SAP HANA on IBM Power Systems

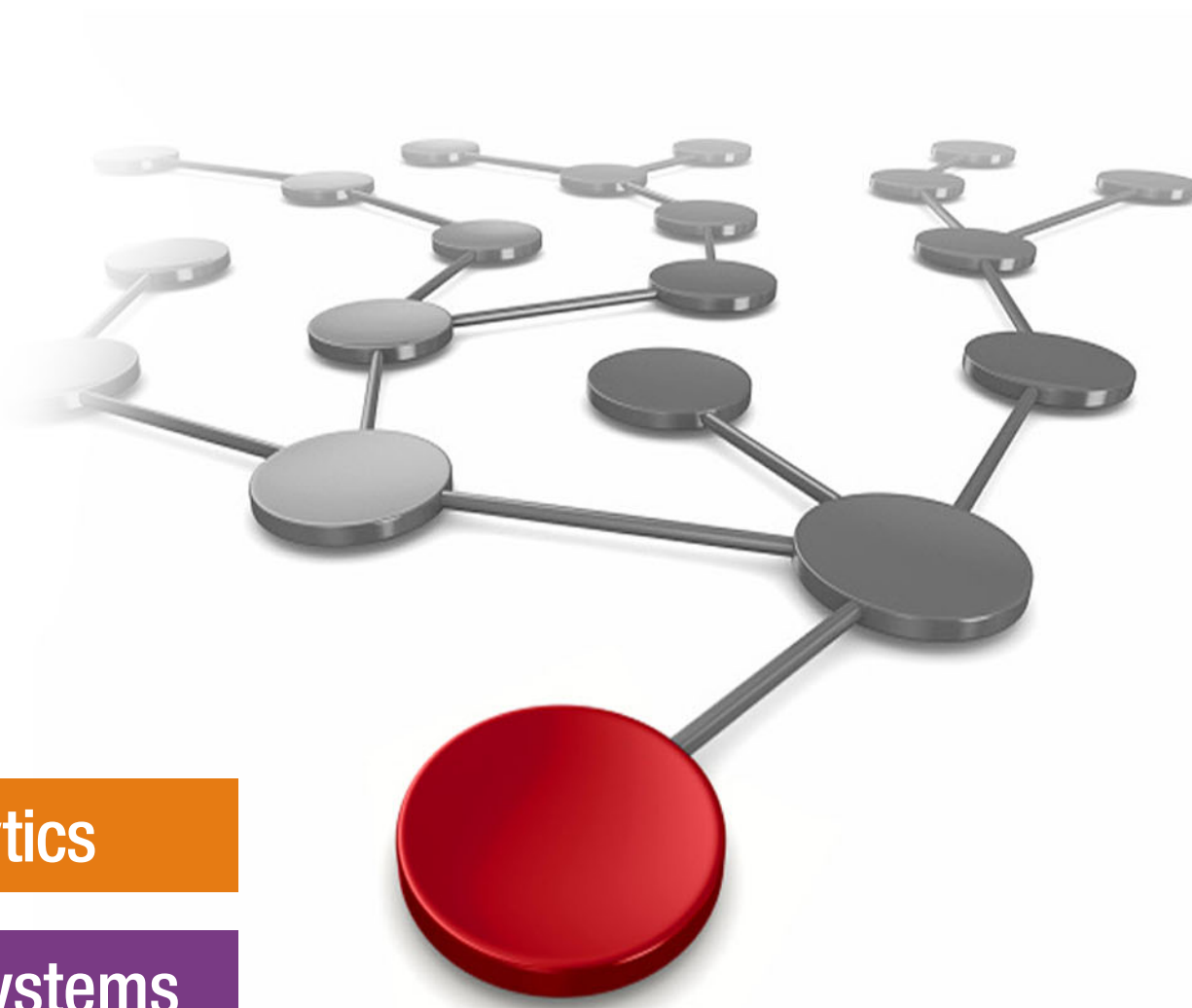## High Availability and Disaster Recovery Implementation Updates

Dino Quintero

Luis Bolinches

Rodrigo Ceron

Mika Heino

John Wright

Analytics

Power Systems

IBM

International Technical Support Organization

SAP HANA on IBM Power Systems: High Availability and Disaster Recovery Implementation Updates

July 2019

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

First Edition (July 2019)

This edition applies to Red Hat Enterprise Linux V7.5, PowerHA SystemMirror for Linux V7.2.2.2, and SAP HANA V2.0.

Contents

Notices

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

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

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

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

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

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

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

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

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

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

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

COPYRIGHT LICENSE:

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

Trademarks

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

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

AIX®	IBM Spectrum™	POWER9™
DB2®	IBM Spectrum Accelerate™	PowerHA®
Db2®	IBM Spectrum Scale™	PowerVM®
Enterprise Storage Server®	IBM Spectrum Virtualize™	Redbooks®
GPFS™	POWER®	Redbooks (logo) ®
IBM®	POWER Hypervisor™	Storwize®
IBM Elastic Storage™	Power Systems™	SystemMirror®
IBM FlashSystem®	POWER8®	XIV®

The following terms are trademarks of other companies:

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

Linux is a trademark of Linus Torvalds in the United States, other countries, or both.

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

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® Redbooks® publication updates *Implementing High Availability and Disaster Recovery Solutions with SAP HANA on IBM Power Systems*, REDP-5443 with the latest technical content that describes how to implement an SAP HANA on IBM Power Systems™ high availability (HA) and disaster recovery (DR) solution by using theoretical knowledge and sample scenarios.

This book describes how all the pieces of the reference architecture work together (IBM Power Systems servers, IBM Storage servers, IBM Spectrum™ Scale, IBM PowerHA® SystemMirror® for Linux, IBM VM Recovery Manager DR for Power Systems, and Linux distributions) and demonstrates the resilience of SAP HANA with IBM Power Systems servers.

This publication is for architects, brand specialists, distributors, resellers, and anyone developing and implementing SAP HANA on IBM Power Systems integration, automation, HA, and DR solutions. This publication provides documentation to transfer the how-to-skills to the technical teams, and documentation to the sales team.

Authors

This book was produced by a team of specialists from around the world working at the International Technical Support Organization, Poughkeepsie Center:

Dino Quintero is an IT Management Consultant and an IBM Level 3 Senior Certified IT Specialist with IBM Redbooks in Poughkeepsie, New York. 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 holds a Master of Computing Information Systems degree and a Bachelor of Science degree in Computer Science from Marist College.

Luis Bolinches has been working with IBM Power Systems servers for over 16 years and has been working with IBM Spectrum Scale™ (formerly known as IBM General Parallel File System (IBM GPFS™) for over 10 years. He works 50% of his time for IBM Lab Services in Nordic, where he is the subject matter expert (SME) for HANA on IBM Power Systems, and the other 50% on the IBM Spectrum Scale development team.

Rodrigo Ceron is an IBM Master Inventor and Senior Managing Consultant at IBM Lab Services and Training. He has 19 years of experience in the Linux and UNIX arena, and has been working for IBM for over 15 years, where he has received eight intellectual property patents in multiple areas. He graduated with honors in Computer Engineering from the University of Campinas (UNICAMP) and holds an IEEE CSDA credential. He is also an IBM Expert Certified IT Specialist. His responsibilities are to engage customers worldwide to deliver highly specialized consulting, implementation, and skill transfer services in his areas of expertise: cognitive and artificial intelligence, SAP HANA, IBM Spectrum Scale, Linux on Power, systems HA, and performance. He has also been fostering business development by presenting these topics at IBM conferences globally, and writing technical documentations. He has written seven IBM Redbooks publications so far, awarding him the tile of ITSO Platinum author.

Mika Heino is a Client Technical Specialist working in IBM Lab Services in IBM Finland for local and international IBM accounts. He has a degree in Telecommunications and Computer Science from Turku University of Applied Sciences. Mika has 20 years experience with Intel and IBM Power Systems with Linux servers, AIX® and IBM i servers, and server virtualization for both Intel and IBM POWER® processor-based servers. He has more than 10 years of experience with storage area networks (SANs), IBM Storage Systems servers, and storage virtualization.

John Wright is a Technical Design Architect at Pure Storage. With over a decade of his 19 years of experience spent at IBM, John has a deep and varied skillset that was gained from servicing multiple industry sectors across multiple vendor technologies. He specializes in cloud (Amazon Web Services (AWS), OpenStack, and IBM PowerVC), Pure Storage products, analytics (SAP HANA on IBM Power Systems and Hortonworks Data Platform on Power Systems), and SUSE Linux. He has a background in traditional AIX and virtualization environments, including complex data center migrations and hardware refresh projects. He holds certifications with AWS and Pure Storage. John splits his time between delivering services, designing new solutions that use the latest technology, and running onsite workshops across the UK and Europe.

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 ITSO 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 books to be as helpful as possible. Send us your comments about this book 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, International Technical Support Organization
   Dept. HYTD Mail Station P099
   2455 South Road
   Poughkeepsie, NY 12601-5400

Stay connected to IBM Redbooks

► Find us on Facebook:

   http://www.facebook.com/IBMRedbooks

► Follow us on Twitter:

   http://twitter.com/ibmredbooks

► Look for us on LinkedIn:

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

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

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

► Stay current on recent Redbooks publications with RSS Feeds:

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

1

Introduction

This chapter describes the goals of this publication, the contents that are covered, key aspects of the SAP HANA solution on IBM Power Systems servers, and what is new since the last publication of this book.

This chapter covers the following topics:

► About this publication
► The SAP HANA platform
► High availability for SAP HANA

1.1 About this publication

This book was written and updated by expert IBM consultants who have experience installing SAP HANA on IBM Power Systems servers based on hundreds of past installations at customers worldwide. The contents of this book follow the best practices that were developed according to SAP recommendations.

This publication provides you with all the information that you need to help you avoid issues with your SAP HANA on IBM Power Systems implementation. The goals of this publication are:

► To be a practical guide for the most common SAP HANA on IBM Power Systems landscapes.

► To inform you of all the SAP directives for HANA on a Tailored Datacenter Integration (TDI) architecture so that the environment is fully supported by SAP.

► To suggest best practice standards for SAP HANA on IBM Power Systems implementations around the globe.

For more materials to complement this publication, see the *SAP HANA Administration Guide*.

The SAP HANA TDI architecture helps you to build an environment by using your existing hardware, such as servers, storage, storage area networks (SANs), and network switches. This architecture gives you freedom over the SAP HANA appliance model that was widely used in the past. However, you must follow the SAP list for the supported hardware models and configurations.

> **Note:** SAP HANA TDI must be performed by TDI certified personnel.
>
> For more information, see the SAP HANA Tailored Datacenter Integration (TDI) Overview and SAP HANA Tailored Datacenter Integration - Frequently Asked Questions.

Although SAP allows flexibility in a TDI implementation of HANA, there is a set of configuration and settings that work best for a HANA on a Power Systems implementation. This configuration and these settings are seen by many customers as the ones that bring the most performance and stability, less management and maintenance effort, and better understanding of the environment. That is why they are called *best practices*.

The audience of this publication consists of the following groups:

► Customers, IBM Business Partners, and IBM consultants planning and installing HANA on IBM Power Systems.

► System administrators managing the installed HANA systems.

1.2 The SAP HANA platform

There are various answers that you can give to the question "What is SAP HANA?" However, the answer that can be emphasized is that *SAP HANA is an SAP solution*.

This is a simple but important definition. As shown in 2.1, "SAP requirements for SAP HANA on IBM Power Systems implementations" on page 14, the core aspects of your HANA implementation are defined by SAP guidelines and requirements. Factors such as supported operating systems (OSes), core to memory ratios, allowed server hardware, allowed storage hardware, networking requirements, and the HANA platform announcements roadmap, are determined by SAP.

SAP HANA is the SAP database (DB) platform for multiple SAP solutions. In changing direction to former classic SAP solutions, HANA is the processing core of it all. Operations that formerly were performed at application layers moved into the DB layer and are now performed by the HANA engines.

This changed the way that data was traditionally processed by using Online Transactional Processing (OLTP), which gave way to the more dynamic Online Analytical Processing (OLAP) or a mixed schema, which required a solution that could work with both types of data processing. SAP was able to combine the processing of these two schemes because it concluded that many similarities existed in both types of processing. The result was a single DB able to use the same source of data for performing both kinds of operations, thus eliminating the need for time-consuming extraction, transformation, and loading (ETL) operations between an OLTP base into an OLAP base. SAP HANA is built to work with both OLTP and OLAP data.

Traditionally, DBs store data by *rows*, with a data entry in each *column*. So, retrieving the data means that a read operation on the entire row is required to build the results of a query. Therefore, many data entries in the columns of a particular row are also read. However, in today's world of analytical processing, the user is interested in building reports that provide an insight into a vast amount of data, but is not necessarily interested in knowing all of the details about that data.

Reading numerous columns of a row to create an answer for an aggregation report that targets only some of the columns' data is perceived as a waste of I/O time because many other columns that are not of interest are also read. Traditionally, this task was minimized by the creation of index tables that lowered the amount of I/O at the expense of consuming more space on disk for the indexes. SAP HANA provided a solution to this issue by storing data in columns as opposed to rows. Analytical processing greatly benefits from this change. Nevertheless, SAP HANA can work with both columnar and row data.

Sparsity of data is an aspect that has been treated by computer scientists since the early days of computing. Countless data structures were proposed to reduce the amount of space for storing sparse data. SAP HANA can potentially reduce the footprint of used memory by applying compression algorithms that treat this sparsity of data and also treat default data values.

SAP HANA works by loading all of the data into memory, which is why it is called an *in-memory DB*. This is the most important factor that allows SAP HANA to run analytical reports in seconds as opposed to minutes, or in minutes as opposed to hours, allowing real-time analysis of analytical data.

In summary, these characteristics of SAP HANA allow SAP to strategically place it at the core of its solutions. SAP HANA is the new platform core for all SAP applications.

1.2.1 What is new in SAP HANA on IBM Power Systems

Some concepts that are mentioned here are not entirely new, but they are new to this publication. Here are the concepts that were added:

► There is SAP HANA support of the new IBM POWER9™ series of servers: IBM Power System S922, IBM Power System H922, IBM Power System S924, IBM Power System H924, and IBM Power System L922. There is also support for all IBM POWER9 models that are based on PowerVM®. This gives you a choice of either IBM POWER8® or POWER9 servers to use for hosting your HANA environment. For more information, see 2.1.1, "Storage and file system requirements" on page 15.

► Red Hat Enterprise Linux V7 is now a fully supported OS for SAP HANA on IBM Power Systems. So, now you have a choice of using either SUSE Linux Enterprise Server or Red Hat Enterprise Linux and you can choose the one with which your company IT development and operations department is more familiar.

► At the high availability (HA) level, HANA can now handle *Invisible Takeover* under an SAP HANA System Replication (HSR) scenario. This works only for *read-only* transactions, where the sessions that were connected to the primary system are restored on the secondary one. Nevertheless, cluster management software at the OS layer is still required for managing failover of the virtual IP address (VIPA). Figure 1-1 shows the whole mechanism of Invisible Takeover.
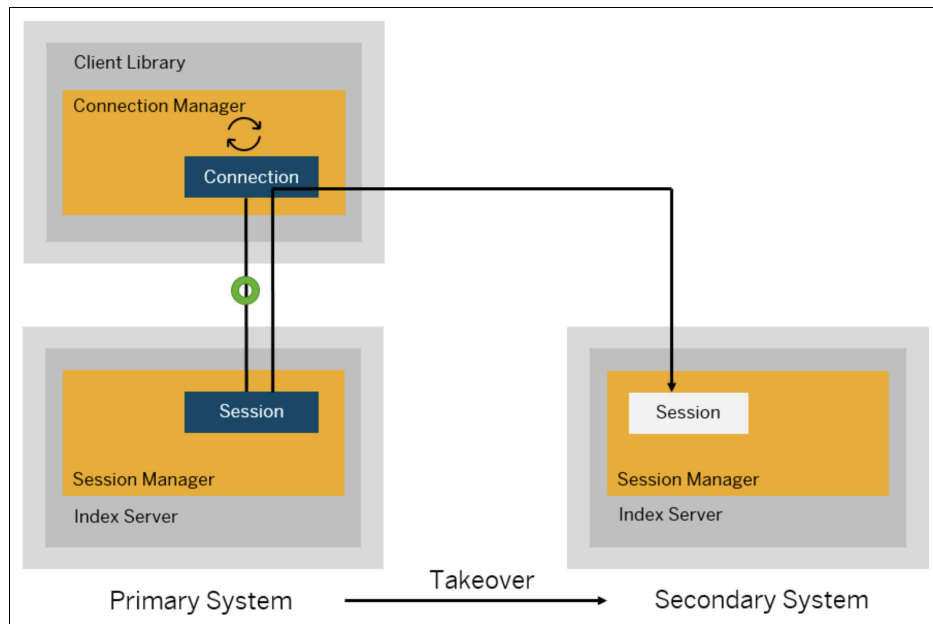
Figure 1-1   SAP HANA System Replication Invisible Takeover

► A same source node can directly replicate to multiple systems without needing to chain the replication along the way. This is called *Multitarget Systems Replication*. Figure 1-2 on page 5 shows this concept, where *server A* is the source of data replication for both *servers B* and *C*.
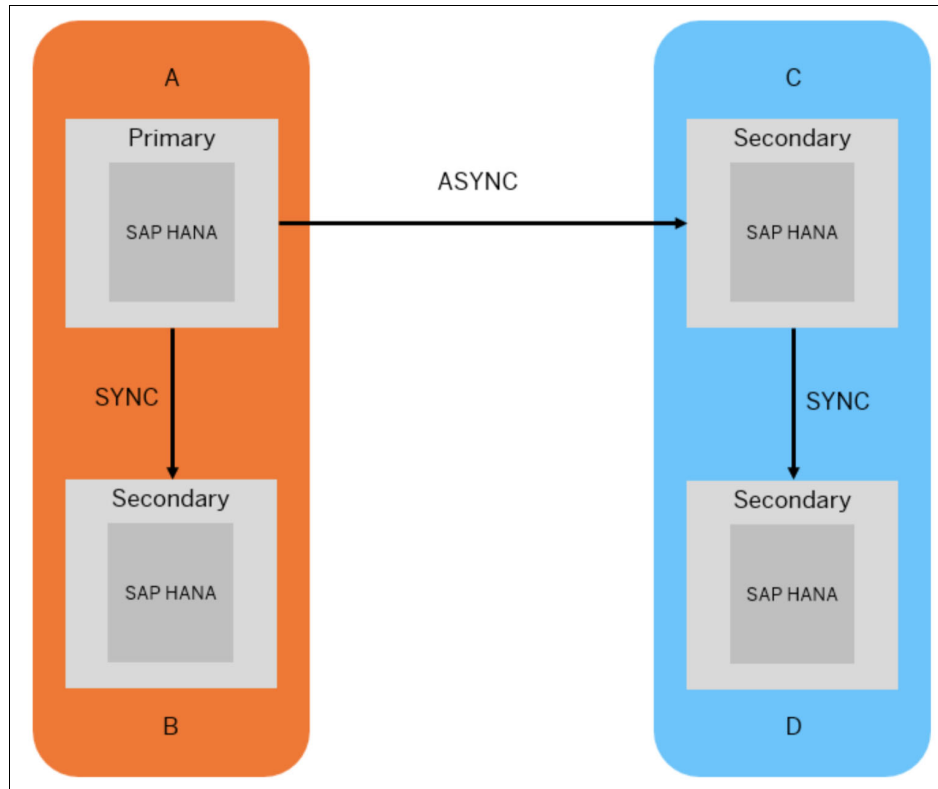
Figure 1-2   SAP HANA System Replication: Sample Multitarget Systems Replication

► From an HA management point of view, there is now a selection of cluster managers to use with SAP HANA on IBM Power Systems:

– SUSE Linux HA (using the SAPHanaSR resource agent)

– Red Hat Enterprise Linux for SAP Applications HA

– PowerHA SystemMirror for Linux

> **Note:** There are also other cluster vendors. The vendors that are listed are the dominant ones among others.

