IBM Redbooks

IBM Power Systems Planning and Monitoring Best Practices for SAP Applications

July 2020
Note: Before using this information and the product it supports, read the information in "Notices" on page v.

First Edition (July 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

This document was created or updated on July 29, 2020.

© Copyright International Business Machines Corporation 2020. All rights reserved.
Note to U.S. Government Users Restricted Rights -- Use, duplication or disclosure restricted by GSA ADP Schedule Contract with IBM Corp.
# Contents

## Notices

<table>
<thead>
<tr>
<th>Contents</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trademarks</td>
<td>vi</td>
</tr>
</tbody>
</table>

## Preface

<table>
<thead>
<tr>
<th>Contents</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Authors</td>
<td>vii</td>
</tr>
<tr>
<td>Now you can become a published author, too!</td>
<td>viii</td>
</tr>
<tr>
<td>Comments welcome</td>
<td>viii</td>
</tr>
<tr>
<td>Stay connected to IBM Redbooks</td>
<td>ix</td>
</tr>
</tbody>
</table>

## Chapter 1. Planning and implementation best practices for SAP workloads

<table>
<thead>
<tr>
<th>Contents</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1 Introduction</td>
<td>2</td>
</tr>
<tr>
<td>1.2 Planning and implementation recommended practices for SAP workloads</td>
<td>2</td>
</tr>
<tr>
<td>1.3 High level SAP installation selection in the Product Availability Matrix (PAM)</td>
<td>2</td>
</tr>
<tr>
<td>1.4 SAP applications configurations</td>
<td>9</td>
</tr>
<tr>
<td>1.5 Linux installation</td>
<td>9</td>
</tr>
<tr>
<td>1.5.1 The Linux matrix</td>
<td>10</td>
</tr>
<tr>
<td>1.5.2 Hardware dimension</td>
<td>10</td>
</tr>
<tr>
<td>1.5.3 Linux distribution dimension</td>
<td>11</td>
</tr>
<tr>
<td>1.5.4 Linux operating system customization dimension</td>
<td>12</td>
</tr>
<tr>
<td>1.5.5 Endianness variants</td>
<td>12</td>
</tr>
<tr>
<td>1.6 Linux customization with SAP base patterns</td>
<td>13</td>
</tr>
<tr>
<td>1.6.1 SAP Application Server Base Pattern</td>
<td>14</td>
</tr>
<tr>
<td>1.6.2 SAP NetWeaver Server Base Pattern</td>
<td>14</td>
</tr>
<tr>
<td>1.6.3 SAP HANA Server Base Pattern</td>
<td>14</td>
</tr>
<tr>
<td>1.7 Linux tuning for SAP workloads</td>
<td>15</td>
</tr>
<tr>
<td>1.7.1 tuned</td>
<td>15</td>
</tr>
<tr>
<td>1.7.2 SLES V11 sapconf</td>
<td>17</td>
</tr>
<tr>
<td>1.7.3 saptune</td>
<td>17</td>
</tr>
<tr>
<td>1.8 Hardware aspects</td>
<td>20</td>
</tr>
<tr>
<td>1.8.1 NUMA</td>
<td>20</td>
</tr>
<tr>
<td>1.8.2 I/O configuration and tuning</td>
<td>28</td>
</tr>
<tr>
<td>1.9 Online references</td>
<td>30</td>
</tr>
</tbody>
</table>

## Chapter 2. Monitoring

<table>
<thead>
<tr>
<th>Contents</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1 SAP monitoring components</td>
<td>31</td>
</tr>
<tr>
<td>2.1.1 SAP Host Agent</td>
<td>32</td>
</tr>
<tr>
<td>2.2 Operating system specific monitoring tools on Linux on POWER</td>
<td>34</td>
</tr>
<tr>
<td>2.2.1 IBM Performance Management for Power Linux Systems (PMLinux)</td>
<td>35</td>
</tr>
<tr>
<td>2.2.2 Utilities for PowerPC hardware</td>
<td>35</td>
</tr>
<tr>
<td>2.2.3 Performance monitor (nmon)</td>
<td>38</td>
</tr>
<tr>
<td>2.3 Differences in SAP monitoring between hypervisor PowerVM and RHV-KVM</td>
<td>39</td>
</tr>
<tr>
<td>2.3.1 SAP monitoring for virtualization configuration and metrics</td>
<td>39</td>
</tr>
</tbody>
</table>

## Appendix A. fsperf

<table>
<thead>
<tr>
<th>Contents</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Introduction to fsperf</td>
<td>43</td>
</tr>
<tr>
<td>Command line options</td>
<td>44</td>
</tr>
<tr>
<td>Running tests</td>
<td>45</td>
</tr>
<tr>
<td>Output of results of initial write test</td>
<td>45</td>
</tr>
</tbody>
</table>
Related publications .................................................. 47
IBM Redbooks ......................................................... 47
Online resources ......................................................... 47
Help from IBM .......................................................... 48
Notices

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

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

IBM may have patents or pending patent applications covering subject matter described in this document. The furnishing of this document does not grant you any license to these patents. You can send license inquiries, in writing, to:

IBM Director of Licensing, IBM Corporation, North Castle Drive, MD-NC119, Armonk, NY 10504-1785, US

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

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

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

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

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

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

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

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

COPYRIGHT LICENSE:

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

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

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

AIX®  POWER®  PureFlex®
DB2®   POWER9™  Redbooks®
IBM®   PowerHA®  Redbooks (logo)
Insight®  PowerPC®  System i™
POWER®  PowerVM®  SystemMirror®

The following terms are trademarks of other companies:

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

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

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

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

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

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 documentation supporting 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 specialist, and systems administrators deploying SAP workloads whom often spend a lot of time and effort managing, provisioning, and monitoring SAP software systems and landscapes on IBM Power Systems.

