, as shown in Figure 2-1 on page 32), the virtualization configuration, and the corresponding metrics can be displayed. It makes no difference to SAP applications that use SAP monitoring if they are running under Red Hat Enterprise Linux or SUSE Linux Enterprise Server. However, the available metrics depend on the capabilities of the hypervisor (PowerVM or RHV-KVM).

Table 2-1 lists the virtualization-related metrics that are gathered by the individual hypervisors.

Table 2-1 Comparison between hypervisors and metrics

Metric	PowerVM dedicated LPAR	PowerVM Shared Processor LPAR with pool utilization authority (PUA) not granted	PowerVM Shared Processor LPAR with PUA granted	RHV-KVM guest
Model	√	√	√	√
Processor	√	√	√	√
Pool ID		√	√	
PUA		√	√	
Pool CPUs		√	√	
Pool CPUs Idle		√	√	
Current Processor Frequency	√	√	√	√
Virtual CPUs	√	√	√	√
SMT Mode	√	√	√	√
Threads	√	√	√	√
Partition Name	√	√	√	
Partition ID	√	√	√	
Partition Type	√	√	√	
Capped		√	√	
Weight		√	√	
Capacity Consumed		√	√	√
Guaranteed Capacity	√	√	√	
Guaranteed Capacity Consumed			√	

Metric	PowerVM dedicated LPAR	PowerVM Shared Processor LPAR with pool utilization authority (PUA) not granted	PowerVM Shared Processor LPAR with PUA granted	RHV-KVM guest
Available Capacity	√		√	√
Available Capacity Consumed			√	√
Additional Capacity Available				√
Capacity Maximum				√

Within the output of transaction ST06 as shown in Figure 2-1 on page 32, the information that is shown in Table 2-1 on page 40 can be found in the CPU Virtualization Host and the CPU Virtualization Virtual System as follows:

- The CPU Virtualization Host shows processor information that is relevant and defined on the host:
  - Model: Host server model type.
  - Processor: Host server processor type.
  - Pool ID: (PowerVM LPAR only) The shared processor pool number for the LPAR.
  - PUA: (PowerVM LPAR only) This field indicates whether the LPAR has the authority to retrieve information about the shared pool, for example, the idle processor capacity in the shared pool. The possible values are:
    - Granted.
    - Not granted.
  - Pool CPUs: (PowerVM LPAR only) This field shows the total number of physical processors in the shared pool to which this LPAR belongs.
  - Pool CPUs Idle: (PowerVM LPAR only) This field indicates the idle capacity of the shared pool in units of physical processors. This value is only available if the PUA is granted.
- The CPU Virtualization Virtual System shows processor information that is relevant and defined on the LPAR or the RHV-KVM guest:
  - Current Processor Frequency: The current operating processor frequency.
  - Virtual CPUs: This field shows the number of administrator-defined virtual processors. Virtual processors are an abstraction of processors, which are mapped to physical processors in a scheduled manner by the hypervisor. The number of virtual processors that is defined places an implicit limit on the number of physical processors that can be used by the LPAR or the RHV-KVM guest.
  - SMT Mode: This field indicates whether SMT is active. Possible values are:
    - On.
    - Off.

- Threads: The number of active SMT threads. For PowerVM LPARs, the value is calculated as follows:
  - For dedicated LPARs, the number of logical processors equals the number of dedicated physical processors that are assigned to the LPAR, multiplied by the number of SMT threads.
  - For shared processor LPARs, the number of logical processors equals the number of virtual processors that are defined for the LPAR, multiplied by the number of SMT threads.
- Partition Name: (PowerVM LPAR only) The HMC-defined LPAR name.
- Partition ID: (PowerVM LPAR only) The HMC-defined LPAR number.
- Partition Type: (PowerVM LPAR only) This field describes the type of the LPAR. The possible values are:
  - Dedicated LPAR.
  - Shared processor LPAR.
- Capped: (PowerVM LPAR only) This field indicates whether a Shared Processor LPAR can exceed its entitlement.