For more information about SAP HANA, see the following websites:

– IBM Power Systems for SAP HANA

– SAP HANA solutions on IBM Power Systems

► With the *Secondary Time Travel* mechanism, you can place the secondary system in online mode and have it load replicated data to a point in the past. With this function, you can easily recover data that was accidentally deleted on the primary system. In order for this function to work, the replication modes must be either `logreplay` or `logreplay_readaccess`. You can control the amount of change history to keep by using the `timetravel_max_retention_time` parameter in `global.ini`. Make sure that the secondary system data and log areas have enough space to handle the amount of time travel data that you want to handle.

1.3 High availability for SAP HANA

The costs of downtime have increased over time, so companies are paying more attention to HA today than in the past. Also, the costs for HA solutions have decreased considerably in a way that makes much more sense to invest in protecting business continuity than to undertake the downtime costs.

No one has 100% business continuity, which is why SAP HANA on IBM Power Systems offers HA and disaster recovery (DR) solutions. Figure 1-3 shows the possible scenarios for HA and DR that you can implement. This publication focuses on HANA and SUSE Linux Enterprise Server, SAP HANA, Red Hat Enterprise Linux and PowerHA SystemMirror, and IBM VM Recovery Manager DR mechanisms to provide HA. Other alternatives are documented in SAP Note 2407186.
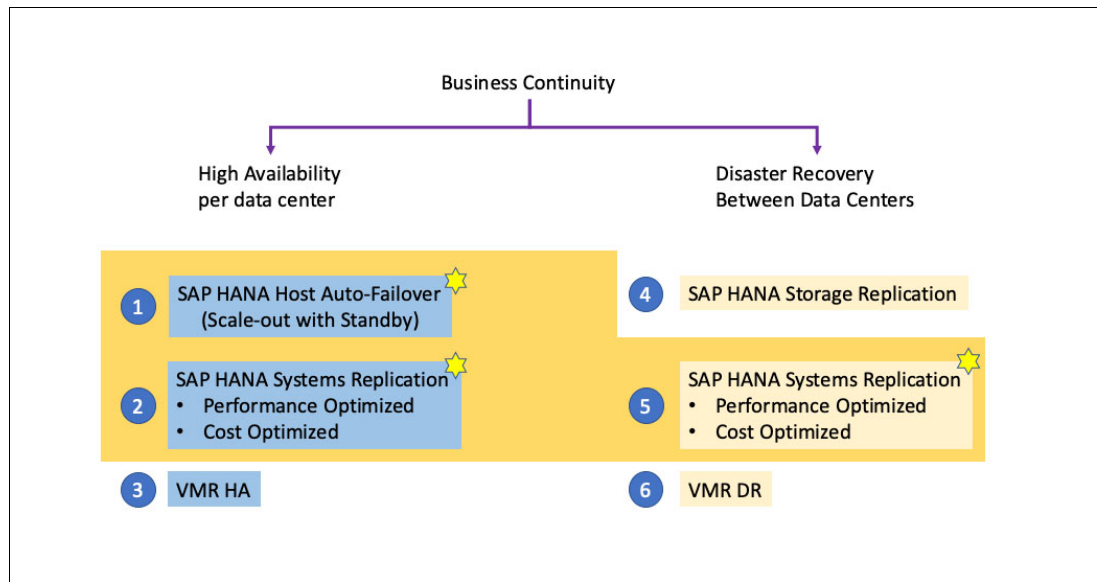
Figure 1-3   Available SAP HANA high availability and disaster recovery options

> **Note:** The numbers in Figure 1-3 represent scenarios and not numbered steps.

From a business continuity perspective, you can protect your systems by creating a local HA plan to ensure the minimum recovery time objective[1] (RTO) possible, and also protect your business from a complete site failure (DR). Scenarios 1, 2, and 3 in Figure 1-3 refer to HA, and scenarios 4, 5, and 6 refer to DR.

In scenario 3, the failed virtual machines (VMs) are restarted on an adjacent server. This is a shared storage topology that is used with Live Partition Mobility (LPM). The secondary server or partition is inactive until the VMs are restarted (booted up) on it or if LPM is used for a planned outage event. There is one physical copy and one logical copy. This solution is outside of the scope of this publication.

---

[1] The amount of time that it takes you to bring your system back online after a failure.

Scenario 4, which is based on bare storage hardware replication mechanisms, is out of the scope for this publication. In this scenario, build a HANA environment the same way as you build the primary system, and leave the secondary system turned off. Only the HANA data and log areas are replicated because each site instance has its own boot disk and HANA binaries disk (`/hana/shared`). The RTO is the highest of all solutions, as shown in Figure 1-3 on page 6, because a full start and mount of the DB happens, and the recovery point objective[2] (RPO) is almost zero, but not zero.

IBM Geographically Dispersed Resiliency (GDR) is also a solution that can be used for HANA DR[3].

Scenario 6 is based on the replication of the VMs to a remote location. IBM Geographically Dispersed Resiliency for Power Systems is the former product name. The DR server is inactive until the replicated VMs are restarted on it. If the production system fails (or tested for DR compliance), the VMs are restarted on a secondary system in the cluster. There are two physical copies of the VMs and one logical copy in this particular configuration. This solution is outside of the scope of this publication.

1.3.1 Disaster recovery: SAP HANA System Replication

This section describes how to create a DR environment by using only an SAP HANA mechanism for data replication: HSR. Figure 1-4 summarizes how this mechanism works, which is the basis for all HSR versions, such as active/passive (performance and cost) optimized, active/active and transparent failover, and multisite replication scenarios.
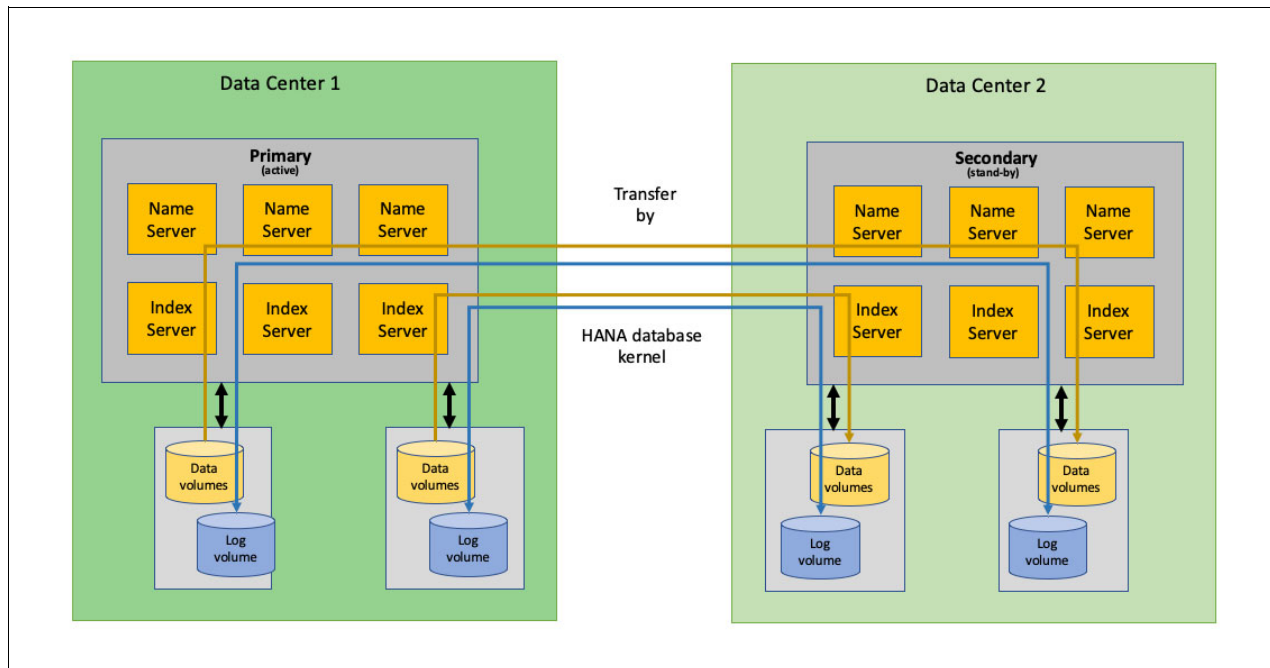
Figure 1-4   SAP HANA System Replication for Disaster Recovery scenario

---

[2] The amount of data that is lost in a failure. Resilient IT systems attempt an RTO of 0.
[3] *IBM Geographically Dispersed Resiliency for Power Systems enables IBM POWER users to reliably realize low recovery times and achieve recovery point objectives*.

In essence, there is one HANA instance at the primary site and another one at the secondary site. Each has their own independent storage areas for the HANA data, log, and shared areas. In this DR scenario, the DR site has a fully duplicated environment for protecting your data from a total loss of the primary site. So, each HANA system has its own IP address, and each site has its own SAP application infrastructure pointing to that site's HANA DB IP address.

The system replication technology within SAP HANA creates a unidirectional replication for the contents of the data and log areas. The primary site replicates data and logs to the secondary site, but not vice versa. The secondary system has a replication receiver status (*secondary system*), and can be set up for read-only DB access, thus not being idle.

If there is a failure in the primary site, all you need to do is perform a *takeover operation* on the secondary node. This is a DB operation that is performed by the basis team and informs the secondary node to come online with its full range of capabilities and operate as a normal, and independent instance. The replication relationship with the primary site is broken. When the failed node comes back online, it is outdated in terms of DB content, but all you need to do is create the replication in the reverse order, from the secondary site to the primary site. After your sites are synchronized again, you can choose to perform another takeover operation to move the DB back to its original primary site.

According to *SAP HANA Network Requirements*, it is a best practice to have a dedicated network for the data replication between the nodes so that HSR does not compete for bandwidth with the data network. In DR implementations, the distance between the primary and DR data centers can be rather long, so the replication is done asynchronously.

According to *SAP High Availability Guide*, this scenario provides an RPO = 0 (synchronous replication) and a low to medium RTO.

1.3.2 High availability: SAP HANA Host Auto-Failover

In this scenario, denoted as 1 in Figure 1-3 on page 6 (note that the numbers in the figure represent scenarios and not numbered steps), the HA of the HANA system is built within the HANA software stack itself. There are no OS tools or extra software that are involved here. Controlling the HA mechanisms for heartbeating, failover, and master, worker, and standby roles is decided by HANA.

This scenario builds a real HANA cluster where the DB itself knows it is working as a cluster, as shown in Figure 1-5.
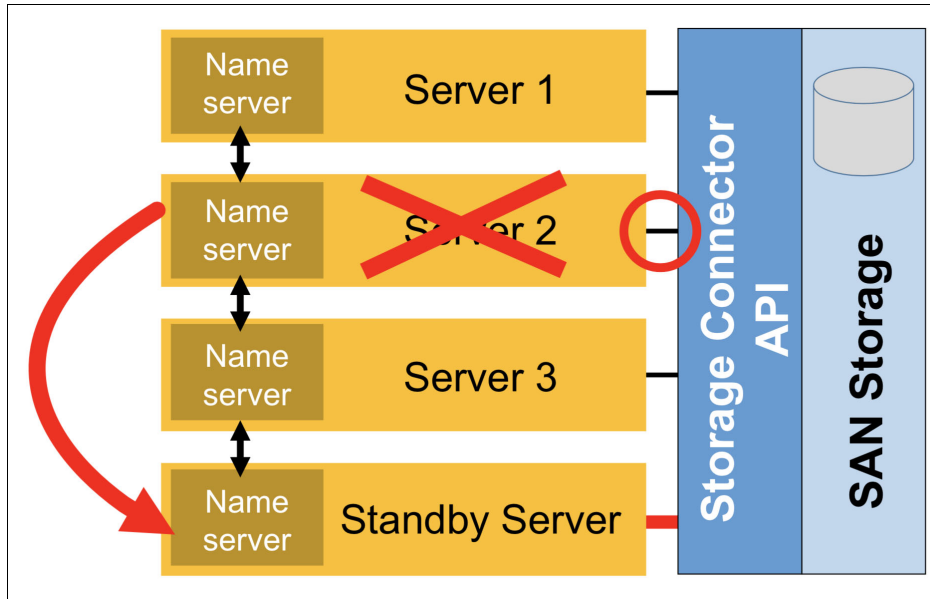
Figure 1-5   HANA scale-out architecture for storage area network deployments (shared disk)

Note: The typical difference between SAN and network-attached storage (NAS) is that an NAS is a single storage device that operates on data files, and SAN is a local network of multiple devices that operate on disk blocks. However, to connect to a SAN, you must have the server class devices with SCSI Fibre Channel.

Each node has its own boot disk. The HANA data and log disks are either assigned to all nodes as shared disks by using the storage connector API to ensure that no two nodes access the same disk at the same time, or shared among the nodes as data and log file systems that use a TDI-supported file system such as Network File System (NFS) or IBM Enterprise Storage Server. Additionally, a third area, the HANA shared file system, is shared among all nodes either through NFS or IBM Spectrum Scale in both deployment options. Also, this architecture needs a dedicated, redundant, and low-latency 10 Gbps Ethernet or InfiniBand network for the HANA nodes to communicate as a cluster environment, which is called the *internode communication network*.

Note: Internode communication cannot run over InfiniBand. You need a minimum of 10 Gb bandwidth that is tuned according to Recommendations for Network Configuration. For filers, you need either InfiniBand (recommend 56 Gbps) or Ethernet. When it comes to Ethernet, the new deployments do not use 10 Gbps. A best practice is to use 40 Gbps single root input/output virtualization (SR-IOV) (no LPM).

This scenario has a master node, a set of worker nodes, and a set of standby nodes. The most common implementations have just one standby node, so the HANA cluster can handle the failure of a single node of either given node type. More standby nodes are required to handle simultaneous node failures.

Whenever a worker node fails, the services on the failed node are taken over by a standby node, which also reloads the portion of the data on the failed node into its memory. The system administrator does not need to perform any manual actions. When the failed node rejoins the cluster, it joins as a standby node. If the master node fails, one of the remaining worker nodes takes over the role as master to prevent the DB from being inaccessible, and the standby comes online as a worker node. For a comprehensive description about how failover occurs, see *SAP HANA Host Auto-Failover*.

In the event of a node failure, the SAP application layer uses a load-balancing configuration to allow any node within the cluster to take on any role. There is no concept of virtual IP addresses for the HANA nodes. Explaining how to set up the application layer for this particular environment is out of the scope for this publication.

> **Note:** There is an option that is supported by SUSE High Availability Extension (HAE) to combine Host Auto-Failover for local HA with HSR for DR. The virtual IPs are an optional step or you can give the application servers a list of candidates to check (which is more effort to maintain).

According to *SAP High Availability Guide*, this scenario provides an RPO=0 and a medium RTO.

From a cost point of view, the standby nodes use all of their entitled processor and memory resources and stay idle until a failover happens. The only room for cost optimization here is to use dedicated donating processors in logical partitions (LPARs). Memory cannot be cost-optimized. Also, in scale-out clusters with less than 2 TB of memory per node, no data is handled by the master node, thus requiring an extra worker node.

1.3.3 High availability: SAP HANA System Replication

This scenario applies to both scale-up and scale-out architectures. However, this publication focuses on the scale-up architectures only.

You can think of this scale-up architecture as a two-node active/stand-by environment. This scenario is what most SAP customers are used to when using other DBs other than HANA, for example, a two-node active-passive SAP + IBM DB2® DB that is controlled by PowerHA SystemMirror on AIX. It is most likely that these users migrate to HANA and apply this kind of architecture to their new HANA environment.

Figure 1-6 depicts this scenario.
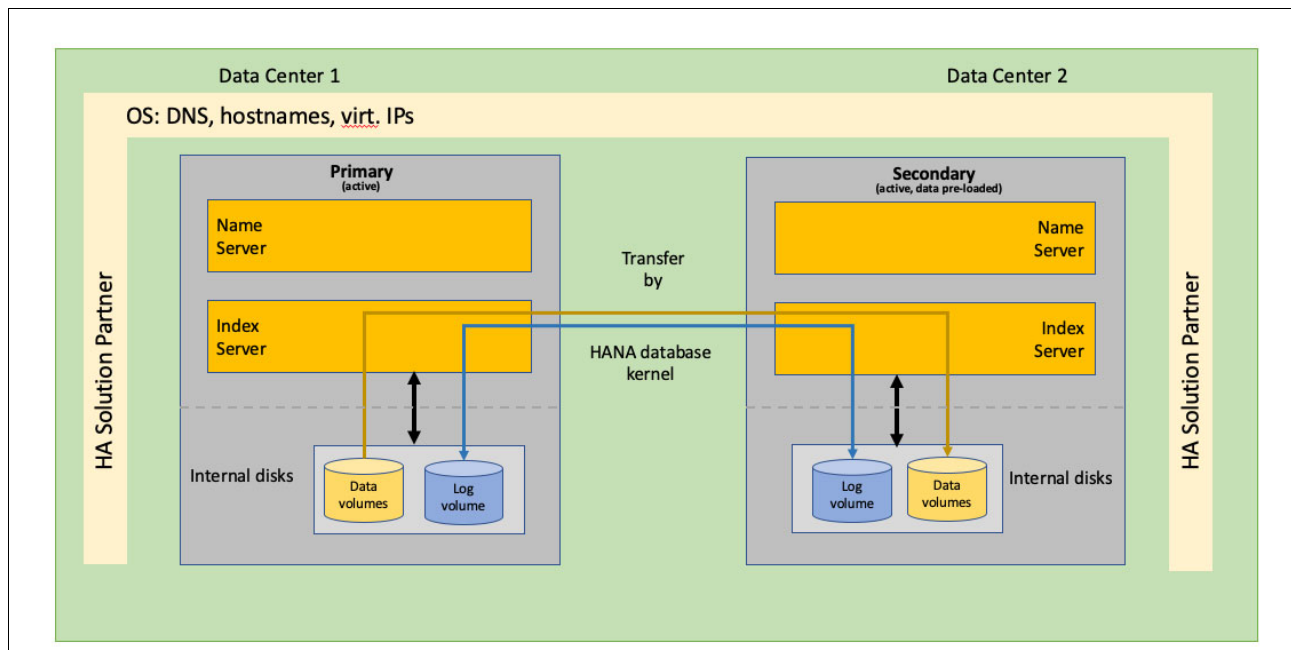
Figure 1-6   Two-node HANA scale-up with SAP HANA System Replication plus SUSE Linux HA

This scenario shows two independent HANA systems, where one system is the primary system and the other is the secondary system. The primary system is in active mode and replicates data by using SAP HANA System Replication to the secondary system, which is in a passive/stand-by mode. The secondary instance can also be in read-only mode. Different from replication for DR, in this HA scenario the replication is synchronous, which ensures an RPO of zero and a low RTO.