Authors

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

Dino Quintero is an IT Management Consultant and an 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, 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. He has worked on functional verification testing for VIOS and IBM AIX®. In his current role, he supports IBM 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, with operating systems AIX and Linux on POWER, provided for SAP applications. In earlier years he was working in different Software Engineering roles on IBM AIX environments. In his current role, he is responsible for the IBM POWER platform support provided for the 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 responsible for the Datacenter Readiness and Eco System enablement of SAP HANA on Power Systems. She has 15 years background on SAP on IBM Power Systems in the area of Storage, Business Continuity, Disaster Recovery AIX and Linux.

Marc-Stephan Tauchert is a System Specialist at IBM Germany. In his 25+ years’ experience in designing and implementing SAP Solutions on IBM Power Systems, he has covered various topics and roles. As 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. Actually he is member of 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, Poughkeepsie Center

Walter Orb, 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
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:

► Stay current on recent Redbooks publications with RSS Feeds:
  http://www.redbooks.ibm.com/rss.html
Chapter 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 recommended practices for SAP workloads
- High level SAP installation selection in the Product Availability Matrix (PAM)
- SAP applications configurations
- Linux installation
- Linux customization with SAP base patterns
- Linux tuning for SAP workloads
- Hardware aspects
- Online references
1.1 Introduction

This publication is part of a series of technical documents supporting clients in planning, implementing and operating 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.

1.2 Planning and implementation recommended practices for SAP workloads

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

The conceptual starting point is the selection of an SAP product to fulfil a given business need. This decision is assumed to already been made.

1.3, “High level SAP installation selection in the Product Availability Matrix (PAM)” on page 2 introduces the procedure on how to select the correct SAP software components for an IBM POWER server based installation. This task can conceptually be thought as selecting a platform, in a generalized sense, for the desired SAP product. The term generalized platform here defines not only the operating system, but also the database and possibly other infrastructural components, which together build the components on which the SAP applications run.

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

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

Finally, 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 (PAM)

After the decision for an SAP product has been made, based on business needs, for example, an enterprise resource planning or developing customized reports, proper technical components need to be selected to implement it.

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

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

The tool provided by SAP for this purpose is called the Product Availability Matrix (PAM), and it is available at the following website:

http://support.sap.com/pam

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

![Figure 1-1  Schematic structure of the SAP Product Availability Matrix (PAM)](image)

In a traditional sense, choosing a platform is generally thought as choosing the operating system and hardware, or even just the database component. SAP from the PAM perspective defines the term platform consisting not only of the operating system and hardware, but also the SAP Business Application Stack. Figure 1-1 shows the platform components include the SAP Kernel (light green), the database (blue) and the operating system and hardware (dark blue).

In the following, an example navigation through the PAM for the appropriate components for SAP NETWEAVER 7.5 is described:

Access the SAP PAM:

http://support.sap.com/pam

---

1 This is accessible by way of the SAP intranet, or through the Internet with a valid SAP user ID.
Using the search tool in the upper right corner of the main PAM page, enter NetWeaver 7.5. From the resulting search results as shown in Figure 1-2, click the first search result.

Figure 1-2  Search for the NetWeaver entry point
Figure 1-3 shows the main page for the NetWeaver V7.5 product.
By selecting the Technical Release Information tab, the possible choices for the generalized platform for NetWeaver 7.5 are shown in Figure 1-4. Then, choosing an option from the Product Instance column allows for further refinement of the details for this platform.

The Application Server ABAP is selected as shown in Figure 1-4. To find details of the database platforms for this solution, use the link → Details under column Database Platforms or click it.
Other columns, such as Languages and Web Browsers can be interpreted as additional dimensions of the overall selection matrix. These cases are not discussed in this publication.

The resulting details, as shown in Figure 1-5, 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.

The operating system choice also implies a hardware selection. Some further 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 IBM POWER hardware.

From conceptual view, the SAP products seen in the PAM are built of two layers, the lower SAP kernel layer and the higher SAP business layer. The SAP kernel layer consists of a set of executables and tools which at runtime provides an application server functionality. The business layer’s objects and entities then run conceptually in this application server environment. This modular approach can later be utilized by exchanging at suitable time an SAP kernel with a newer one although the SAP business layer can remain as is.
The column on 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 finally 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.

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 currently available. Components that are excluded are not select able as they are not currently validated or available. For example, any option for LINUX ON POWER LITTLE ENDIAN is not selectable with DB2 LUW 64-BIT (Linux UNIX Windows), while LINUX ON POWER BIG ENDIAN options exist (for example, Red Hat Enterprise Linux V6, V7, and SUSE Linux Enterprise Server V11).

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 additional 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 need to 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, just to name a few.
- These specifications can then be mapped into SAP application parameters, such as the amount of extended memory for the required business objects and the number of parallel work processes needed to meet the need for the business processes and concurrent users.
- The result of the analysis at the application layer is then 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. These 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 consultants and sizing tools.

Visit [www.sap.com/sizing](http://www.sap.com/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 re-evaluation of the solution’s resource requirements to evaluate if they are still meeting the business processes as they change and evolve or if 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, additional hardware resources (for example, more CPU capacity, additional memory, and others) are needed or after business process optimizations the LPAR size can be reduced. New demands can include need for high availability, a disaster recovery solution, implementing new solution monitoring and management tools, backups, and others.