  The possible values are:
  - On.
  - Off.
- Weight: (PowerVM LPAR only) This value is used to determine the allocation of spare pool resources to uncapped shared processor LPARs.
- Capacity Consumed: This field indicates the computing power that the LPAR or RHV-KVM guest consumes in units of physical processors.
- Guaranteed Capacity: This field shows the guaranteed physical processor capacity of an LPAR or RHV-KVM guest in units of fractional physical processors.
- Guaranteed Capacity Consumed: This field indicates the ratio of the consumed physical processor to the guaranteed capacity as a percentage. In the case of an uncapped shared processor LPAR, the value can exceed 100%.
- Available Capacity: This field indicates the possible available computing power for the LPAR or RHV-KVM guest. For shared processor LPARs, it is based on the entitlement, which is guaranteed to the LPAR from the current idle capacity in the pool. This value is available only if the PUA is granted.
- Available Capacity Consumed: This field indicates the ratio of physical processors that the LPAR or RHV-KVM guest consumes to the available capacity for the LPAR or RHV-KVM guest as a percentage. This value is available only if the PUA is granted.
- Additional Capacity Available: The amount of physical processor capacity that can still be attained by the LPAR or RHV-KVM guest.
- Capacity Maximum: The maximum amount of physical processor capacity that can be acquired by the LPAR or RHV-KVM guest.

Note: SAP Notes 1131691 and 1379340 document the preferred or eligible values for SAP applications.

A

The fsperf tool

This appendix discusses the **fsperf** open source file system performance tool.

This appendix contains the following sections:

- Introduction to fsperf
- Command-line options
- Running tests
- Output of results of the initial write test

Introduction to fsperf

This open source tool is designed as a file system performance testing framework. For more information, see GitHub.

Usage information for **fsperf** is available by running **fsperf -help**. The most important flags for **fsperf** are **-m** and **-t**:

```
-t, --test-mode TEST_MODE
-m, --measurements MEASUREMENT_MODE
```

These flags specify which performance test or tests to run:

```
-initial_write
-overwrite
-write (= initial_write & overwrite)
-read
-all (= write & read, Default)
```

These flags specify which performance measurements are taken:

```
-throughput (Default)
-latency
-all (= throughput & latency)
```

These flags define the kind of test to be run with defaults specified. The **-t all** flag means to run all the test mode, and **-m -all** means test for both latency and throughput even though the default is only throughput.

The following output is a sample of the run, with a file size of 16 GB (specified by the **-f** flag), a block size of 16 K (**-b** flag), **-m all** (running all the performance measurements), a kernel I/O queue of 2 K, a maximum parallel I/O request of 2 K, and a queue length of 8. The files system to be tested is `/hana/data/dummy`. This file system is backed by a Fibre Channel adapter, and the disk that is under test is configured as NPIV.

```
# ./fsperf -i random -o verbose -f 16G -b 16K -m all  --param
size_kernel_io_queue=2048 --param max_parallel_io_requests=2048 --param
num_submit_queues=8 /hana/data/dummy
```

Command-line options

The original command was:

```
./fsperf -i random -o verbose -f 16G -b 16K -m all --param
size_kernel_io_queue=2048 --param max_parallel_io_requests=2048 --param
num_submit_queues=8 /hana/data/dummy
```

The tests are run with the following settings:

```
Test mode:.............................all
Measurements...........................all
Program output:........................verbose
Block size:............................16kB
File size:.............................16GB
I/O access order:......................random
Randomize data:........................yes
Interactive:...........................no
```

```
No cleanup: ..........................no
No initialization:....................no
Directory path:......................./hana/data/dummy
```

More program parameters are:

```
size_kernel_io_queue=2048
max_parallel_io_requests=2048
num_submit_queues=8
```

Running tests

The configuration parameters that are set are:

```
async_write_submit_active=auto,async_write_submit_blocks=all,async_read_submit=on,
num_submit_queues=8,num_completion_queues=1,size_kernel_io_queue=2048,max_parallel
_io_requests=2048,min_submit_batch_size=16,max_submit_batch_size=64
```

- Running test for initial write latency:

  ```
  Opening and initializing file "/hana/data/dummy/datavol.dat"... done.
  Preparing I/O buffer... done.
  Test of latency for initial write has completed.
  ```

- Running test for initial write throughput:

  ```
  Opening and initializing file "/hana/data/dummy/datavol.dat"... done.
  Preparing I/O buffer... done.
  Test of initial write has completed.
  ```

- Running overwrite latency test:

  ```
  Preparing I/O buffer... done.
  Test of overwrite latency has completed.
  ```