Each supported OS, SUSE Linux and Red Hat Enterprise Linux, have their own mechanisms to create and manage the cluster at the OS level. Those mechanisms are defined as *HA Solution Partner* in Figure 1-6.

Compared to the HA scenario that is described in 1.3.2, "High availability: SAP HANA Host Auto-Failover" on page 8, this design does not use a network for HANA inter-node communication, but instead uses a separate network for *replicating* the data from one node to the other. Even though you can replicate data through the existing data network, use a dedicated, redundant network based on 10 Gbps technologies to avoid competing for bandwidth on the data network. Our best practices throughout this publication use a dedicated network for data replication.

Important: As data is replicated from the source system to the destination system by using HSR, you need twice as much space for the HANA data, log, and shared areas because the disks are not shared between the two nodes, and each node has its own disks.

According to the *SAP HANA High Availability Guide*, this scenario provides an RPO=0 and a low RTO, being the most preferred HA architecture by SAP.

2

Planning your installation

This chapter provides an overview of the most important SAP HANA on IBM Power Systems requirements based on published SAP Notes. This chapter guides you through what you need to know in terms of hardware infrastructure and software.

This chapter covers the following topics:

- ► SAP requirements for SAP HANA on IBM Power Systems implementations
- ► Preparing your software

2.1 SAP requirements for SAP HANA on IBM Power Systems implementations

The following sections explain the SAP requirements for SAP HANA on IBM Power Systems implementations. Each requirement is illustrated by an official SAP Note, which is published and updated by SAP. As SAP Notes are constantly updated, always check them before implementing SAP HANA on IBM Power Systems, no matter how familiar you are with them.

> **Hint:** SAP Notes change constantly. Validate all notes before you start implementing your HANA environment because SAP guidelines and statements change frequently. For SAP Notes, see SAP ONE Support Launchpad.

It is a best practice to read the release notes for familiarity of features and requirements. Table 2-1 shows a summary of some important SAP Notes to which you must pay special attention.

Table 2-1   SAP Notes that are related to SAP HANA on IBM Power Systems implementations

SAP note	Title
2055470	*SAP HANA on IBM Power Systems planning and installation specifics - Central note*
2188482	*SAP HANA on IBM Power Systems: Allowed hardware*
2218464	*Supported products when running SAP HANA on IBM Power Systems*
2230704	*SAP HANA on IBM Power Systems with multiple LPARs per physical host*
2235581	*SAP HANA: Supported operating systems*
2205917	*Recommended OS settings for SLES 12 / SLES for SAP Applications 12*
2684254	*Recommended OS settings for SLES 15 / SLES for SAP Applications 15*
2292690	*Recommended OS settings for RHEL 7*
2656575	*HANA 2 SPS4 release note*
2551355	*SAP HANA Platform V2.0 SPS 03 Release Note*
2613646	*SAP HANA TDI Phase 5*

The following sections describe important aspects of an SAP HANA on IBM Power Systems implementation that uses the guidelines that are described in the notes in Table 2-1. These rules must be followed in order for the system to be compliant and supported by SAP. It is also considered a best practice to discuss these guidelines with SAP before starting the implementation because they can have an impact on your systems architecture. There are no comments that are documented in the following sections regarding the day-to-day requirements, but we certainly apply all of them throughout the implementations in this publication, and mention them when doing so.

2.1.1 Storage and file system requirements

SAP HANA requires a minimum of four file systems:

- ► The data file system: Where all the data is stored.
- ► The log file system: Where the logs are stored.
- ► The shared file system: Where the binary file and file-based backups are stored.
- ► The `/usr/sap` file system: Where the local SAP system instance directories are stored.

As a best practice, for implementations that use storage area network (SAN) storage disks with Extents File System (XFS), the data area must be divided into a minimum of four LUNs[1], the log area can be divided into multiple of four LUNs as well, and the shared area can be on a single LUN or multiple ones. Their sizes vary according to the following SAP rules, which are documented in *SAP HANA Storage Requirements*:

- ► The *minimal* data area size requirement is 1.2 times the anticipated net data size on disk if an application-specific sizing program can be used (for example, SAP HANA Quick Sizer). If no sizing program can be used, then the minimum becomes 1x the amount of RAM memory. Although there is no maximum limit, three times the size of the memory is a good upper limit. Use multiples of four for the number of LUNs (4, 8, 12, and so on).

- ► The *minimal* log area size is 0.5 times the size of memory for systems with less than or equal to 512 GB of memory, or a fixed 512 GB for systems with more than 512 GB of memory. As a best practice from our implementation experiences, using a log area equal to the memory size for systems with less than 512 GB of memory is adequate to ensure optimal performance.

- ► The shared area size is 1x the size of the memory, up to the limit of 1 TB. For scale-out configurations, this requirement is per group of four worker nodes, not per node.

SAP Note 2055470 requires the use of one of three file systems types for *production* SAP HANA on IBM Power Systems environments for the *data* and *log* areas: XFS, Network File System (NFS) (with a 10 Gbps dedicated, redundant network), or IBM Spectrum Scale in an Elastic Storage Server configuration with a minimum 10 Gbps Ethernet or InfiniBand connection. No other file system type is supported.

In addition to the file system type, the storage unit providing the LUNs must be certified by SAP to work with HANA in a Tailored Datacenter Integration (TDI) methodology. A storage list can be found at Certified and Supported SAP HANA Hardware Directory.

For the log area, you *must* use either low-latency disks, such as flash or solid-state drives (SSDs), or ensure that the storage unit has a low-latency write cache area. This setup allows changes to the data content in memory to be quickly written to a persistent device. These two alternatives ensure that the speed of making the changes persistent on disk is as fast as possible. After all, what good does an in-memory database (DB) provides if commit operations must wait on slow disk I/O operations?

> **Note:** Finally, and most important, the storage areas for data and log must pass the SAP HANA Hardware Configuration Check Tool (HWCCT) file system tests.

---

[1] Based on the number of paths to the storage. Our implementations use four N_Port ID Virtualization (NPIV) paths.

Non-production SAP HANA on IBM Power Systems can follow relaxed guidelines for the storage and file systems, as described in SAP Note 2055470. Non-production systems can:

► Use IBM Spectrum Scale in any kind of configuration, such as Elastic Storage Server, and data and log disks.

► Use ext3 for data and log disks.

► Use standard network connectors (non-high-performance) for disk access when accessing disks over the network.

Additionally, non-production systems can be relaxed in the following ways:

► No need to pass the HWCCT file system benchmarks.

► Therefore, there is no need to place logs on low-latency disks or use a storage low-latency write-cache.

2.2 Preparing your software

This section provides guidelines about where to get the software that you need to perform an SAP HANA on IBM Power Systems installation, including the operating system (OS), IBM software for Linux on Power, and the HANA installer itself.

2.2.1 Getting your operating system image

You can obtain the SUSE Linux Enterprise Server image directly from the SUSE Downloads. You can download the no-charge trial ISO images to start, but you *must* have a valid SUSE Linux Enterprise Server license to register the system later, or your environment will not be supported after 60 days. Also, the versions that have the high availability (HA) packages that are commercially supported are the *for SAP Application* ones, and they are also the only versions that are supported for production environments. So, ensure that you get the image that you need and that you have a license to apply later.

Similarly for Red Hat, you can obtain the image directly from the Red Hat downloads page. You can download the no-charge trial ISO images to start, but you *must* have a valid Red Hat Enterprise Linux Server license to register the system later, or your environment will not be supported after 30 days.

Every customer who purchases SAP HANA on IBM Power Systems receives either a SUSE Linux Enterprise Server license or Red Hat Enterprise Linux license from either IBM or SUSE or Red Hat, depending on from whom the license was purchased. If the license is acquired from IBM, then IBM supports any issues with the OS and is the starting point for opening OS support tickets. If the license is acquired directly from SUSE or Red Hat, then SUSE or Red Hat supports any issues with the OS and is the starting point for opening OS support tickets.

Important notice: The OS license code comes in a white envelope with the IBM hardware if you purchased the license from IBM. Do not lose this envelope because if you do, you must engage your sales representatives to obtain another license, and this process is time-consuming and impacts your project schedule.

2.2.2 Getting the IBM service and productivity tools for Linux on Power

IBM Power Systems is known for its high levels of reliability, availability, and serviceability (RAS). The difference between an ordinary Linux for x86 image and a Linux on Power image is that the latter has a layer of extra added value software to enable Linux to take advantage of Power System hardware and virtualization features, such as dynamic logical partition (DLPAR) operations, resource monitoring and control (RMC) communication with the Hardware Management Console (HMC), and other functions.

Parts of the IBM RAS tools are distributed to Linux Business Partners such as SUSE, Red Hat, and Ubuntu, and some others are available for download from IBM at no charge. So, when you install Linux on Power, a subset of the RAS tools are already there. Nevertheless, download the other packages from the IBM website, and any updates to the packages that are included with the Linux distribution.

The RAS tools are based on the OS version that you use. To download and use them, see Service and productivity tools.

> **Notice:** Installing and updating the IBM Linux on Power RAS tools is a best practice for SAP HANA on IBM Power Systems environments and other Linux on Power environments. For packages that you cannot install in HANA logical partitions (LPARs), see SAP Note 2055470.

2.2.3 Getting the SAP HANA on IBM Power Systems installation files

The SAP HANA on IBM Power Systems installation files are downloadable from the SAP Support Portal. You must have an SAP user ID (SAP user) with enough credentials to download it. Only customers who purchased HANA licenses have access to the software.

What you must download from SAP's support portal are the *installation files*, not each individual SAP HANA component (server, client, studio, and so on). What you need to get is a set of compressed RAR files. The first of them has a `.exe` extension, but these RAR files work on Linux on Power as well.

Click **Download software** on the SAP Support Portal website. Then, click **By Alphabetical Index (A-Z)** →**H** →**SAP In-Memory (SAP HANA)** →**HANA Platform Edition** →**SAP HANA Platform Edition** →**SAP HANA Platform Edition 2.0** →**Installation** to get the HANA software.

Download all files for *Linux on Power*, including the HANA platform edition files and the HANA Cockpit. The Cockpit is available for HANA 2.0 only.

3

IBM PowerVM and SAP HANA

This chapter describes the considerations for configuring IBM PowerVM when providing logical partitions (LPARs) to run SAP HANA.

This chapter covers the following topics:

► Introduction to IBM PowerVM and SAP HANA
► Virtual I/O Server
► Other considerations

3.1 Introduction to IBM PowerVM and SAP HANA

IBM Power Systems servers provide flexibility to meet the individual needs of organizations that deploy SAP HANA. One aspect of this flexibility is that robust virtualization is supported and ready for use. It helps consolidate multiple SAP HANA virtual machines (VMs) on a single Power Systems server. IBM PowerVM virtualization is fully supported by SAP, enabling customers to deploy SAP HANA in a virtual environment that supports both dedicated and shared processor resources, running both production and non-production workloads in a single server.

Virtualization with PowerVM also enables you to handle the varying utilization patterns that are typical in SAP HANA workloads. Dynamic capacity sizing allows for fast, granular reallocation of compute resources among SAP HANA VMs. This approach to load-balancing and tailoring the workload enhances agility compared to competing processor architectures that require capacity to be allocated in larger chunks.

Another contributor to the flexibility of Power Systems servers is that they are deployed as part of the SAP Tailored Datacenter Integration (TDI) model. The goal of this approach is to reuse existing IT resources, such as server, storage, and networking assets. By supporting TDI in the deployment of SAP HANA, Power Systems servers give organizations a choice of the technology that they use compared to the rigidly defined hardware appliances that are used in many competing SAP HANA infrastructures.

For more information about the SAP HANA TDI, see SAP HANA Server and Workload Sizing.

For more information about PowerVM, see IBM PowerVM: Overview.

For more information about PowerVM and SAP HANA, see SAP HANA server infrastructure with Power Systems.

For technical details about the PowerVM configuration for systems that run SAP HANA, see SAP HANA on IBM Power Systems and IBM System Storage - Guides.

> **Note:** Any information in this guide is superseded by the information at the links in this chapter. Check those links for any updated information about SAP HANA on IBM Power Systems.
>
> These links provide and build a basic set of documents, but might not be complete for all cases.

3.2 Virtual I/O Server

A medium level of knowledge about PowerVM and Virtual I/O Server (VIOS) is assumed. If that is not the case, you must become familiar with the topic. A starting point is setting up a dual VIOS by using the information from IBM Knowledge Center.

> **Note:** The following statements are based on multiple technical items, including NUMA allocations, IBM POWER Hypervisor™ dispatcher wheel, multipath, network communications optimization, and others.
>
> It is not the goal of this chapter to explain in detail the reasons behind these statements. If the reader wants to understand the reasoning behind them, see the linked documentation in this chapter.

The specifics of the VIOS configuration when using LPAR with a production SAP HANA are:

► If I/O virtualization is used, a dual-VIOS setup is mandatory. You can have more than two VIOSes in the system separating different environments, such as production and test or multiple customers. At the time of writing, Novalink and KMV is not supported for virtualization.

► Each VIOS must be configured with at least two dedicated or dedicated donating cores for any serving SAP production systems. Size them as needed and monitor CPU usage to adapt to workload changes over the lifetime of the system.

► At least one Fibre Channel card per VIOS is needed. For high-end systems, be sure to use optimal PCI placement. The HBA port speed must be at least 8 Gb minimum. A best practice is 16 Gb, including the infrastructure.

► At least one Ethernet card per VIOS is needed. Interfaces with 10 GbE are needed at a minimum for scale-up systems. For scale-out systems, a speed of at least 10 GbE is mandatory. For more information, see *SAP HANA Network Requirements*.

> **Note:** There are strict PCI placement rules for optimal performance that are not explicitly HANA-related. These rules are server- and card-dependant, and follow the required PCI slot placement. If you require assistance for your particular system, contact IBM Support.

► Either dedicate a PCI card for the LPAR or use Ethernet virtualization with a Shared Ethernet Adapter (SEA). Although single root input/output virtualization (SR-IOV) vNIC is not yet explicitly used, use SR-IOV-capable cards, in particular if other LPARs are going to be hosted in a system that already can use SR-IOV vNIC technology.

► Use only supported storage virtualization with N_Port ID Virtualization (NPIV) when using an SAP storage connector. Otherwise, use NPIV over other storage virtualizations on PowerVM. Four NPIV ports per HANA LPAR must be used. Alternatively, as on Ethernet you can also dedicate a PCI card to the LPAR and not use any virtualization; the 4-port requirement remains regardless of your approach.

► Jumbo frames with an MTU size of 9000 are required for native and VIOS-attached 10 Gb Ethernet adapters to achieve the throughput key performance indicators (KPIs) that are demanded by the SAP HANA Hardware Configuration Check Tool (HWCCT) tool. For scale-up systems, there are no network KPIs.

► Use Platform Large Send (PLSO).

For more information about setting up PLSO, MTU, and other SEA tuning, see Configuring traditional largesend for SAP HANA on SLES with VIOS.

A simple overview of the configuration of a stand-alone system with HANA LPAR is shown in Figure 3-1.
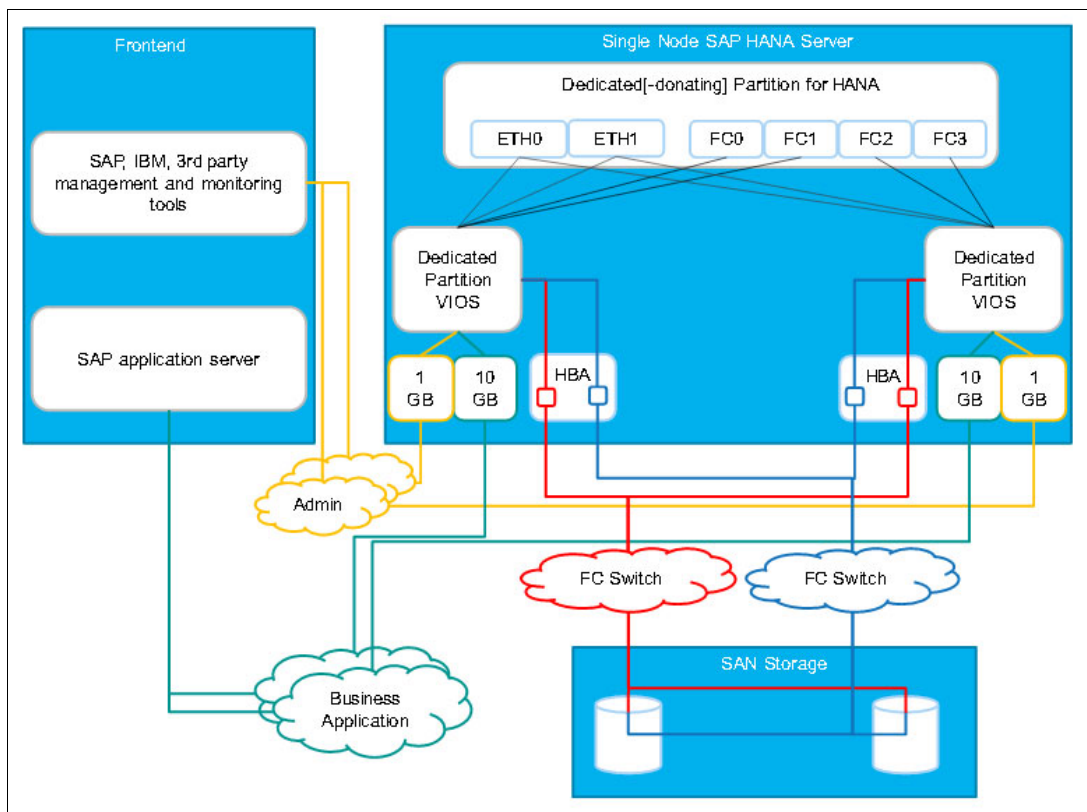
Figure 3-1   PowerVM overview

Virtual ISO image installation

Although there are other ways to do a Base Operative System (BOS) installation than using a virtual ISO image, on virtual environments that lack other types of automation, such as IBM PowerVC, use the virtual ISO. For more information about how to use the virtual ISO installation media, see *How to Assign a VIOS Hosted Virtual Optical Device Using the New HMC V8 GUI*.

3.3 Other considerations

This section shares other topics that are requirements to run SAP HANA on IBM Power Systems, but that ease your overall experience with the solution.

3.3.1 IBM PowerVC

IBM PowerVC can accelerate the deployment of LPARs in general. It performs Hardware Management Console (HMC) configurations, VIOS, storage area network (SAN) zoning, storage mapping, and BOS deployments. For more information, see IBM PowerVC: Overview IBM PowerVC: Overview.

3.3.2 IBM Systems Lab Services

The IBM Power to Cloud Rewards Program can help you implement a cloud solution on Power Systems servers. Either with or without this reward program, you can contact IBM Systems Lab Services about using SAP HANA on IBM Power Systems. Installations, health checks, and performance reviews are the standard offerings from IBM Systems Lab Services. Other tailored engagements can be designed. To contact the IBM Systems Lab Services team, see Lab Services for Power Systems Lab Services for Power Systems.