There are many solutions that can extend the functionality of every SAP installation. Choosing to deploy new components to enhance the existing system require 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 consultants or by the vendor of a product (for example, snapshot backup or security software).

1.5 Linux installation

Linux has evolved over the years to an enterprise level solution with a vast ecosystem.
As an Open Source project, the Linux operating system is independent of the distribution vendor. With a big development and support community, a large variety of supported hardware platforms, and an ever increasing array of software solutions, the Linux platform provides a big variety from still not native to enterprise ready. Moreover IBM as a major contributor also provides an optimal support for Power Systems.

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

1.5.1 The Linux matrix

When a choice of one solution is to be made 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 IBM Power Systems, the following sections illustrate some of the main dimensions and attributes for running Linux on POWER:

► Hardware dimension.
► Linux distribution dimension.
► Linux operating system customization dimension.
► Endianness variants.

1.5.2 Hardware dimension

Considering each hardware platform has different capabilities, the platforms’ hardware features can be exploited by enabling it in the kernel. For the POWER platform, some of these features include:

► Running Linux on IBM PowerVM® or KVM Hypervisor or bare metal.
   - In SAP Landscapes, the typical deployment on Power Systems is to host many instances of the operating system in virtual machines by installing a hypervisor layer that the Linux operating system runs upon to benefit from workload consolidation, and use the POWER features coming only with the E-class models on PowerVM. For a more detailed discussion see IBM Power Systems Virtualization Operation Management for SAP Applications, REDP-5579 (https://ibm.co/2WWHVrR).
   - The small IBM Power Systems servers are able to 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 are typically not used for SAP applications and are often deployed in HPC environments.
► Simultaneous Multi-Threading (SMT).
► Live Partition Mobility.
► An LPAR running on PowerVM can be configured to allow for live migration of the LPAR to another machine addressing 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, to help facilitate problem determination.
Visit the following site to learn more about these and other features and capabilities of Linux on Power Systems servers:


1.5.3 Linux distribution dimension

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

➤ SUSE distributes SUSE Linux Enterprise Server, or SLES for short:
  - As of 2019, the supported versions are SLES V11, SLES V12, and SLES V15.
  - Each version further subdivides into a number of service packs. For instance, SLES V12 SP3.
  - Each service pack can support a different Linux kernel version, which in turn can define specific software solution options that are able to run on that kernel version. For instance, SLES V12 SP1 is based on the Linux kernel V3.1, although SLES V12 SP3 is based on the Linux kernel V4.4.
  - For more information about SLES, visit the following website:
    https://en.wikipedia.org/wiki/SUSE_Linux_Enterprise
    - Both companies ship a specific SAP Edition where support for SAP centric tooling come along ready for use. Furthermore, the release cycles are adapted to the ones from SAP.

➤ 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 instance Red Hat Enterprise Linux V7.6.
  - For more information about Red Hat Enterprise Linux, visit the following website:
    https://en.wikipedia.org/wiki/Red_Hat_Enterprise_Linux

➤ Each Linux distribution usually has some unique tools that have been designed to manage the operating environment:
  - Package distribution methods, such as RPM (RPM Package Manager), APT (Advanced Packaging Tool), or even tar.gz.
  - Red Hat and SUSE both use RPM.

➤ Installation and package update tools:
  - Red Hat provides the gpk-application GUI tool (X-windows), or the YUM (Yellowdog Update Manager) command line tool.
  - SUSE provides the YAST (Yet Another Setup Tool) graphical tool (X-windows or cursors), and zypper command line.

➤ Analysis tools, including support config for SLES and sosreport for Red Hat Enterprise Linux. Also SAP specific tools are provided by the distributors as described in the following section.
1.5.4 Linux operating system customization dimension

Both SUSE and Red Hat provide a customized version of their enterprise operating system that provide utilities, settings and features that are designed to help make the administration of the system easier with SAP specific tools and predefined settings. Furthermore, the lifecycle of the releases help to stay in synchronization with the SAP lifecycles in a supported and competitive manner.

- In addition to the standard SUSE Linux Enterprise Server distribution, SUSE provides SUSE Enterprise Server for SAP Application as shown in the following website:
  
  https://www.suse.com/products/sles-for-sap/

  For more details, visit the following website:
  

- In addition to the standard Red Hat Enterprise Linux distribution, Red Hat provides the Red Hat Enterprises Linux for SAP Solutions version as shown in the following website:
  

  For more details, visit the following website:
  
  https://access.redhat.com/solutions/34169

- 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 installs automatically additional packages.

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

Note: Both vendors provide along with the technical features also 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 endiannahses for POWER: Big Endian and Little Endian. Data in the operating system is represented as bytes, and the different endiannahses 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 endianess.
Linux customization with SAP base patterns

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

Note: Currently, most of IBM SAP on Linux on POWER installations are based on SUSE Linux Enterprise Server (SLES). But Red Hat delivers same functionality and integration and is equally supported by SAP and IBM.

For information regarding required software packages for Red Hat Enterprise Linux, refer to the following website:
https://access.redhat.com/solutions/3082481

In collaboration with SAP and IBM, the Linux distributors have made customizations to their operating systems targeting SAP applications to run easier and perform better. These customizations are defined as base patterns, and are customized for many SAP application configurations.

In the following scenario SUSE is chosen as the example:

SUSE has developed SAP-specific patterns. With these SAP patterns SUSE pursues the path to make the implementation of an SAP NetWeaver or SAP HANA based on SLES or SLES for SAP Applications easier.