- Running overwrite throughput test:

  ```
  Preparing I/O buffer... done.
  Test of overwrite has completed.
  ```

- Running read latency test:

  ```
  Preparing I/O buffer... done.
  Test of read latency has completed.
  ```

- Running read throughput test:

  ```
  Preparing I/O buffer... done.
  Test of read has completed.
  Tests completed.
  ```

Output of results of the initial write test

- Throughput test results:

```
Trigger time:................... 0.473454 s   (Throughput:     34605.2 MB/s)
Asynchronous submit time:.......  10.6405 s   (Throughput:     1539.76 MB/s)
Synchronous submit time:........        0 s   (Throughput:           0 MB/s)
I/O time:.......................  85.1065 s   (Throughput:     192.511 MB/s)
Ratio trigger time to I/O time:.0.00556307
```

- Latency test results:

```
I/O time:.......................  647.863 s   (Throughput:     25.2892 MB/s)
Latency:........................      617 us
```

- Overwrite test:
  - Throughput test:

```
Trigger time:................... 0.466746 s   (Throughput:     35102.6 MB/s)
Asynchronous submit time:.......  1.73355 s   (Throughput:     9451.08 MB/s)
Synchronous submit time:........        0 s   (Throughput:           0 MB/s)
I/O time:.......................  14.0032 s   (Throughput:     1170.01 MB/s)
Ratio trigger time to I/O time:.0.0333312
```

  - Latency test:

```
I/O time:.......................  679.438 s   (Throughput:      24.114 MB/s)
Latency:........................      647 us
```

- Results of the read test:
  - Throughput test:

```
Trigger time:................... 0.477828 s   (Throughput:     34288.4 MB/s)
Asynchronous submit time:....... 0.814007 s   (Throughput:     20127.5 MB/s)
Synchronous submit time:........        0 s   (Throughput:           0 MB/s)
I/O time:.......................  6.71068 s   (Throughput:     2441.47 MB/s)
Ratio trigger time to I/O time:. 0.071204
```

  - Latency test:

```
I/O time:.......................  4219.07 s   (Throughput:     3.88331 MB/s)
Latency:........................     4023 us
```

The output of the command is self-explanatory, and gives information about what test is running and completed at a particular instance. At the end, the command gives a summary of throughput and latency for all test modes that were run.

These key performance indicators (KPIs) numbers can be matched with what IBM or SAP recommend for their applications, and tuned for better performance.

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 topic in this document. Some publications that are referenced in this list might be available in softcopy only.

- *IBM Power Systems Infrastructure I/O for SAP Applications*, REDP-5581
- *IBM Power Systems Virtualization Operation Management for SAP Applications*, REDP-5579
- *Performance Optimization and Tuning Techniques for IBM Power Systems Processors Including IBM POWER8*, SG24-8171

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:

- *Installation Quick Start SUSE Linux Enterprise Server for SAP Applications 15 SP1*:

  https://documentation.suse.com/sles-sap/15-SP1/pdf/SLES4SAP-quick_color_en.pdf
- Installing and setting up Linux on Power Systems servers:

  https://www.ibm.com/support/knowledgecenter/en/linuxonibm/liabx/installsetupparent.htm

  For SAP HANA for Linux on power, tailored documentation was published through the SAP Launchpad.
- Overview of the Red Hat Enterprise Linux for SAP Applications subscription:

  https://access.redhat.com/solutions/34169
- Prepare your Linux for your SAP solution with saptune:

  https://blogs.sap.com/2017/12/22/prepare-your-linux-for-your-sap-solution-with-saptune/
- Resource Optimized High Availability in PowerHA SystemMirror:

  https://www.ibm.com/support/knowledgecenter/en/SSPHQG_7.2/admin/ha_roha.html

- sapconf – A way to prepare a SLES system for SAP workload – Part 1:

  https://www.suse.com/c/sapconf-a-way-to-prepare-a-sles-system-for-sap-workload-part-1/
- SAP Note 2535891 - Run your SAP workload best on AIX and POWER8 or POWER9 in SMT8 mode:

  https://support.sap.com/en/my-support/knowledge-base.html

Help from IBM

IBM Support and downloads

ibm.com/support

IBM Global Services

ibm.com/services

REDP-5580-00

ISBN 0738459127

Printed in U.S.A.