4

Operating system installation and customization

This chapter describes the installation of SUSE Linux Enterprise Server V12 SP3 for SAP Applications and Red Hat Enterprise Linux Server V7.4 on IBM Power Systems logical partitions (LPARs) to host SAP HANA databases (DBs).

This chapter covers the following topics:

► Introduction
► Creating the logical partition for SAP HANA
► Installation to the logical partition

4.1 Introduction

The information in this chapter is valid at the time of writing. Before planning an operating system (OS) installation for an SAP HANA instance, see SAP Note 2055470. Then, look at SAP Note 2235581, which then leads to two SAP Notes that are OS-specific: SAP Note 2009879 and SAP Note 2205917.

In addition to these SAP Notes, all the documentation that is specified in Chapter 2, "Planning your installation" on page 13, and at SAP HANA on IBM Power Systems and IBM System Storage - Guides also are useful.

Start the LPAR and install the Base Operative System (BOS) by using the serial console on the Hardware Management Console (HMC) until the SUSE or Red Hat Enterprise Linux installer is available over the network with Virtual Network Computing (VNC). From this point, follow the GUI installation procedure.

> **Note:** There are other ways to install the BOS, such as using the command-line interface (CLI). However, for this exercise, we use the GUI whenever possible.

4.2 Creating the logical partition for SAP HANA

There are multiple ways to create an LPAR. All methods are documented in IBM Knowledge Center.

There are specific recommendations for LPARs to be used for the HANA DB that you must be aware of and follow. These recommendations are specified in Chapter 2, "Planning your installation" on page 13, and in subsequent SAP Notes and SAP HANA on IBM Power Systems and IBM System Storage - Guides.

> **Note:** An LPAR is not the only way to install a HANA DB. It also can be installed in a full system partition configuration. The only prerequisite is that it is installed on top of PowerVM and its size is not over the limits. The process for installing the BOS is similar either way.

4.3 Installation to the logical partition

After the LPAR is created, a BOS installation must be done, as specified in Chapter 2, "Planning your installation" on page 13 and in SAP Note 2235581. SUSE Linux Enterprise Server 12 SP3 for SAP Applications ad Red Hat Enterprise Linux Server 7.4 are supported and are the versions that are installed in this chapter.

This chapter also uses Virtual I/O Servers (VIOS) for I/O without a dedicated PCI slot to the LPAR. We use N_Port ID Virtualization (NPIV) for the storage virtualization and Shared Ethernet Adapters (SEAs) for the network virtualization, as shown in Chapter 3, "IBM PowerVM and SAP HANA" on page 19.

> **Important:** NPIV is a must for Fibre Channel because VDisk mapping introduces too much latency for high-speed storage and increases CPU requirements on the VIOS.

4.3.1 Starting the logical partition in SMS mode

Complete the following steps:

1. From the HMC GUI, select the HANA LPAR, which in this example is hana001. Click **Actions** →**Activate**, as shown in Figure 4-1.

> **Note:** Figure 4-1 shows the Hardware Management Console (HMC) running V9R1 M920. Your view can differ depending on which version of the HMC code you are running.
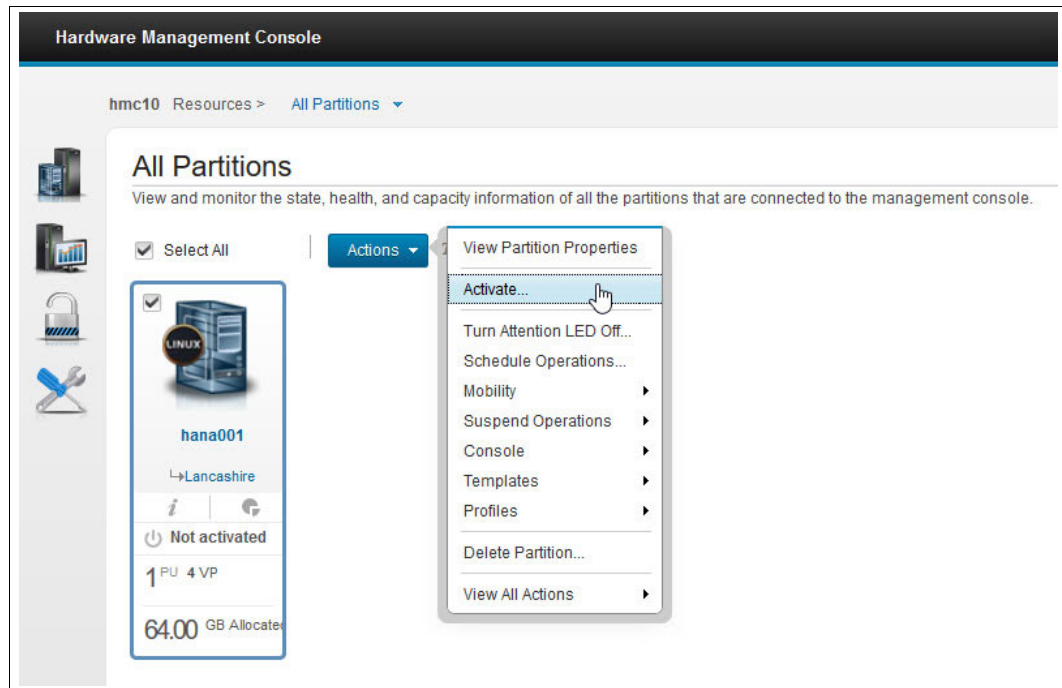
Figure 4-1   Activating the logical partition from the Hardware Management Console

This action opens the System Activation Options window, as shown in Figure 4-2.
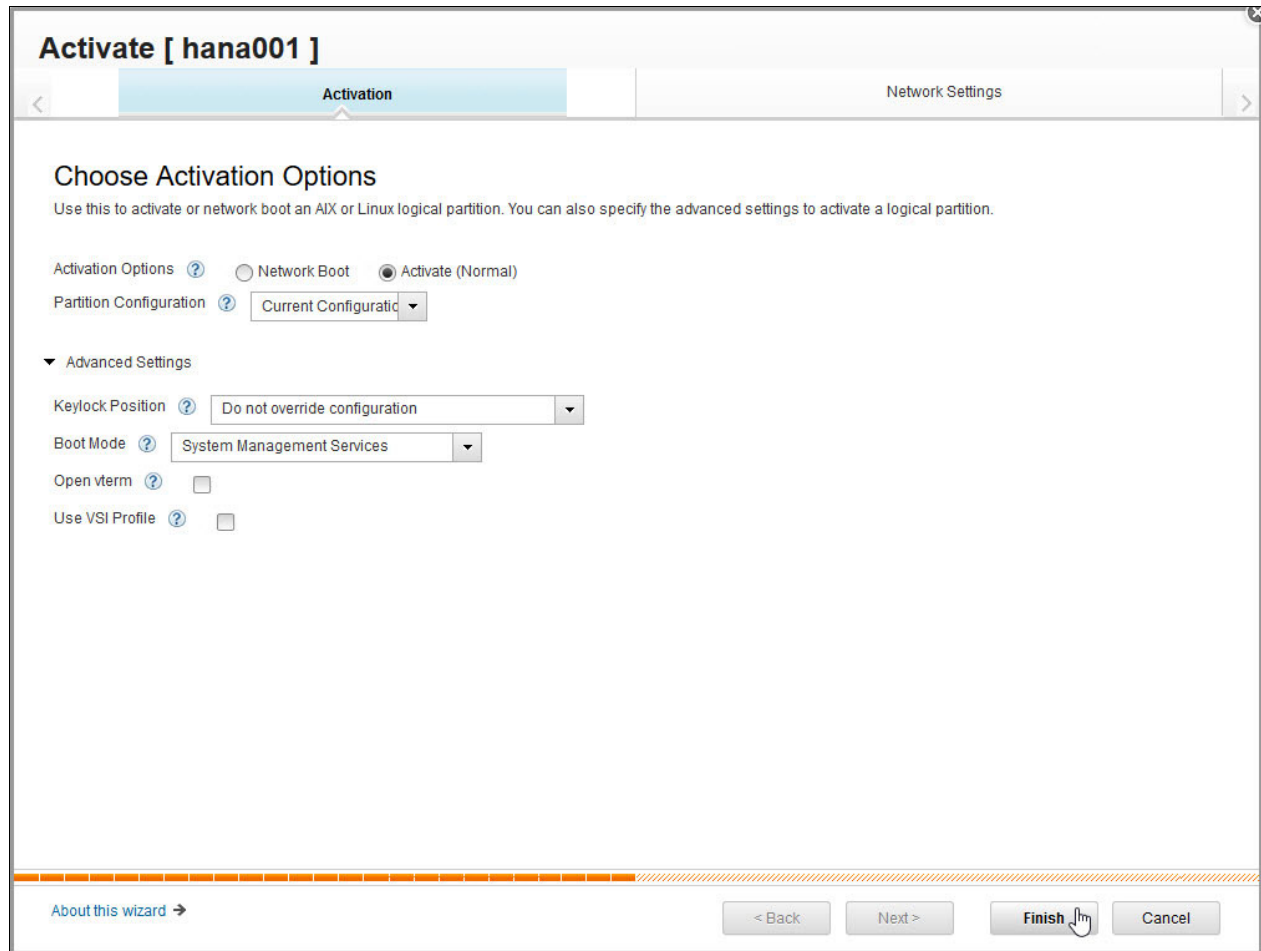
Figure 4-2   Choose Activation Options for your logical partition

2. After you select the appropriate profile, click **Advanced Settings** and select **Systems Management Services** under **Boot Mode**. Click **Finish** to start the LPAR. When the LPAR starts, you see the window that is shown in Figure 4-3 on page 29.
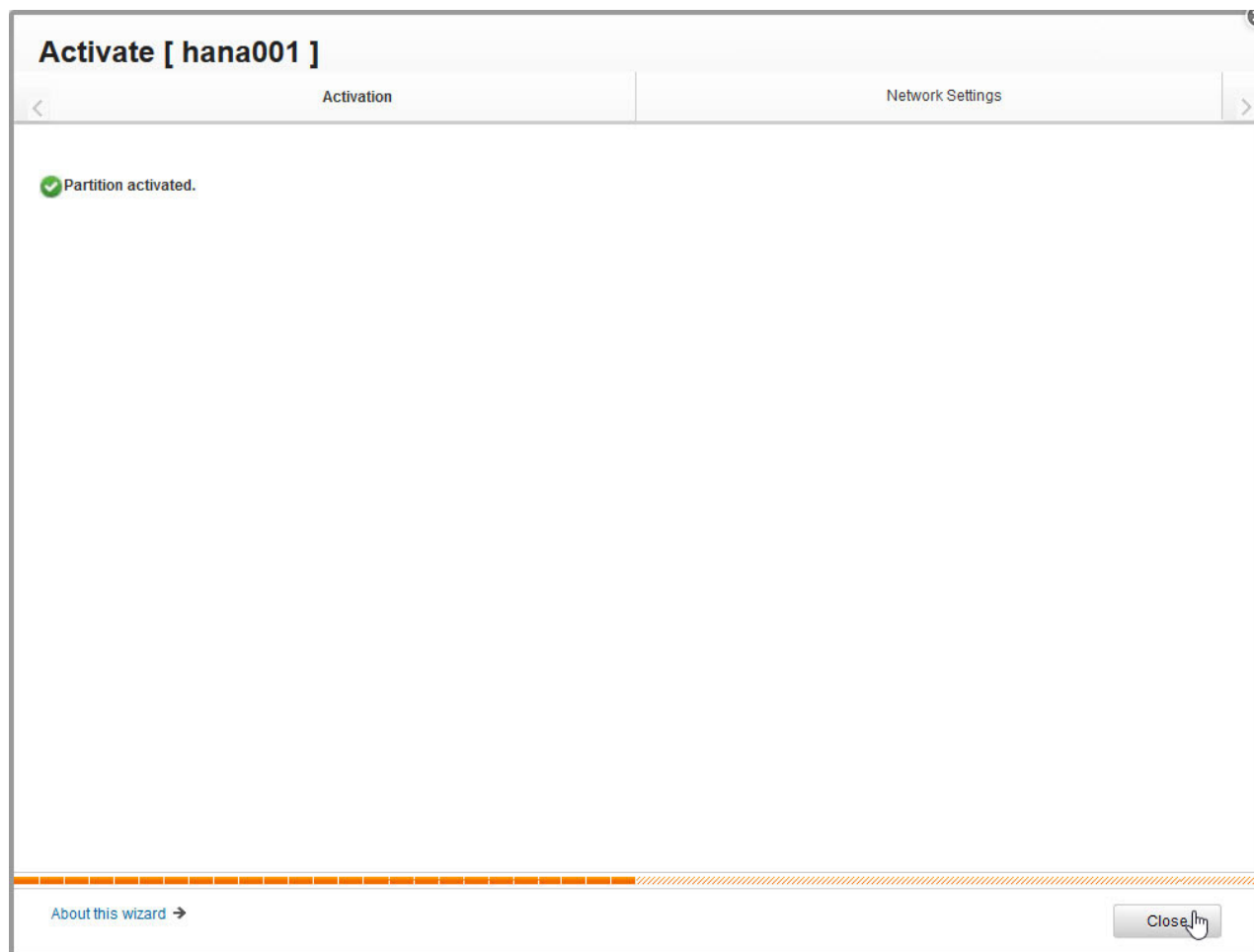
Figure 4-3 Confirming the partition activation

3. Close the wizard by clicking **Close**.

> **Note:** Be sure to boot into SMS so that you can choose the installation device.
>
> For more information, see IBM Knowledge Center, where this method, among others, is explained in detail.

4.3.2 Installing the Base Operative System from the Hardware Management Console virtual terminal

This section uses the virtual terminal (vterm) by way of an SSH connection to the HMC. Complete the following steps:

1. Using SSH, connect to the HMC that manages the frame that hosts the SAP HANA LPAR that you want to install, run the **vtmenu** command, and select the frame and the partition that is being installed. You see the initial SMS menu entry, as shown in Example 4-1.

Example 4-1   Initial SMS menu

```
PowerPC Firmware
 Version FW860.51 (SV860_165)
 SMS (c) Copyright IBM Corp. 2000,2016 All rights reserved.
-------------------------------------------------------------------------------
 Main Menu
 1.    Select Language
 2.    Setup Remote IPL (Initial Program Load)
 3.    I/O Device Information
 4.    Select Console
 5.    Select Boot Options




-------------------------------------------------------------------------------
 Navigation Keys:

                                             X = eXit System Management
Services

-------------------------------------------------------------------------------
 Type menu item number and press Enter or select Navigation key:
```

> **Note:** For more information about selecting a device to boot, see Example Using SMS To Choose Boot Device.

2. Select **Select Boot Options** and press Enter, as shown in Example 4-2.

Example 4-2   SMS multiboot menu

```
PowerPC Firmware
 Version FW860.51 (SV860_165)
 SMS (c) Copyright IBM Corp. 2000,2016 All rights reserved.
-------------------------------------------------------------------------------
 Multiboot
 1.    Select Install/Boot Device
 2.    Configure Boot Device Order
 3.    Multiboot Startup <OFF>
```

```
   4.  SAN Zoning Support




-------------------------------------------------------------------------------
 Navigation keys:
 M = return to Main Menu
 ESC key = return to previous screen        X = eXit System Management
Services


-------------------------------------------------------------------------------
 Type menu item number and press Enter or select Navigation key:
```

3. Select **Select Install/Boot Device** and press Enter. The panel that is shown in Example 4-3 opens.

Example 4-3   Select Device Type SMS menu

```
PowerPC Firmware
 Version FW860.51 (SV860_165)
 SMS (c) Copyright IBM Corp. 2000,2016 All rights reserved.
-------------------------------------------------------------------------------
 Select Device Type
 1.   Tape
 2.   CD/DVD
 3.   Hard Drive
 4.   Network
 5.   List all Devices








-------------------------------------------------------------------------------
 Navigation keys:
 M = return to Main Menu
 ESC key = return to previous screen        X = eXit System Management
Services


-------------------------------------------------------------------------------
 Type menu item number and press Enter or select Navigation key:
```

4. Select **CD/DVD** and press Enter. The panel that is shown in Example 4-4 opens.

Example 4-4   Select Media Type SMS menu

```
PowerPC Firmware
 Version FW860.51 (SV860_165)
 SMS (c) Copyright IBM Corp. 2000,2016 All rights reserved.
 -------------------------------------------------------------------------------
 Select Media Type
 1.    SCSI
 2.    SAN
 3.    SAS
 4.    SATA
 5.    USB
 6.    List All Devices




 -------------------------------------------------------------------------------
 Navigation keys:
 M = return to Main Menu
 ESC key = return to previous screen        X = eXit System Management
 Services


 -------------------------------------------------------------------------------
 Type menu item number and press Enter or select Navigation key:
```

5. Because you are using vSCSI, select **1. SCSI** and press Enter. The panel that is shown in Example 4-5 opens.

Example 4-5   Select Media Adapter SMS menu

```
PowerPC Firmware
 Version FW860.51 (SV860_165)
 SMS (c) Copyright IBM Corp. 2000,2016 All rights reserved.
 -------------------------------------------------------------------------------
 Select Media Adapter
 1.          U8286.42A.21576CV-V19-C5-T1    /vdevice/v-scsi@30000005
 2.    List all devices
```

```
--------------------------------------------------------------------------------
 Navigation keys:
 M = return to Main Menu
 ESC key = return to previous screen         X = eXit System Management
Services

--------------------------------------------------------------------------------
 Type menu item number and press Enter or select Navigation key:
```

6. If the vSCSI adapter is properly configured, you see it in the SMS menu, as shown in Example 4-5 on page 32. Select the correct vSCSI adapter and press Enter. The panel that is shown in Example 4-6 opens.

Example 4-6   Select Device SMS menu

```
PowerPC Firmware
 Version FW860.51 (SV860_165)
 SMS (c) Copyright IBM Corp. 2000,2016 All rights reserved.
--------------------------------------------------------------------------------
 Select Device
 Device  Current  Device
 Number  Position  Name
 1.         -       SCSI CD-ROM
        ( loc=U8286.42A.21576CV-V19-C5-T1-L8100000000000000 )




--------------------------------------------------------------------------------
 Navigation keys:
 M = return to Main Menu
 ESC key = return to previous screen         X = eXit System Management
Services

--------------------------------------------------------------------------------
 Type menu item number and press Enter or select Navigation key:
```

7. Select the vCD device number and press Enter. This panel that is shown in Example 4-7 opens.

> **Note:** If the CD/DVD is already set up as Current Position of boot number 1, choosing the boot device is optional.