The SAP base patterns from different SAP solutions on Linux on POWER are available as RPM packages. These packages include all necessary files and programs needed to enhance the configuration, implementation and management of an SAP solution. Each package has a small documentation included with the components of the package. Installation is simple, as all dependent operating system packages are installed with the base pattern RPM. The following sections give an overview of which operating system packages for the different SAP solutions are required.

There are three SAP solutions supported by SUSE:
  a. SAP NetWeaver.
  b. SAP HANA.
  c. SAP NetWeaver + HANA.
In each use case, customization of the Linux operating system happens by installing one or more base patterns. For each use case, the following are the suggested base pattern packages to install (described in the following sections):

- For use case a) install SAP Application Server Base Pattern (currently required for any SAP installation), then install SAP NetWeaver Server Base Pattern.
- For use case b) install SAP Application Server Base Pattern, then the SAP HANA Server Base Pattern.
- For use case c) 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 which can change. Refer to your SAP and operating system product installation guides for the current selection of patterns.

### 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 helps to automatically apply the base tuning, can be adjusted to specific needs, and ensures over time compliance of the tuning and consistency in the landscape.

### 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 LUKS-based disk encryption.
- Installation wizard for SAP applications.
- Symlink to firefox.
- Integration of SAP Network Interface Router into systemd.
- SLES for SAP White Papers.
- Dynamic adaptive system tuning daemon.
- Collection of basic system utilities.
- Helper daemon to guarantee uniqueness of time-based 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:

- Set up system firewall for SAP HANA database.
- 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.
- Virus Scan Adapter (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 runtime, International Components for Unicode, JPEG Support, PNG support and SSH Support.
Support Utilities for NFS.
Network Time Protocol daemon.
NUMA Policy Control.
Installation wizard for SAP applications.
Symlink to firefox.
Integration of SAP Network Interface Router into systemd.
SLES for SAP White Papers.
Hardware health monitoring for Linux.
Execute 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 SLES ensure that all needed operating system packages are installed. To start with a 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 higher, the tuned service package is available, which contains some specific tuned profiles for SAP solutions.

On SUSE SLES and SLES for SAP, the package saptunev2 optimizes SAP solutions depending on the SLES version.

1.7.1 tuned

Tuned is an optional Linux operating system daemon that monitors connected devices and tunes system settings for CPU tuning according to a selected profile adjusted to key SAP Notes or any other application specific best practice. It builds the foundation for the automation of Red Hat and SUSE.

The daemon helps to tune the base operating system configuration by monitoring the system and adjusting operating system parameters for changing workload characteristics. It is provided by an independent RPM package and is installed by way of the RPM command line tool (rpm -i tuned), or by the operating system package manager zypper or YaST on SLES, and yum on Red Hat Enterprise Linux.

SLES provides the additional packages sapconf and saptune that provide additional input to tuned related to the running of SAP workloads.

Since Red Hat Enterprise Linux V6.5, Red Hat provides the RPM package tuned-profiles-sap and sapconf containing specific tuned profiles for SAP solutions.

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

- sysctl.
The static tuning mainly consists of predefined sysctl and sysfs settings and one-time activation of several configuration tools such as ethtool. Tuned 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.

Tuned 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, usage of the configuration profiles, and their placement in the system environment.

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

When saptune is installed (see 1.7.2, “SLES 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 defined in the new profile.
1.7.2 SLES V11 sapconf

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

- tuned.
- uuidd.
- sysstat.

The tuned profiles 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, “tuned” on page 15).

Note: Starting with SLES V12 SP2, it is highly recommended to only use saptune and not sapconf.

Starting sapconf starts the tuned daemon and setups 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 the script.sh. The sapconf executable controls the active profile, and can start and stop the service.

For more information about installation, configuration, update and architecture for sapconf, visit the SUSE blog at the following website:


1.7.3 saptune

With SLES V12 SP2, saptune2 is the recommended mechanism to use. Saptune is a program package offered by SUSE and it is only available in SLES for SAP Applications. This program package adapts parameter recommendations from SAP notes or SUSE to specific SAP solutions. Currently there are two versions of saptune available. The old V1 and the new complete redesigned V2. The saptune V2 contains the V1 and can be switched to that by way of the general configuration profile. This option allows customers using V1 to use saptune with their old settings.

Note: It is mandatory to use saptune V2 starting with SLES V12 SP2. Do not use any older version, and it is recommended to be used to tune.

The new V2 of saptune 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 pages saptune-note.

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

Example 1-1  saptune note list

# saptune note list
All notes (+ denotes manually enabled notes, * denotes notes enabled by solutions, - denotes notes enabled by solutions but reverted manually later, 0 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)
Furthermore, saptune can create, verify, apply, simulate, customise 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 a overview of which kind of solutions are available, which notes are included in a solution and which of the solutions are active, use the command saptune solution list as shown in Figure 1-8. Notes which 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 23382421 2534844
  MAXDB          - 941735 1771258 1984787
  NETWEAVER      - 941735 1771258 1984787
  NETWEAVER+HANA - 941735 1771258 1980196 1984787 2205917 23382421 2534844
  S4HANA-APP+DB  - 941735 1771258 1980196 1984787 2205917 23382421 2534844
  S4HANA-APPSEVER - 941735 1771258 1984787
  S4HANA-DBSERVER - 941735 1771258 1980196 1984787 2205917 23382421 2534844
  SAP-ASE         - 941735 1410736 1608003 1771258 1984787

linux:- #
```

Figure 1-8  Output saptune solution list

Operations like apply, verify, simulate and revert for the installed solutions can be executed by way of the command line tool saptune.

To activate and view the overall status of saptune, there are several options available. To view the status of saptune run the command 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 command saptune daemon start starts saptune. If no solution or note is selected, there is a warning to apply one of the available notes and solutions. The command saptune revert all removes all parameter changes and resets the operating system to the defaults.
Saptune is enabled by the tuned daemon at operating system boot time. Figure 1-10 shows the file and directory structure of the saptune environment including the used file structure of the tuned daemon.

Figure 1-10   Directory structure of the saptune

Saptune 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 manual page for saptune.

1.8 Hardware aspects

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

1.8.1 NUMA

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

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

With IBM POWER9 processors, the interconnect bandwidth on each node and across multiple nodes has been improved by 4X compared to IBM POWER8 processors. This has resulted in improvement of throughput and reduction of latency for SAP applications running on IBM POWER9 systems compared to their IBM POWER8 predecessors.

**View the NUMA topology from inside the operating system**

Looking at the NUMA topology from inside a Linux operating system requires knowledge on what information is updated as the LPAR placement can change after a logical partition mobility (LPM) operation or similar tasks. The command used is only updated on restart and hence might display outdated information. For current information, a dump on the HMC has to be pulled as shown in “View the LPAR placement from the HMC” on page 23.

Log into the LPAR and ensure the command `numactl --hardware` has a symmetric output. Specially confirm every NUMA node with a core holds memory as well. Nodes without cores are fine to hold no memory too. This case is often seen for NUMA node 0 not having cores or memory assigned. In such case, SAP HANA internally maps all NUMA nodes, and work only with nodes having CPU and memory both. Hence it has no impact to HANA.

**Background on empty NUMA node0**

The NUMA node 0 comes historically from a bare metal installation where it was a given to have a sequence of NUMA nodes starting with 0. Later, in first virtualized environments, this was kept as a constant to always represent an anchor point. Today, an empty NUMA node 0 is nothing more than a heritage. Hence, in current implementations troubleshooting an empty node0 must be ignored as it has no functions and no impact as HANA is fully aware of this concept.

Example 1-2 shows the NUMA node 0 has no core or memory associated with it. SAP HANA has 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**

```bash
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 active
NUMA node as 1, in this case as shown in Example 1-3, ignoring NUMA node with no
memory and no CPU.

Example 1-3  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 on page 23 shows 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 what degrades the
performance of SAP HANA (more on POWER8 than POWER9).

In this type of setup, when you run your HANA workload, as processor and memory do not
reside on the same NUMA node, it leads to far memory fetches and higher latency. This can
happen when your Power System is already running multiple LPARs, and newly created
LPARs can access only remaining resources. The hypervisor tries to allocate resources, but fails to provide an ideal setup due to resource constraints. It is always desirable to get the best allocation, but in a few cases it is not possible. Refer to the following corrective action that can be performed.

For a better performance, it is always recommended to have memory and processor cores on the same NUMA node. In multicores setup, you have multiple NUMA nodes with some memory and some cores, in such cases:

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

```
ls300021:~ # numbtxt --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**

### View the LPAR placement from the HMC

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

1. Go to the command line login on the HMC console by way of SSH.
2. On the HMC, check first which resource dumps are already in the dump directory:

   ```
   hscroot@:~> lsdump -h
   dump_type=resource,name=RSCDUMP.109130D.09000001.20140624065824,size=2490992,source_size=0
   dump_type=resource,name=RSCDUMP.109130D.0B000001.20141114162422,size=16128,source_size=0
   ```

3. Issue a `startdump` command:

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

   **Note:** Creating the dump can take a few seconds.

4. Identify the creation of the dump.

Use the `ls -ltr /dump` or with the `lsdump` command to check if the new dump is available:

```
ls300021:~> lsdump -h
```

```
dump_type=resource,name=RSCDUMP.109130D.09000001.20140624065824,size=2490992,source_size=0
dump_type=resource,name=RSCDUMP.109130D.0B000001.20141114162422,size=16128,source_size=0
```

This dump contains binary and human readable data.
5. Identify the dump:

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

This displays a set of data. Verify if in the following table none of the LPAR IDs belonging to a HANA LPAR is spanning the drawers. Or download it to another machine using the following command:

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

After you downloaded and checked the dump file, it is probably a good idea to delete it on the HMC to prevent accumulating garbage on it:

```
hscroot@:~> rmdump -f RSCDUMP.<my Dump ID>
```
Read the HMC resource dump

The resource dump contains beside the binary data human readable tables. Figure 1-12 illustrates the key items to look at specially LPAR 6 (green) and LPAR 8 (red). To understand the table, use the following steps:

Table 1:

<table>
<thead>
<tr>
<th>Domain</th>
<th>Procs</th>
<th>Units</th>
<th>Memory</th>
<th>Proc Units</th>
<th>Memory</th>
<th>Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>SEC</td>
<td>PRI</td>
<td>Total</td>
<td>Free</td>
<td>Total</td>
<td>Free</td>
<td>LP</td>
</tr>
<tr>
<td>--------</td>
<td>-------</td>
<td>-------</td>
<td>--------</td>
<td>-------</td>
<td>------</td>
<td>----</td>
</tr>
<tr>
<td>a</td>
<td>0</td>
<td>3200</td>
<td>400</td>
<td>20</td>
<td>4095</td>
<td>0</td>
</tr>
<tr>
<td>c</td>
<td>0</td>
<td>800</td>
<td>100</td>
<td>20</td>
<td>1024</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>a</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>c</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>b</td>
<td>1</td>
<td>3200</td>
<td>700</td>
<td>0</td>
<td>4095</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>800</td>
<td>0</td>
<td>0</td>
<td>1024</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>800</td>
<td>100</td>
<td>0</td>
<td>1024</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure 1-12  Resource dump

1. Step 1: 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 8 sockets, 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 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 memory (1024 blocks).

2. Step 2: LPAR placement:

▶ LPAR ID 6 - green:

Figure 1-12 on page 25 (Table1) the LPAR is spread across 4 sockets in enclosure 0 (see green lines). The column LP describes the LPAR ID (in this case 6). The reason that this LPAR is divided across the 4 sockets is that it does not fit into one socket (it uses 20 Cores, which is 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.

In case the LPAR is smaller than 8 Cores and 256 GB and there is a socket which can provide those resources, then the hypervisor places the LPAR into one socket.

▶ LPAR ID 8 - red:

This LPAR is a more difficult one as it has been created as the last LPAR, and has an inconvenient size (8 Cores and ~1 TB of memory). When the LPAR was created, the hypervisor had to take the rest of the available resources on the machine, and as a result the LPAR has been scattered across 7 Cores, 2 enclosures and in one case has gotten 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 expect to see 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 works transparent for the applications and can be executed while every LPAR is up and running as long as there is available memory on the machine, and this feature is enabled. To list the servers use the following command:

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

To analyze the current situation, run the command `lsmemopt` 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  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 the perfect rating.

Running the command as shown in Figure 1-13 on page 26 with the option -o calcscore shows to what degree Dynamic Platform Optimizer (DPO) can optimize the LPAR based on the current situation.

To use DPO to optimize the LPARs use the command optmem as follows:

```
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 man pages for optmem).

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

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

**Analyzing and fixing memory placement manually**

In case DPO does not optimize the memory placement to the required extend, 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 then await the hypervisor chooses a better placement when the LPAR is activated again.

This obviously works only if you changed the profile of the LPAR to fit better to the available resources, or if other LPARs occupying the needed resources have been changed.

To get the memory assignment for the LPAR run the command lshwres 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>,0,1032704
```

*Figure 1-14 Command lshwres*

To remove the fixed memory assignment, the command chhwres can be used 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 option `-q` is the quantity to remove and this is the `curr_mem` from the previous list.

The strategy to fix the memory placement is then the following:

1. Shut down every LPAR which needs to be changed.
2. Get the correct memory settings with the command `lshwres` (see Figure 1-14).
3. Remove the memory assignment with the command `chhwres`.
4. Start the LPARs in the correct priority order and await the hypervisor performs 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 a new LPAR on Power Systems:

1. Check the `lpar_currscore` value using the `lsmemopt` command to verify how the LPAR is placed with respect to CPU and memory.
2. When using the `optmem` command, this task must be coordinated with the application owner as some applications such as HANA do not adapt automatically.
3. On new servers always create or import the most relevant LPARs first to give them the best placement.
4. 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 SAP HANA the best practices on LPAR placement are documented in the SAP HANA Advanced Operation Guide as the following website:

http://www.ibm.com/support/techdocs/atsmastr.nsf/WebIndex/WP102502

### 1.8.2 I/O configuration and tuning

When a new SAP system is setup, I/O configurations, storage and network need utmost attention. Performance of an I/O configuration is defined by their bandwidth (throughput when using large packages) and latency (when using many small packages). Both storage and network configurations performance are determined by these attributes. End 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 in general.

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

**Note:** For details about aspects of storage and network I/O configuration including latency, refer to the IBM Power Systems Infrastructure I/O for SAP Applications, REDP-5581 at the following link:


**Network bandwidth test**

Network is always measured in speed at which data can be transferred between two hosts over Internet Protocols (IPs). Each host is recognized with an IP, one host becomes a host and the other becomes a receiver. This is also similar to a client-server architecture. Most of the network measurement tools give results in Mbps or Gbps.
Niping
Niping is an SAP tool to measure line speed of a network device. For more details about nipping, how to use it, and sample output refer to:

SAP NOTE 500235 - Network Diagnosis with NIPING

iperf
iperf is another open source tool which measures network bandwidth. This tool easily checks network bandwidth between a client and a server. To learn more on this tool, refer to the following websites:

https://iperf.fr/iperf-doc.php
https://iperf.fr/

When measure against AIX, validate you have the same iperf version. The results between iperf2 and iperf3 cannot be compared 1 to 1.

Storage I/O bandwidth test
Storage I/O bandwidth test or line speed test needs to be understood as throughput that a storage adapter gives, when attached directly to a system or LPAR. This is measured in:

- **IOPS** (input/output operations per second) is a measurement of performance for hard drives and storage area networks. IOPS represents how quickly a given 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. 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 HWCCT bundle downloadable from the SAP Service Market. For the intend described here, the match to a specific HANA version is not important. Although clients can use it through the SAP intended way by configuring a json file, it can be also executed alone as this allows for more selective test modes and by that faster return. For a sample call of fsperf refer to Appendix A, “fsperf” on page 43.

dd commands
dd command is a command line utility available on Linux. Its primary purpose is to convert and copy files, as per the operands passed to the command. It is a light weight command, which can be used to copy files in any format. But it can also be used to measure line speed and throughput of a storage device.

To measure throughput of a storage device, use the dd command as follows:

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

The command requires (Figure 1-15):

- An input file which is the test file to test throughput with.
- An output file which is a 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.
Figure 1-15 shows a throughput of 1.4 GBps has been measured, for example around 9.6 GBps. A block size of 1 GB had been used. The `dd` command was called with count=10, meaning it will write 10 blocks of 1 GB each, on output file /hana/data/test.

1.9 Online references

These websites are also relevant as further information sources:

- Prepare your Linux for your SAP solution with saptune:
- `sapconf` – A way to prepare a SLES system for SAP workload – Part 1:
- Overview of the Red Hat Enterprise Linux for SAP Applications subscription:
  https://access.redhat.com/solutions/34169
- Installation Quick Start - SUSE Linux Enterprise Server for SAP Applications 15 SP1:
- Installing and setting up Linux on Power Systems servers:

  For HANA on Linux on POWER, tailored documentation has been published through the SAP Launchpad.
  
  - Network Configuration for networks demanding for Jumbo Frames:
    http://www.ibm.com/support/techdocs/atsmastr.nsf/WebIndex/WP102502
  - Resource Optimized High Availability in PowerHA® SystemMirror®
  - SAP note 2535891 - Run your SAP workload best on AIX and POWER8 or POWER9 in SMT8 mode
  - Performance Optimization and Tuning Techniques for IBM Power Systems Processors Including IBM POWER8:
Chapter 2. Monitoring

This chapter describes SAP monitoring components for measuring 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 hypervisor PowerVM and RHV-KVM
2.1 SAP monitoring components

This chapter gives a quick overview of the SAP monitoring components for measuring performance of the underlying operating system and its components like CPU, memory, network and I/O subsystem. Most SAP solutions are communicating with the SAP Host Agent to gather performance matrixes from the operating system. The SAP Host Agent is the focal point for the operating system monitoring, but also 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 using transaction ST06.

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

2.1.1 SAP Host Agent

The SAP Host Agent is a service agent 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 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 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 (CIM) client (SAPCIMC)**
The CIM is an open standard that defines how managed elements in an IT environment are represented as a common set of objects and relationships between them. The Distributed Management Task Force maintains the CIM to allow consistent management of these managed elements, independent of their manufacturer or provider. The SAP Host Agent contains a stand-alone CIM client (sapcimc). The following classes are provided by CIM/operating system plugin specific for Linux on POWER (as of 2019):

- CPUVirtCIMhost.
- CPUVirtCIMhost_HardwareId.
- CPUVirtCIMhost_Model.
- CPUVirtCIMhost_PhysCPUIdle.
The CIM/operating system plugin 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 execute different operations in the SAP application called Webmethods. A Webmethod can have different options, depending on their function. By way of the `saphostctrl` command, it is possible to start, stop, manage, verify, observe and control an SAP application instance on the same or on a different host system. On all hosts running an SAP instance, an SAP Host Agent has to be active. The following command example shows how to extract capacity data from the local system:

```
linux:~ # saphostctrl -function ListOSMetrics -metype VirtualComputerSystem
```

<table>
<thead>
<tr>
<th>Timestamp</th>
<th>Metric</th>
<th>MeasuredElement</th>
<th>Value Unit</th>
</tr>
</thead>
<tbody>
<tr>
<td>2019-05-24-16:21:49</td>
<td>VCS.AMSenable</td>
<td>VirtualComputerSystem</td>
<td>0 Count</td>
</tr>
<tr>
<td>2019-05-24-16:21:49</td>
<td>VCS.GuaranteedCPUCapacity</td>
<td>VirtualComputerSystem</td>
<td>12.00 Count</td>
</tr>
<tr>
<td>2019-05-24-16:21:49</td>
<td>VCS.MemoryMaximum</td>
<td>VirtualComputerSystem</td>
<td>536870912 KiloBytes</td>
</tr>
</tbody>
</table>

**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. The most 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 LPAR type (shared or dedicated).
- LPAR mode like capped or uncapped.
- Entitlement capacity.
- Capacity limits.
The majority of the examples in the following sections are based on SMT2 configured systems to reduce the example output. In SAP environments, it is recommended to use SMT8.

### 2.2.1 IBM Performance Management for Power Linux Systems (PMLinux)

Your IBM Power Systems model (including IBM Power Blades, PowerLinux, POWER based 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 fulfil this task.

This easy-to-use yet powerful tool provides critical resource and performance information about the current and long-term utilization trends; from a total system view to an individual logical partition view. It helps to enable capacity planning for your Power Systems running IBM i, AIX, and Linux (Red Hat, SUSE, Fedora, CentOS, and Ubuntu operating systems) 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 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 a **no additional charge** and a **nominal charge** options depending on the level of detail required.

This is available with minimal initial set up. After initial set up, the remaining process is automated, helping to relieve system administrators of tedious and recurring tasks involved with systems management. As a result IBM provides secure Internet access to the analysis graphs.

In summary, performance management for Power System’s asset management and virtualization capabilities provide a comprehensive step to manage today’s environment, and at the same time planning for tomorrow’s challenges and opportunities.

### 2.2.2 Utilities for PowerPC hardware

The RPM package `powerpc-utils` contains utilities required for 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, these must be run as root. The command line tools `amsstat`, `lparstat` and `ppc64_cpu` are explained in more detail in this section.

The following is a reference entry point to download the correct `powerpc-utils` package for all supported hypervisors and Linux distributions:


**Active Memory Expansion (AME)**

AME can reduce the amount of memory used for NetWeaver LPARs (in the double digit range). This tool consist of two main areas planning and viewing:

- Active Memory Expansion Planning and Advisory Tool amepat
AME stats can be found in 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.

This program captures memory statistics relevant in an active AMS environment. This tool can run once or set to run repeated with a specified timeout, in seconds, between instances of data collection. For most SAP Solutions, it is recommended to disable the AMS feature. Only for some restricted configurations of SAP NetWeaver, AMS can be enabled. SAP HANA does not support AMS. To check if AMS is enable or disabled run (as root user) `amsstat` without any parameter:

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

The command is only supported with 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 (lparstat)**

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. Requirement for valid LPAR statistics is a working 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 and example of `lparstat` with 5 iterations and 2 seconds interval time. To reduce the size of the output an SMT level of 2 has been 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  52.36  4873155174  28.36  4873207632  26028487
11.22  16.90  0.00  71.88  0.524164  26.20  4873180426  28.12  4873234255  26028487
09.31  19.25  0.00  71.45  1.048336  52.41  4873207632  28.55  4873207632  26028487
10.94  17.74  0.00  71.32  0.524170  26.20  4873234255  28.12  4873234255  26028487
10.48  18.06  0.00  71.46  0.524168  26.20  4873261114  28.54  4873261114  26028487