Example 4-7   Select Task SMS menu

```
PowerPC Firmware
 Version FW860.51 (SV860_165)
 SMS (c) Copyright IBM Corp. 2000,2016 All rights reserved.
--------------------------------------------------------------------------------
 Select Task

SCSI CD-ROM
     ( loc=U8286.42A.21576CV-V19-C5-T1-L8100000000000000 )

 1.    Information
 2.    Normal Mode Boot
 3.    Service Mode Boot




--------------------------------------------------------------------------------
 Navigation keys:
 M = return to Main Menu
 ESC key = return to previous screen          X = eXit System Management
Services

--------------------------------------------------------------------------------
 Type menu item number and press Enter or select Navigation key:
```

8. Select **2. Normal Mode Boot** and press Enter. The panel that is shown in Example 4-8 opens.

Example 4-8   Confirming the exit of the SMS mode menu

```
PowerPC Firmware
 Version FW860.51 (SV860_165)
 SMS (c) Copyright IBM Corp. 2000,2016 All rights reserved.
--------------------------------------------------------------------------------
 Are you sure you want to exit System Management Services?
 1.    Yes
 2.    No
```

```
   -------------------------------------------------------------------------------
    Navigation Keys:


                                                     X = eXit System Management
   Services


   -------------------------------------------------------------------------------
    Type menu item number and press Enter or select Navigation key:
```

9. Select **1. Yes** and press Enter. You exit from SMS mode and boot from the SUSE or Red
   Hat Enterprise Linux installation media. In this example, you boot from the SUSE
   installation media.

4.3.3 SUSE Linux Enterprise Server V12 SP3 for SAP applications installation

After you boot from the SUSE installation media, complete the following steps:

1. After a few seconds, you see the GRUB boot main menu, as shown in Example 4-9.

Example 4-9   SUSE Linux Enterprise Server 12 SP3 GRUB boot main menu

```
SUSE Linux Enterprise 12 SP3


 +------------------------------------------------------------------------------+
 |*Installation                                                                 |
 | Rescue System                                                                |
 | Upgrade                                                                      |
 | Check Installation Media                                                     |
 | local                                                                        |
 | Other options...                                                             |
 |                                                                              |
 |                                                                              |
 |                                                                              |
 |                                                                              |
 |                                                                              |
 |                                                                              |
 +------------------------------------------------------------------------------+


      Use the ^ and v keys to select which entry is highlighted.
      Press enter to boot the selected OS, `e' to edit the commands
      before booting or `c' for a command line.
```

2. Move the arrow up until the **Installation** menu entry is highlighted. Then, press the e key to edit the commands. The window that is shown in Example 4-10 opens.

> **Note:** If no key is pressed before the countdown ends on the GRUB main boot menu, a default installation is performed. Restart the LPAR and attempt to interrupt this panel before the countdown expires. To do this, start from the boot partition step that is described in 4.3.1, "Starting the logical partition in SMS mode" on page 27.

Example 4-10   GRUB Edit Installation submenu

```
SUSE Linux Enterprise 12 SP3


 +--------------------------------------------------------------------------+
 |setparams 'Installation'                                                  |
 |                                                                          |
 |   echo 'Loading kernel ...'                                              |
 |   linux /boot/ppc64le/linux                                              |
 |   echo 'Loading initial ramdisk ...'                                     |
 |   initrd /boot/ppc64le/initrd                                            |
 |                                                                          |
 |                                                                          |
 |                                                                          |
 |                                                                          |
 |                                                                          |
 |                                                                          |
 +--------------------------------------------------------------------------+

        Minimum Emacs-like screen editing is supported. TAB lists
        completions. Press Ctrl-x or F10 to boot, Ctrl-c or F2 for
        a command-line or ESC to discard edits and return to the GRUB menu.
```

For this installation, we know that the network device is `eth0`, which is going to be used for the VNC installation. We also know the IP address that is going to be used on that interface, which in this case is `10.10.12.83/24`. We also know that the IP gateway is `10.10.12.1` and the DNS servers are `10.10.12.10` and `10.10.12.9`. The host name is going to be `hana001` and the proxy IP address and port has no user authentication required. Append to the `linux` line the text that is shown in Example 4-11.

The reason that we are setting a proxy is that in this example is so that we can use the option to access the SUSE registration and updates. If you have a system that has direct access to the internet or uses the Subscription Management Tool, you can ignore the proxy configuration.

> **Note:** If you are not sure of the network interface name, you can try `eth*` instead of `eth0`, which sets the information in all eth devices.

Example 4-11   Options to append to the Linux entry in GRUB

```
ifcfg=eth0=10.10.12.83/24,10.10.12.1,10.10.12.10 hostname=hana001 vnc=1
vncpassword=Passw0rd proxy=http://10.10.16.10:3128
```

> **Note:** You can adapt this line for your environment by using the following syntax:
>
> ```
> ifcfg=eth0=<IP Address>/<Netmask range>,<Gateway>,<nameserver>
> hostname=<host name> vnc=1 vncpassword='VNCPASSWORD'
> proxy=http://USER:PASSWORD@proxy.example.com:PORT
> ```

After appending your information to the Linux entry in GRUB, the panel that is shown in Example 4-12 opens.

Example 4-12   Appended configuration to the GRUB linux line

```
SUSE Linux Enterprise 12 SP3


 +-----------------------------------------------------------------------------+
 |setparams 'Installation'                                                     |
 |                                                                             |
 |   echo 'Loading kernel ...'                                                 |
 |   linux /boot/ppc64le/linux ifcfg=eth0=10.10.12.83/24,10.10.12.1,10.10.12.1\|
 |0 hostname=hana001 proxy=http://10.10.16.10:3128 vnc=1 vncpas\|
 |sword=Passw0rd                                                               |
 |   echo 'Loading initial ramdisk ...'                                        |
 |   initrd /boot/ppc64le/initrd                                               |
 |                                                                             |
 |                                                                             |
 |                                                                             |
 |                                                                             |
 +-----------------------------------------------------------------------------+


        Minimum Emacs-like screen editing is supported. TAB lists
        completions. Press Ctrl-x or F10 to boot, Ctrl-c or F2 for
        a command-line or ESC to discard edits and return to the GRUB menu.
```

3. Press Ctrl+x, which starts the SUSE installation with the chosen parameters on GRUB. After a couple of minutes, the panel that is shown in Example 4-13 opens.

Example 4-13   Starting YaST2 and Virtual Network Computing boot message

```
starting VNC server...
A log file will be written to: /var/log/YaST2/vncserver.log ...

***
***             You can connect to <host>, display :1 now with vncviewer
***             Or use a Java capable browser on http://<host>:5801/
***

(When YaST2 is finished, close your VNC viewer and return to this window.)

Active interfaces:

eth0: 2e:82:88:20:14:1e
  10.10.12.83/24
  fe80::2c82:88ff:fe20:141e/64

*** Starting YaST2 ***
```

Installing SUSE by using YaST2 and Virtual Network Computing

Complete the following steps:

1. Use the VNC client or the web browser to connect to the GUI that is hosting the YaST2 installer. We use a VNC client. To connect to the YaST2 installer, input the IP address of the server (10.10.12.83) with VNC port 5901. When prompted for a password, use the password that is passed as the parameter, as shown in Example 4-11 on page 36.

   After you connect, the Language, Keyboard, and License agreement window opens, as shown in Figure 4-4.
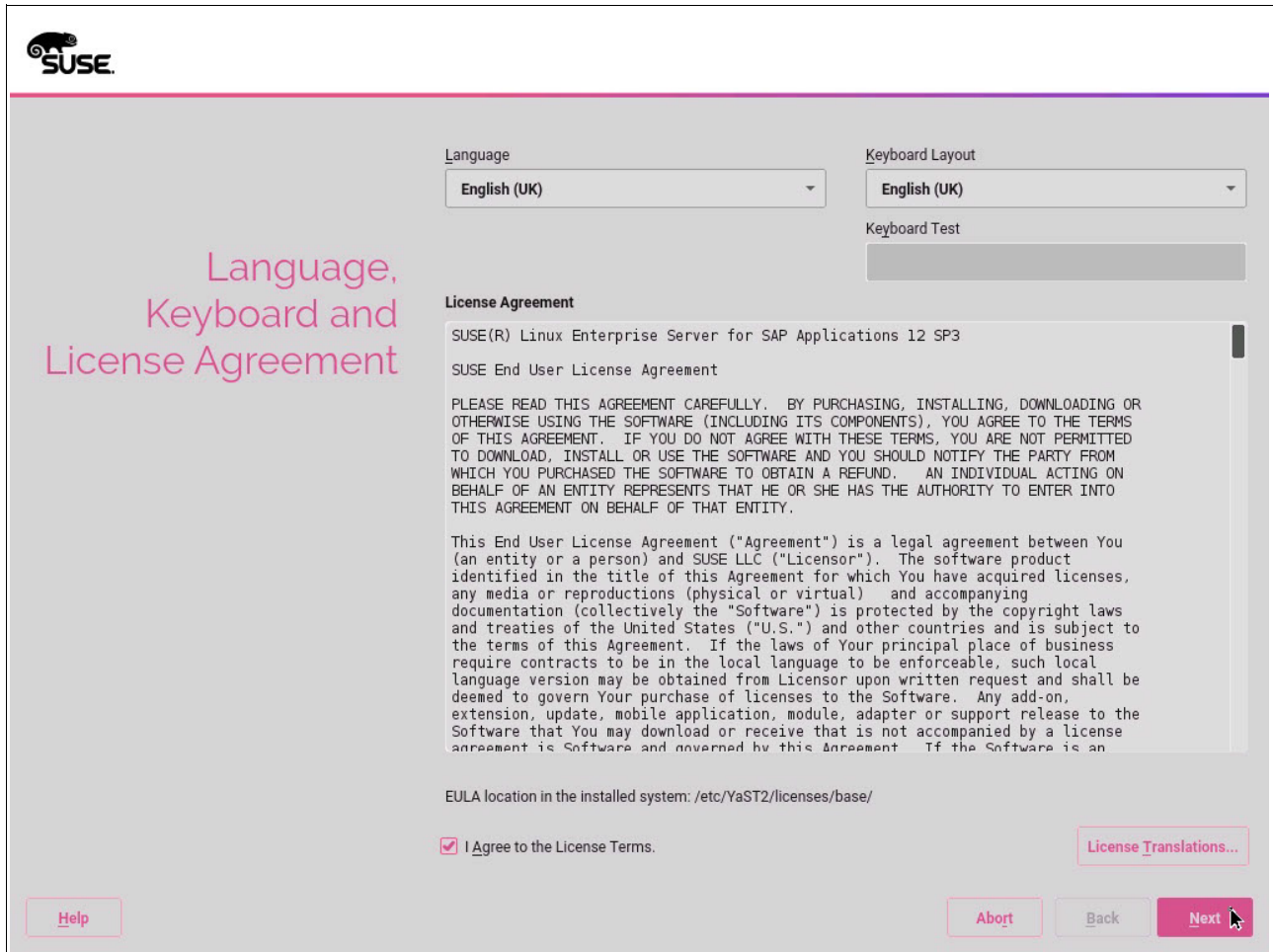

Figure 4-4   YaST2 Language, Keyboard, and License Agreement window

2. Select your language and keyboard settings, and if you agree with the SUSE license terms, select the **I agree to the License Terms** check box, and click **Next**.

The System Probing Yast2 window opens, where you can enable multipath in this setup, as shown in Figure 4-5.
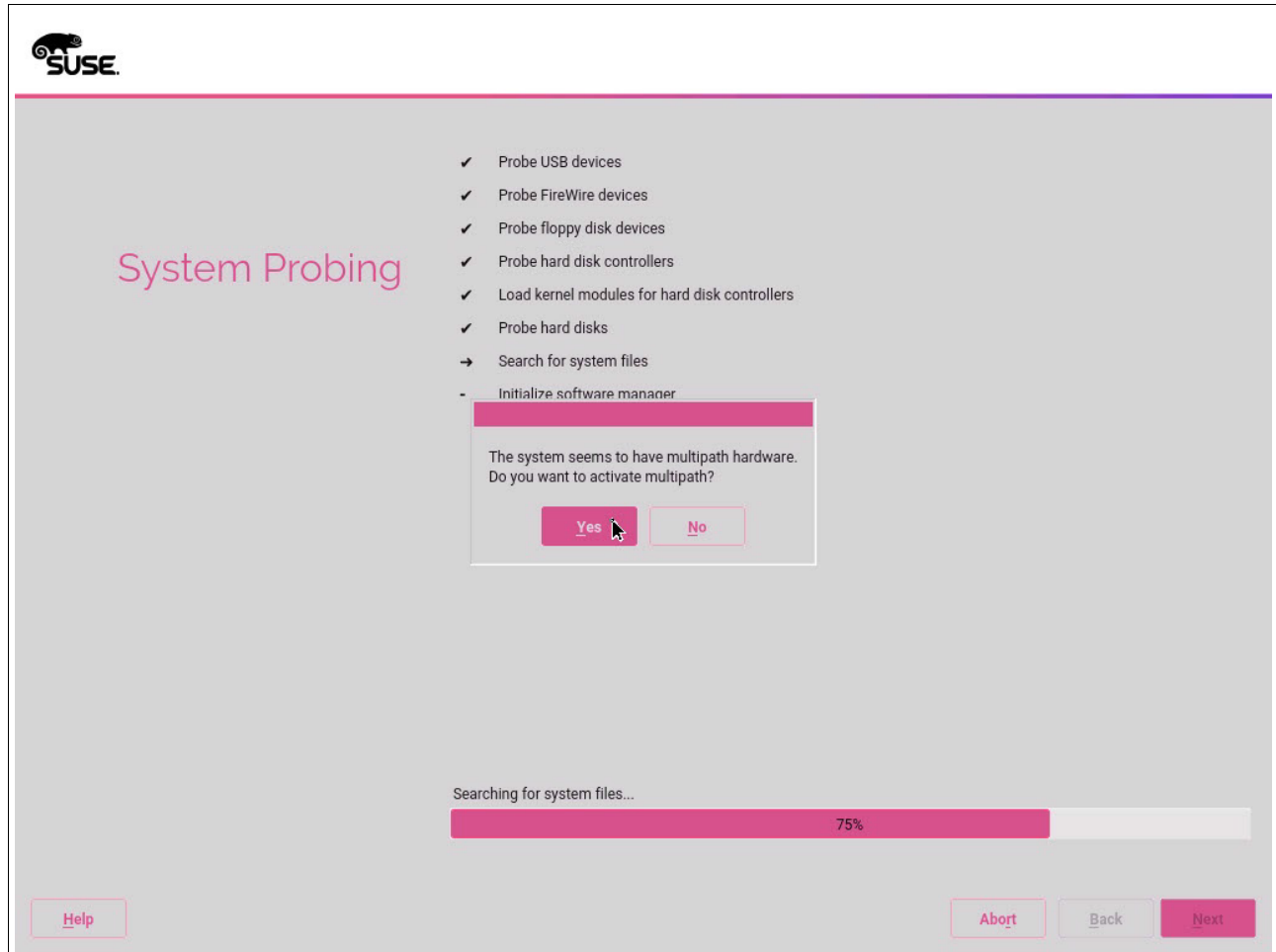

Figure 4-5   Yast2 System Probing multipath window

3. Click **Yes** to activate multipath in this installation.

> **Note:** If you do not see the multipath window, there is a problem with the storage configuration. Before continuing with the installation, see 3.2, "Virtual I/O Server" on page 20.

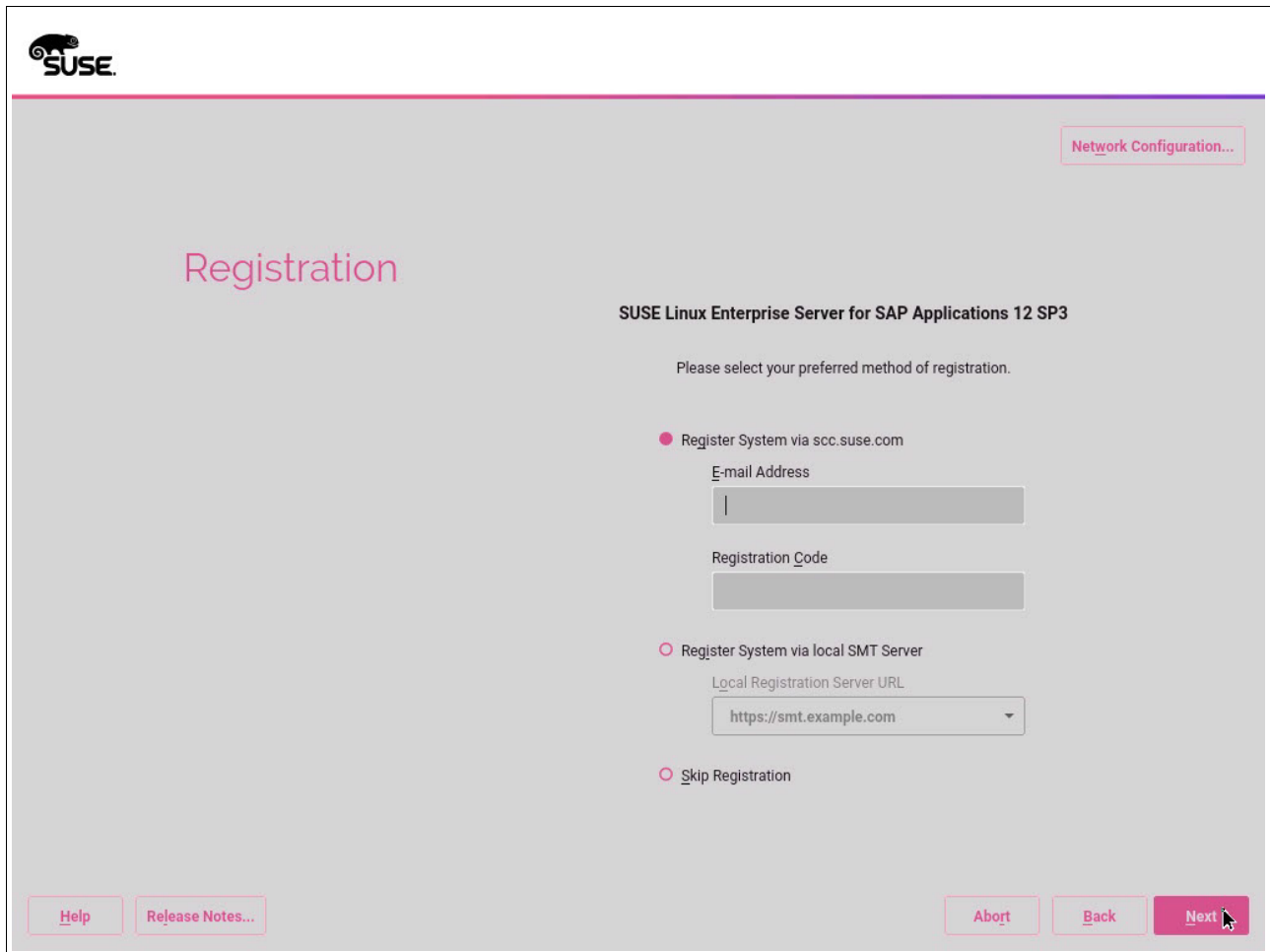The Registration window opens, as shown in Figure 4-6.

Figure 4-6   YaST2 Registration window

4. In this scenario, we use the `scc.suse.com` system to register this installation. If your setup has a local SMT server, you can use it instead.

> **Note:** If you do not register now, you must do it later before the HANA DB software is installed.

After you input the registration information, click **Next**. When the registration completes, a window opens that offers to enable the update repositories, as shown in Figure 4-7.
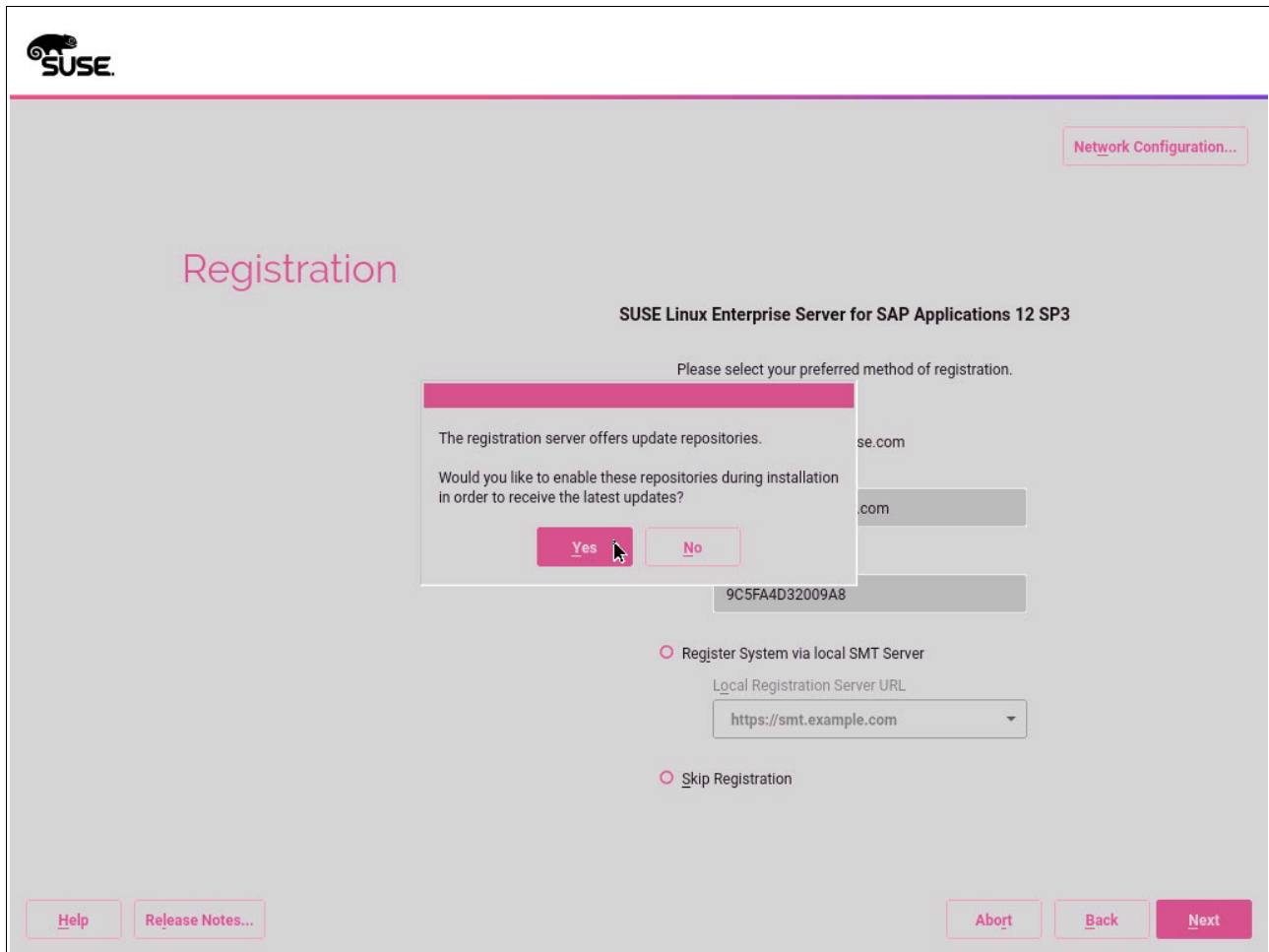

Figure 4-7   YaST2 update repositories window

5. Click **Yes** to enable the update repositories. The Extension and Module Selection window opens, as shown in Figure 4-8.
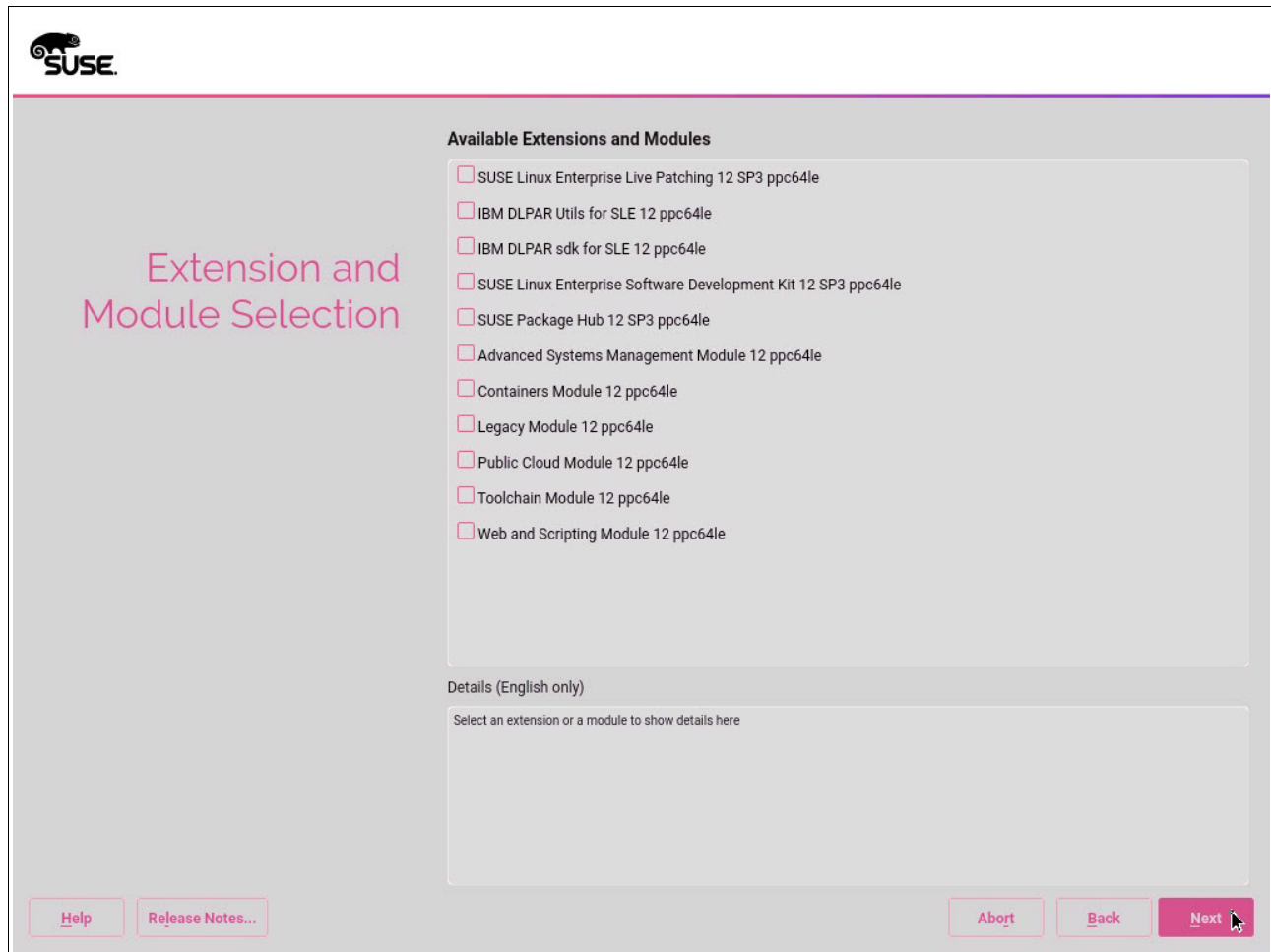

Figure 4-8   YaST2 Extension and Module Selection window

6. Select the **IBM DLPAR Utils for SLE 12 ppc64le** and **IBM DLPAR sdk for SLE 12 ppc64le** extensions, and click **Next**. The Extension and Module Selection window opens, as shown in Figure 4-9.
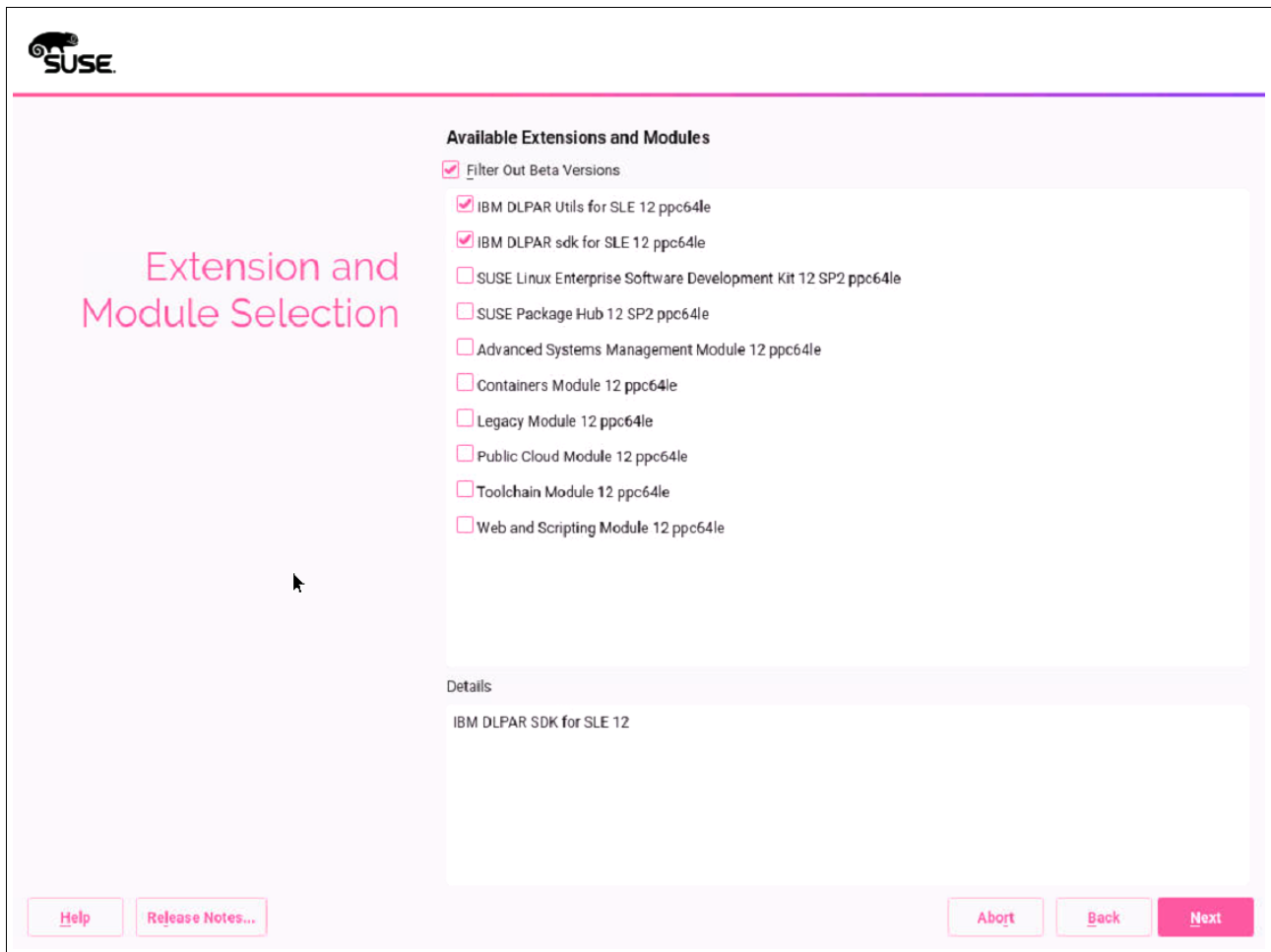

Figure 4-9   Yast2 IBM DLPAR Utils for SLE 12 ppc64le License Agreement window

7. If you agree with the license terms, select **I Agree to the License Term** and click **Next**. The IBM DLPAR sdk for SLE 12 ppc64le License Agreement window opens, as shown in Figure 4-10.
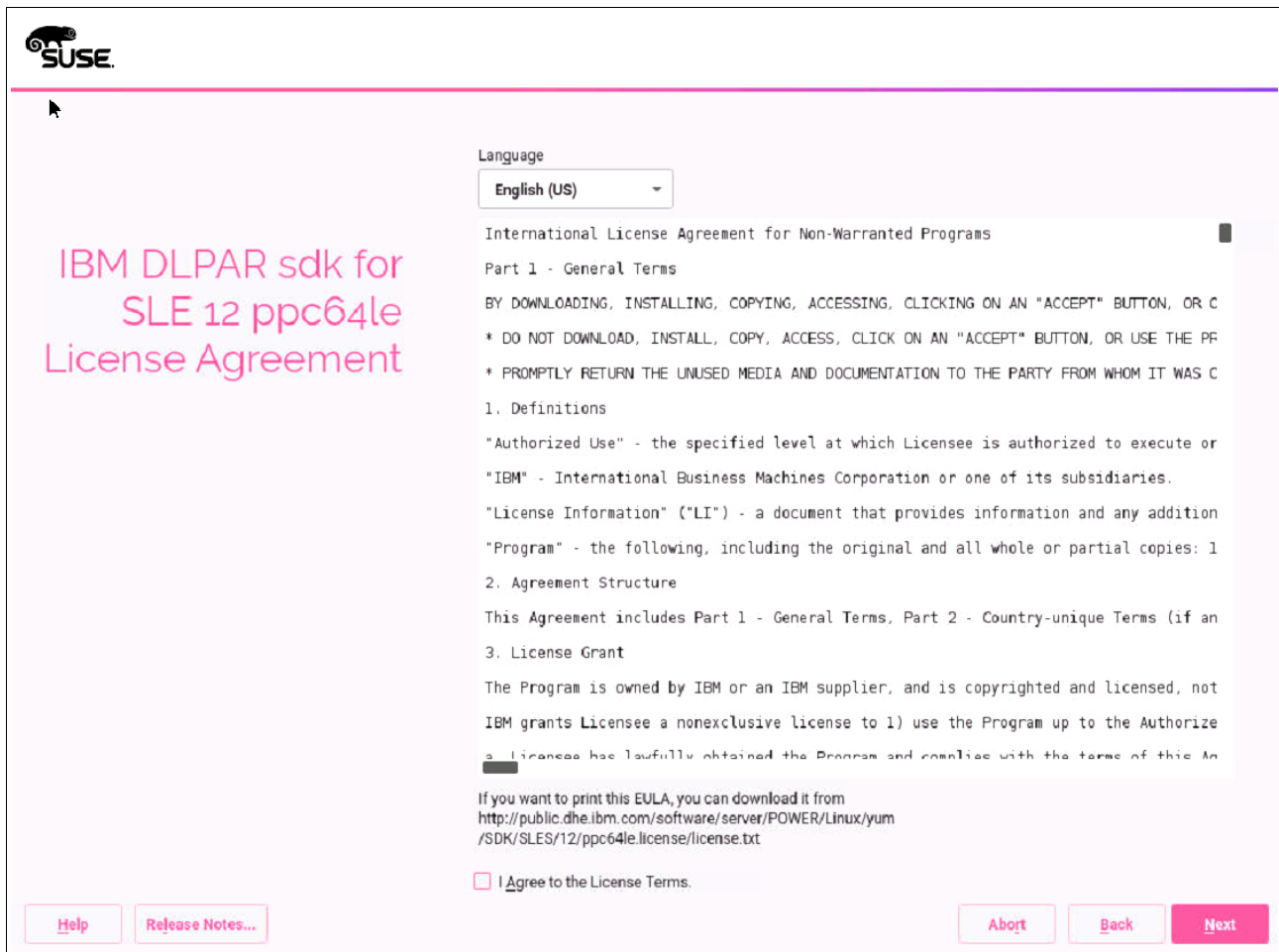

Figure 4-10   Yast2 IBM DLPAR sdk for SLE 12 ppc64 License Agreement window

8. If you agree with the license terms, select **I Agree to the License Terms** and click **Next**. The Import GnuPG Key for IBM-DLPAR-utils repository window opens, as shown in Figure 4-11.