```

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

```
Example 2-1  Output command lparstat

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 CPU 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

The command is only supported with 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
```

CPU characteristics of PowerPC systems (ppc64_cpu)
The ppc64_cpu command is used to display and modify CPU characteristics like SMT level, DSCR settings, run-mode, and subcore settings on the POWER platform. The following command shows which logical CPUs are located 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*
```

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
```
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 (nmon)

The program `nmon` (Nigel's monitor) 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 the following website:

http://nmon.sourceforge.net/pmwiki.php

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 version 16g, the PowerVM LPAR section is available. When running on the PowerVM hypervisor `nmon` collects comparable stats on AIX and Linux. Figure 2-3 shows the statistics of a 2 core LPAR in dedicated mode running with SMT2 level.

![nmon performance monitor](image)
Figure 2-4 shows the `nmon` output of a shared processor pool LPAR.

![nmon performance monitor output - shared processor pool LPAR](image)

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 hypervisor 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 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 using SAP monitoring if these are running under Red Hat Enterprise Linux or SUSE Linux Enterprise Server. The available metrics however are depending on the capabilities of the hypervisor (PowerVM or RHV-KVM).

Table 2-1 on page 40 lists the virtualization related metrics gathered by the individual hypervisors.
Within the output of transaction ST06 as shown in Figure 2-1 on page 32, the information shown in Table 2-1 can be found in the CPU Virtualization Host and the CPU Virtualization Virtual System as follows.

The CPU Virtualization Host shows processor information 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.
- **Pool Utilization Authority**: (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:

### Table 2-1 Comparison between hypervisors and metrics

<table>
<thead>
<tr>
<th>Metric</th>
<th>PowerVM dedicated LPAR</th>
<th>PowerVM Shared Processor LPAR with Pool Utilization Authority (PUA) not granted</th>
<th>PowerVM Shared Processor LPAR with PUA granted</th>
<th>RHV-KVM guest</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Processor</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Pool ID</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pool Utilization Authority</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pool CPUs</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pool CPUs Idle</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Current Processor Frequency</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Virtual CPUs</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>SMT Mode</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Threads</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Partition Name</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Partition ID</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Partition Type</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Capped</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Weight</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Capacity Consumed</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Guaranteed Capacity</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Guaranteed Capacity Consumed</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Available Capacity</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Available Capacity Consumed</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Additional Capacity Available</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Capacity Maximum</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
– 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 to.

- **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 pool utilization authority is granted.

The CPU Virtualization Virtual System shows processor information 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 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 if Simultaneous Multi-Threading (SMT) is active. Possible values are:
  – On.
  – Off.

- **Threads**: The number of active SMT threads. For PowerVM LPARs, this 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 logical partition. 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 actual 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 actual 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 only available if the pool utilization authority 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 only available if the pool utilization authority 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.
This appendix discusses the fsperf open source file system performance tool.

This appendix contains the following:

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

This open source tool is designed as a file system performance testing framework. For more information refer to the following website:

https://github.com/josefbacik/fsperf

Usage information for fsperf is available by executing 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. -t all means to run all the test mode, and -m all means test for both latency and throughput, though the default is only throughput.

The following is a sample output of the run, with file size of 16 GB (specified with -f flag), block size of 16K (-b flag), -m all (running all the performance measurements), kernel I/O queue being 2K, max parallel I/O request is also 2K, and queue length is 8. /hana/data/dummy is the file system to be tested. This file system is backed with a Fibre Channel adapter and the disk 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 line 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 executed 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

Additional program parameters:
size_kernel_io_queue=2048
max_parallel_io_requests=2048
num_submit_queues=8

Running tests

The configuration parameters set are:
async_write_submit_active=auto,async_write_submit_blocks=all,async_read_submit=on,
um_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 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:
I/O time:....................... 647.863 s (Throughput: 25.2892 MB/s)
Latency:........................ 617 us

Results of 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 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 of what test is running and completed at a particular instance. At the end, it gives out a summary of throughput and latency, for all test modes run.

These KPIs numbers, can be matched with what IBM or SAP recommend for their applications, and hence tuned accordingly for better performance.
Related publications

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

IBM Redbooks

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

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

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

[ibm.com/redbooks](http://ibm.com/redbooks)

Online resources

These websites are also relevant as further information sources:

- Prepare your Linux for your SAP solution with saptune
- sapconf – A way to prepare a SLES system for SAP workload – Part 1
- Overview of the Red Hat Enterprise Linux for SAP Applications subscription
  [https://access.redhat.com/solutions/34169](https://access.redhat.com/solutions/34169)
- Installation Quick Start - SUSE Linux Enterprise Server for SAP Applications 15 SP1
- Installing and setting up Linux on Power Systems servers
- Resource Optimized High Availability in PowerHA SystemMirror
Help from IBM

IBM Support and downloads
ibm.com/support

IBM Global Services
ibm.com/services