Figure 4-11   Yast2 Import GnuPG Key for IBM-DLPAR-utils repository window

9. After you check that the ID and Fingerprint are correct, click **Trust**. The Import GnuPG Key for IBM-DLPAR-utils repository window opens, as shown in Figure 4-12.
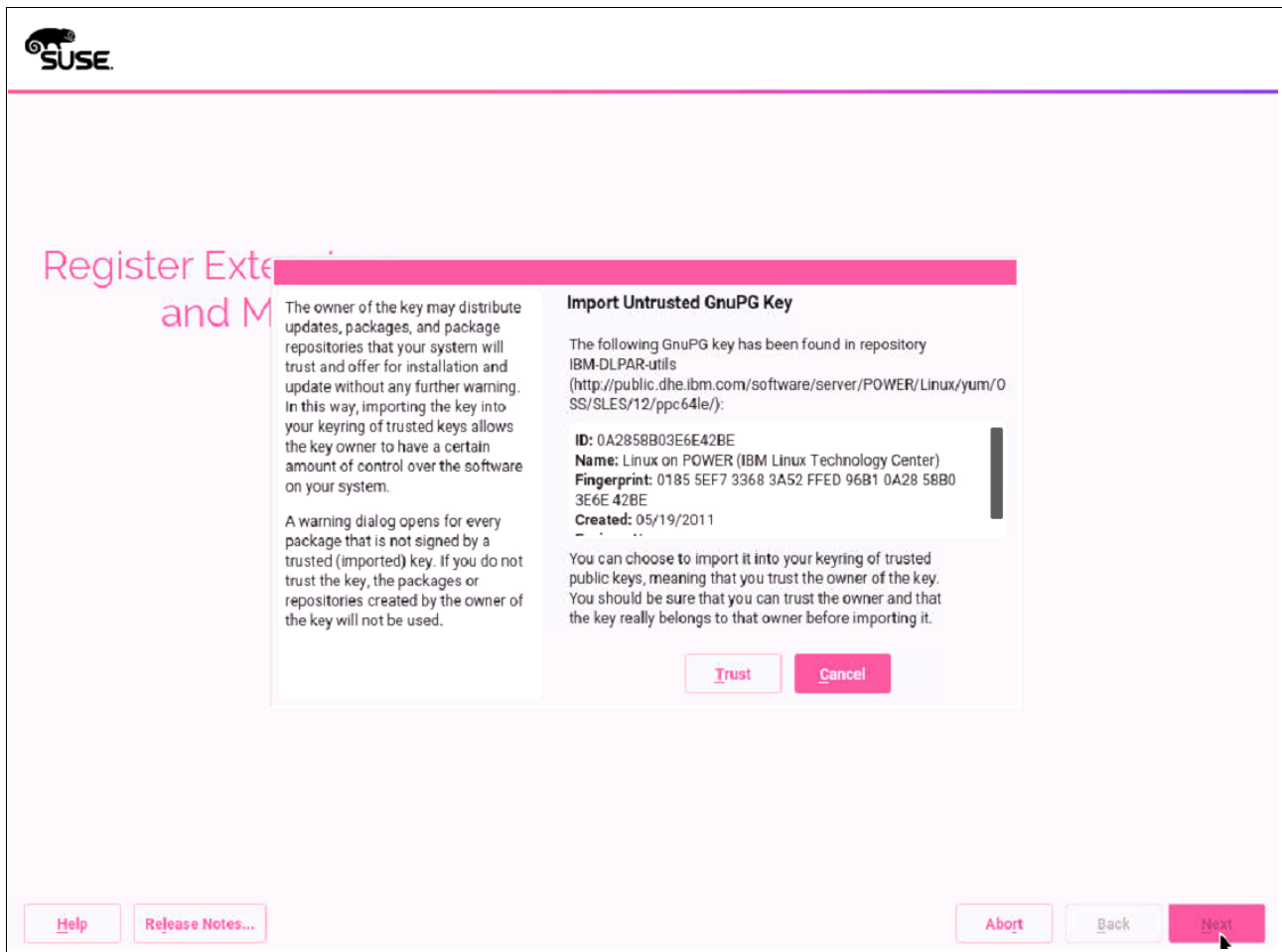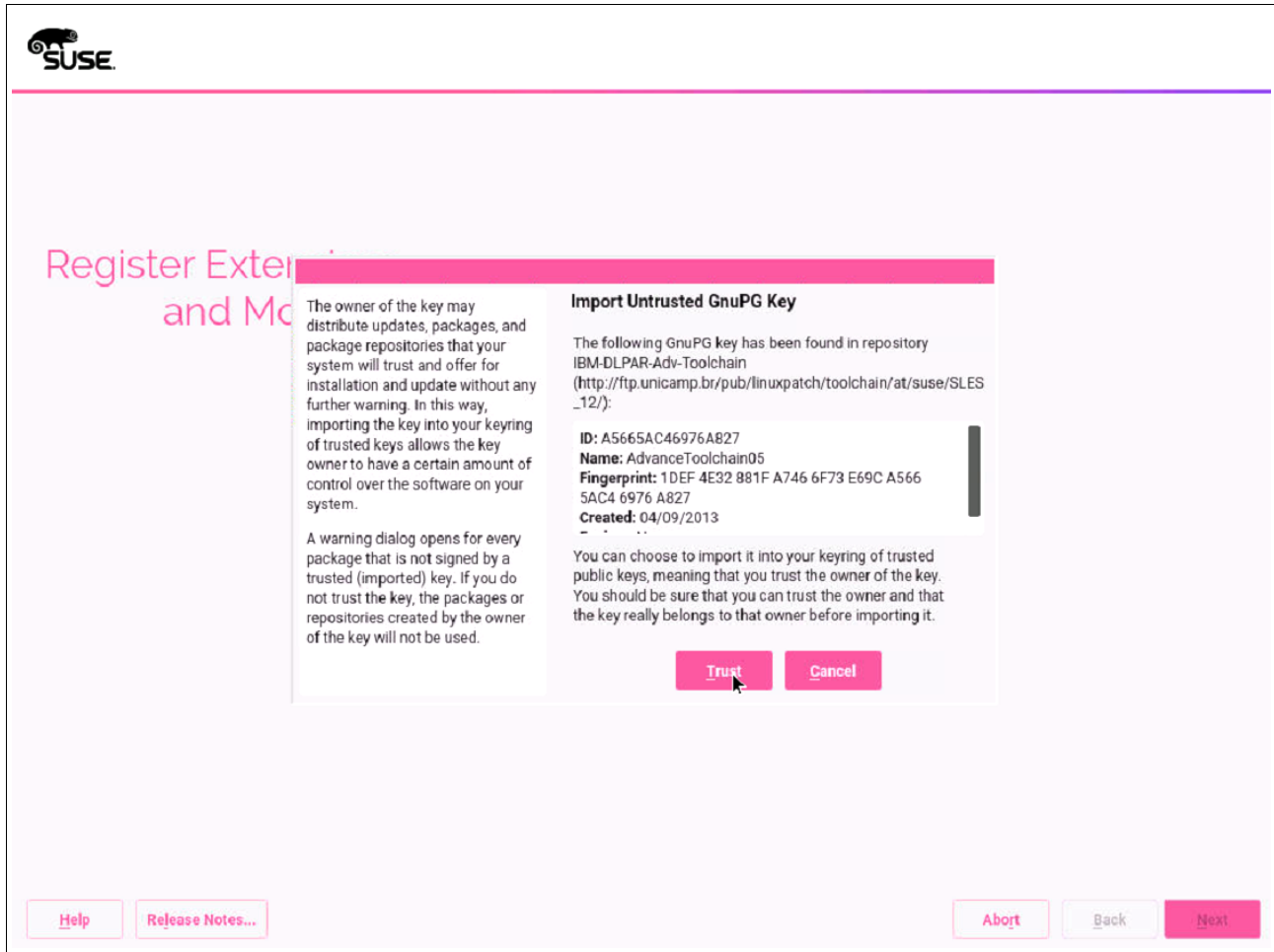

Figure 4-12   Yast2 Import GnuPG Key for IBM-DLPAR-Adv-toolchain repository window

10.After you check that the ID and Fingerprint are correct, click **Trust**. The Choose Operation System Edition selection window opens, as shown in Figure 4-13.
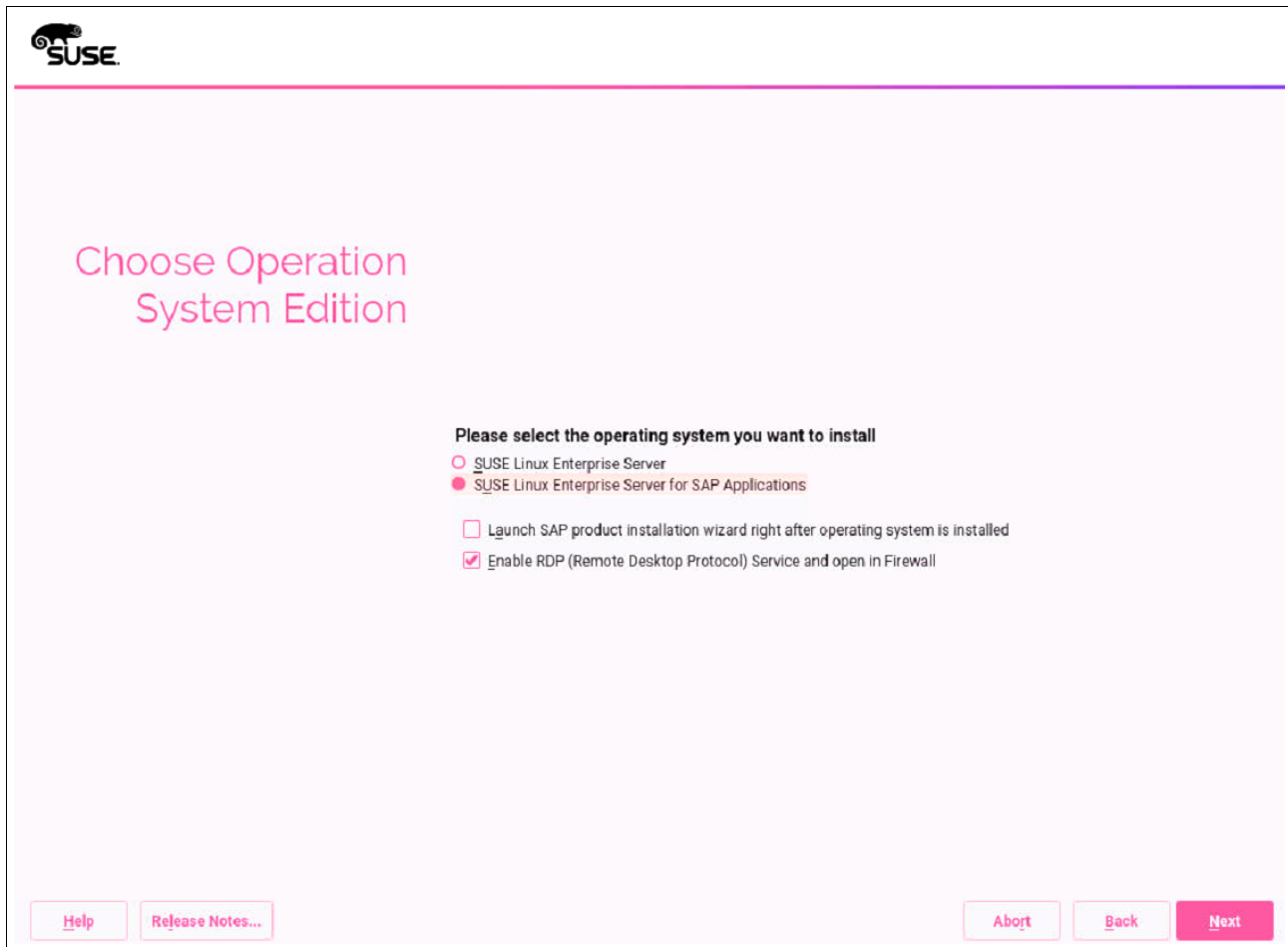
Figure 4-13   YaST2 Choose Operation System Edition window

11.Select **SUSE Linux Enterprise Server for SAP Applications**. Leave **Enable RDP (Remote Desktop Protocol) Service and open in Firewall (the default)** selected. Click **Next**. The Add-On Product Installation window opens, as shown in Figure 4-14.

> **Note:** Although RDP is not a requirement, we found that using RDP for operations is a best practice in SAP Landscape. If that is not your case, clear the check box.
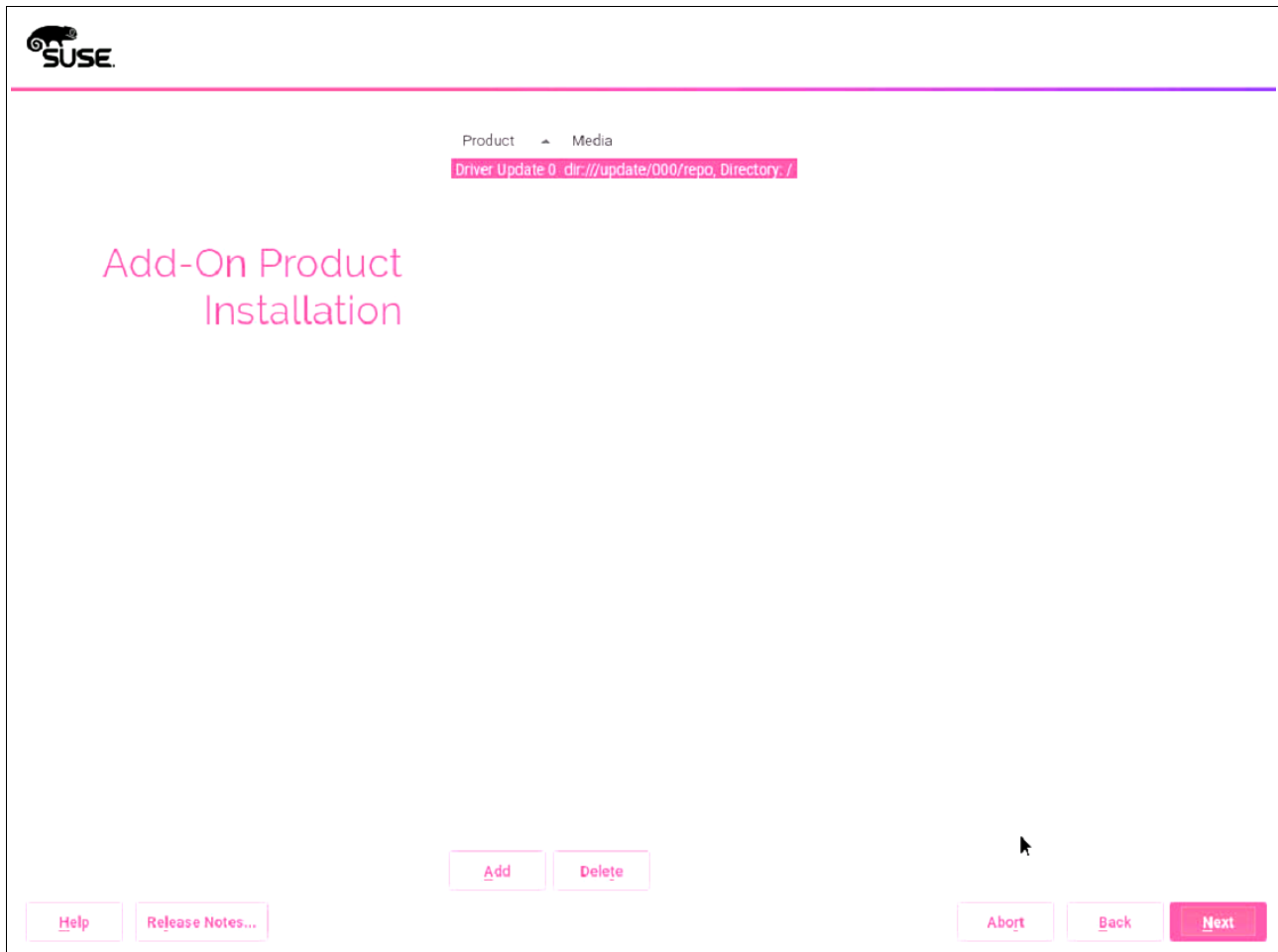
Figure 4-14   YaST2 Add-On Product Installation window

12. No changes are needed in this window. Click **Next**. The Suggested Partitioning window opens, as shown in Figure 4-15.
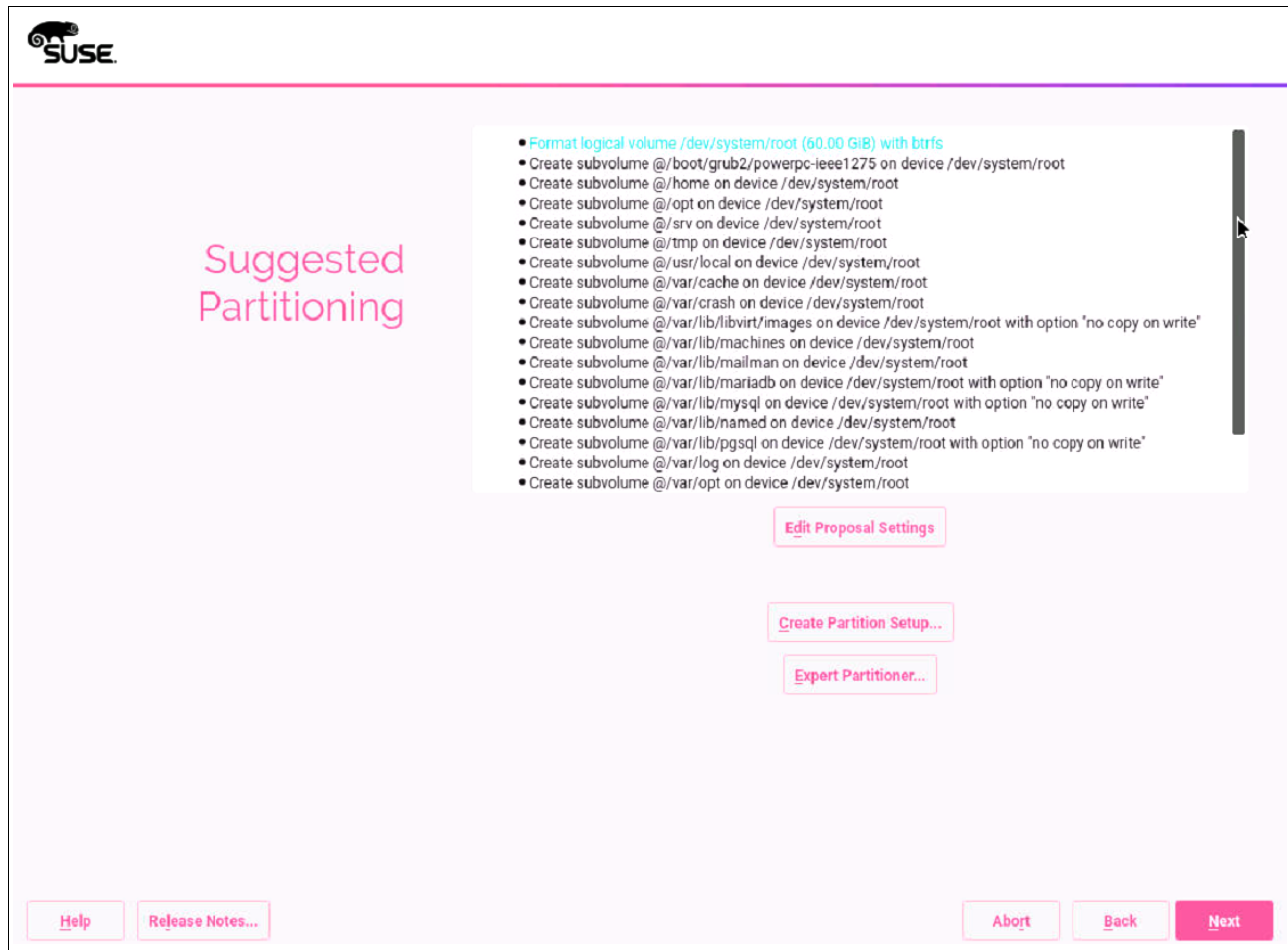
Figure 4-15   YaST2 Suggested Partitioning window

13. There is no need to change the suggested partitioning. Click **Next**.

> **Note:** Although you can change the partitioning, there is no actual requirement to do so. For consistency, use the suggested defaults for the partitioning. If you change them, use Btrfs because it is the default for SUSE V12.

The Clock and Time Zone window opens, as shown in Figure 4-16.
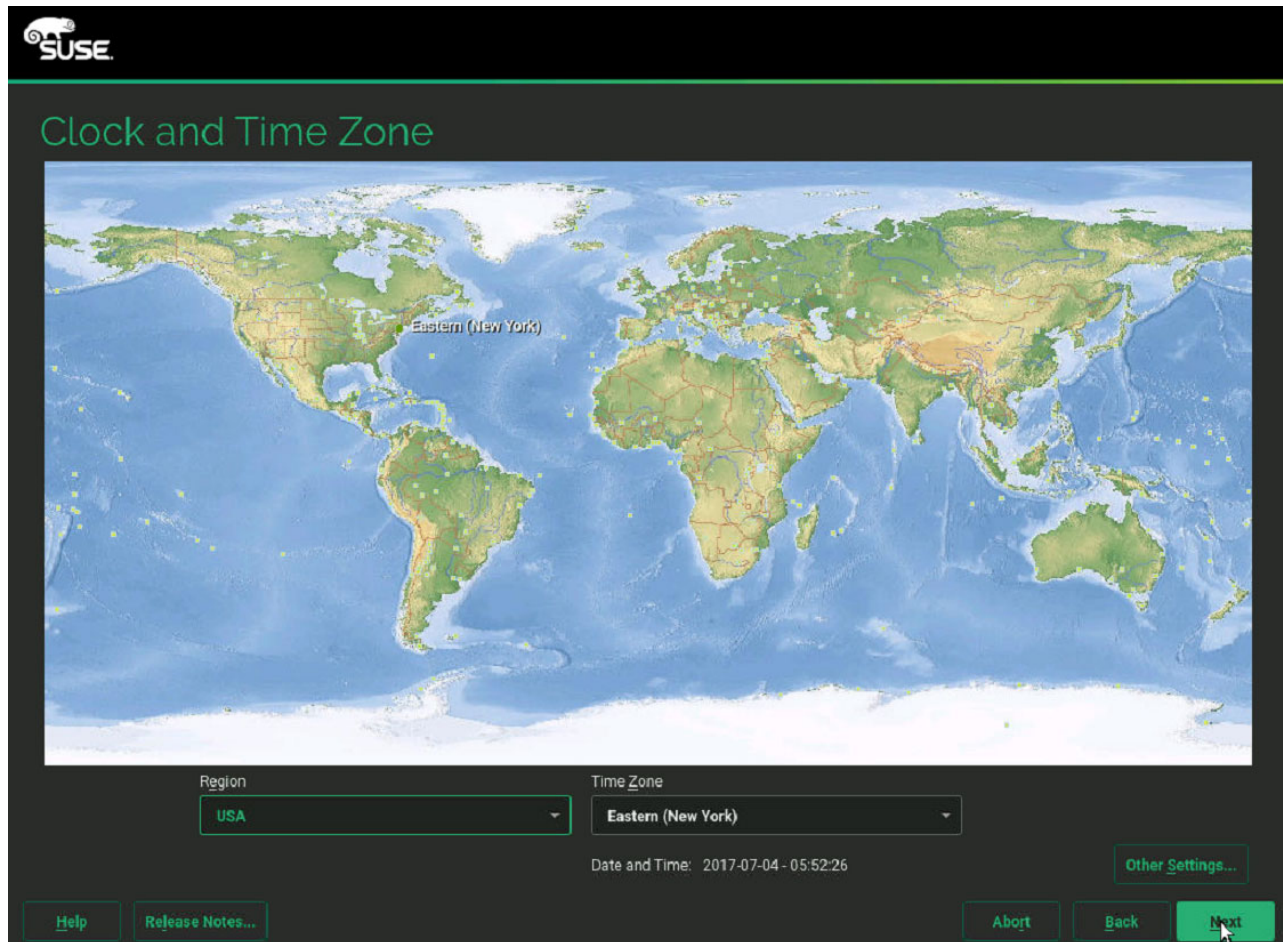
Figure 4-16   YaST2 Clock and Time Zone window

14.Select the time zone for your system and click **Next**.

> **Note:** You can click **Other Settings** and configure the Network Time Protocol (NTP) settings. However, if you join a domain, we do not show those steps at installation time to keep these instructions more generic. For more information about how to set up the NTP client after installation, see SUSE Doc: Administration Guide - Time Synchronization with NTP.

The Password for the System Administrator root window opens, as shown in Figure 4-17 on page 51.
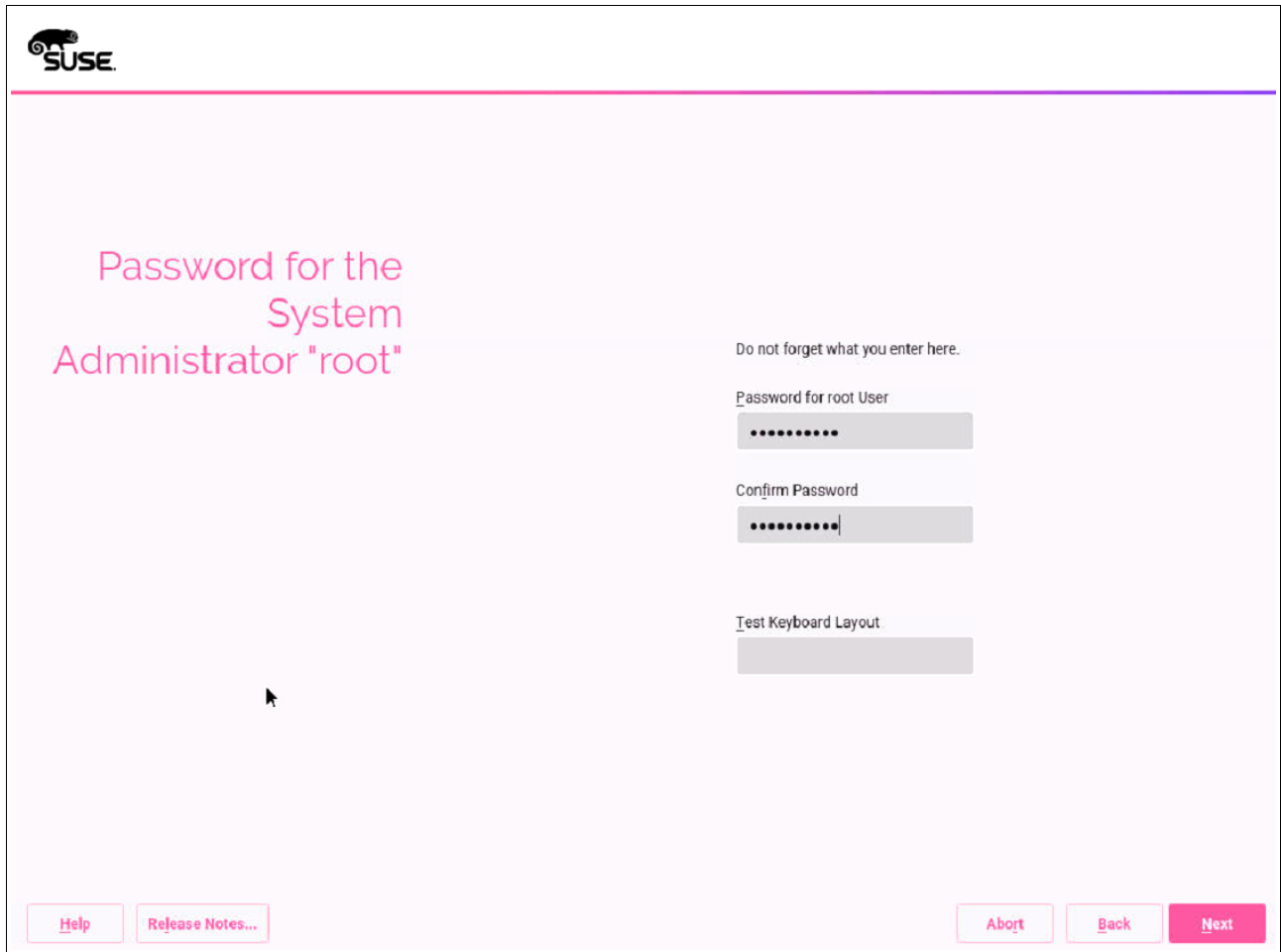
Figure 4-17 YaST2 Password for the System Administrator root window

15.After you input the password and confirm it, click **Next**.

Note: If YaST2 states that your password is weak, you see a warning. Input a stronger password.

The Installation Settings window opens, as shown in Figure 4-18.
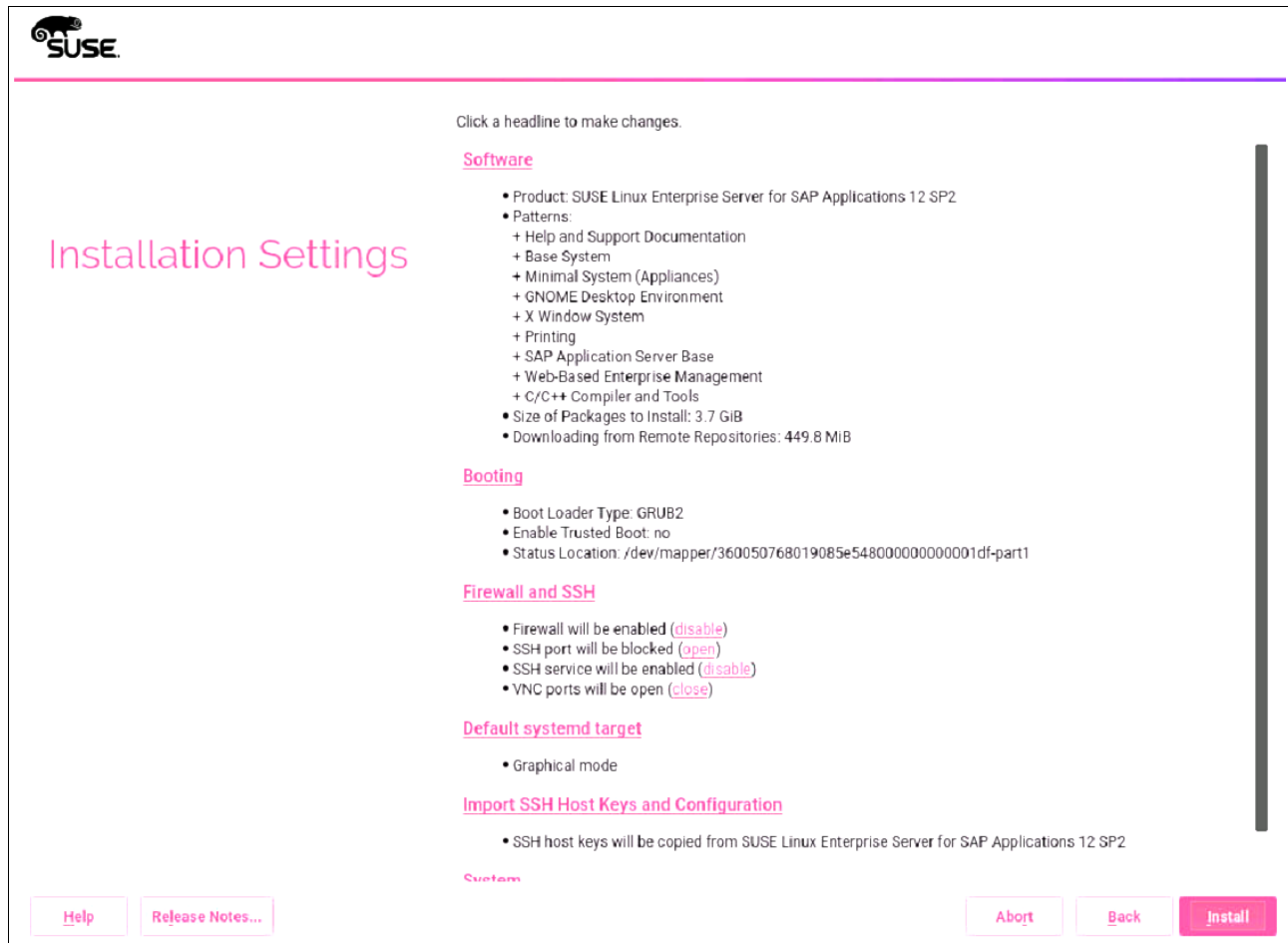
Figure 4-18   YaST2 Installation Settings default window

16.Click the **Software** link. The window that is shown in Figure 4-19 on page 53 opens.

> **Note:** In our scenario, we install Gnome and other patterns that are not listed as required but optional. You can modify the patterns that you want to install. What you select must adhere to SAP Note 1984787. Use the current version.
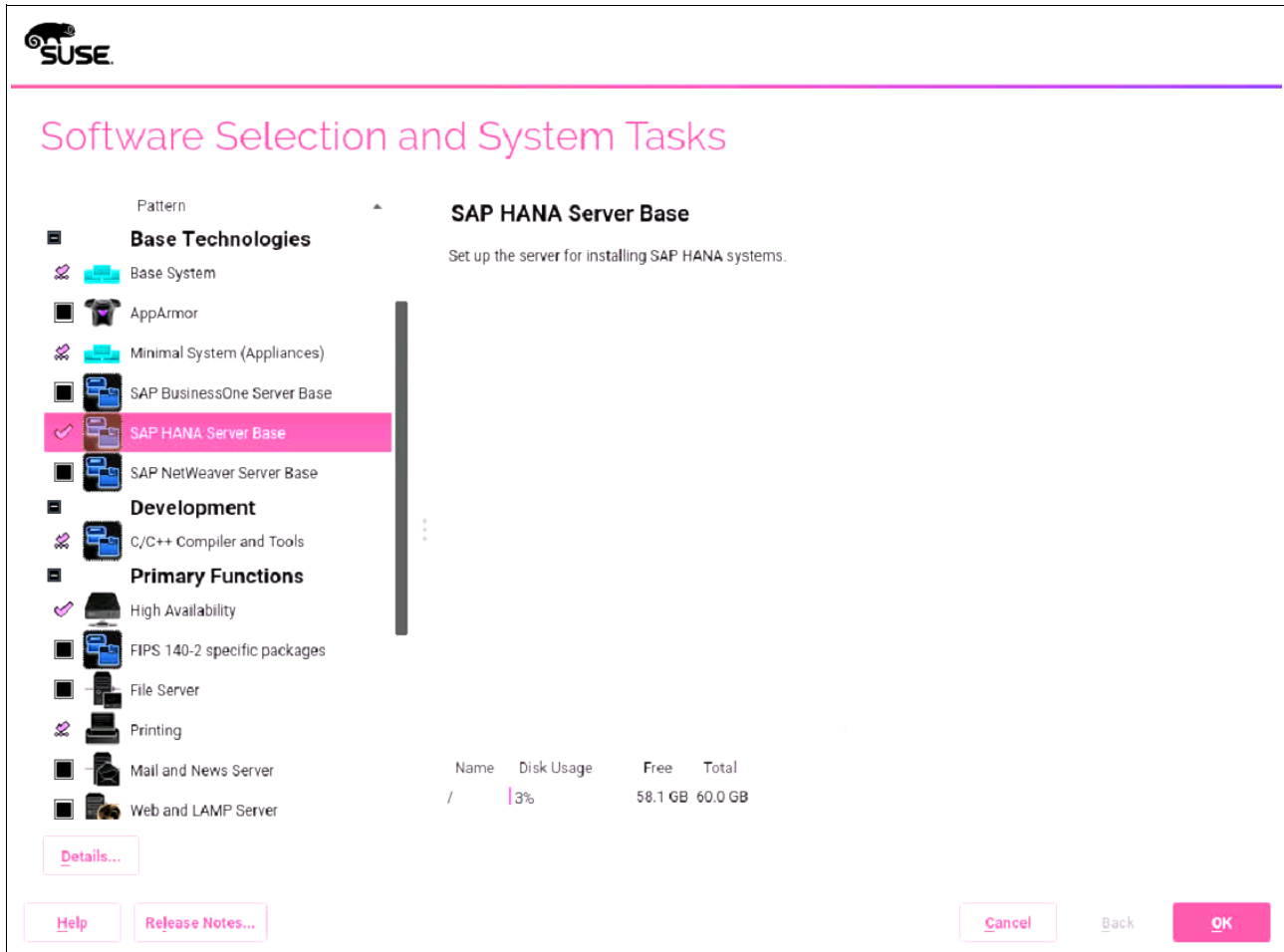
Figure 4-19   YaST2 Software window

17. In the Software selection, click the **SAP HANA Server Base** and **High Availability** patterns. Click **OK**.

> **Note:** If you are installing a stand-alone HANA DB without SUSE high availability (HA), you can select only the **SAP HANA Server Base** pattern.

Depending on the installation source on your DVD, you might see a warning when you select the SAP HANA Server Base pattern, as shown in Figure 4-20.
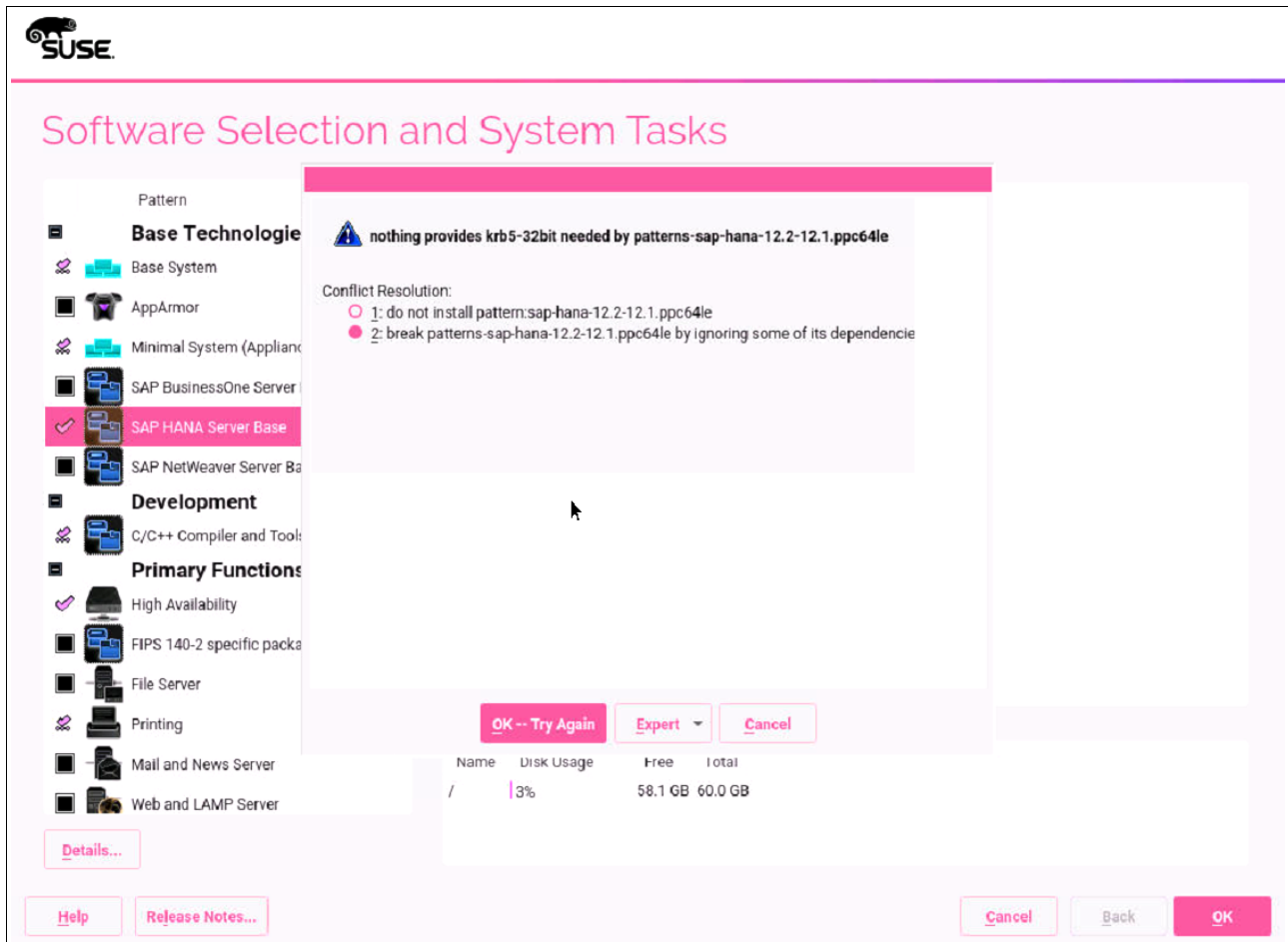
Figure 4-20   YaST2 SAP HANA Server Base pattern krb5-32-bit warning window

If this is your case, either select **2 break patterns-sap-hana-12.2.12.1.ppc64le by ignoring some of its dependencies**. Click **OK -- Try Again**, or request from SUSE an updated installation source that has this issue already fixed.

> **Note:** SUSE is aware of this issue, and at the time of writing is working on a fix for it.

18. Accept the patterns by clicking **OK**. You are back in the Software window but with the patterns selected, as shown in Figure 4-21 on page 55.
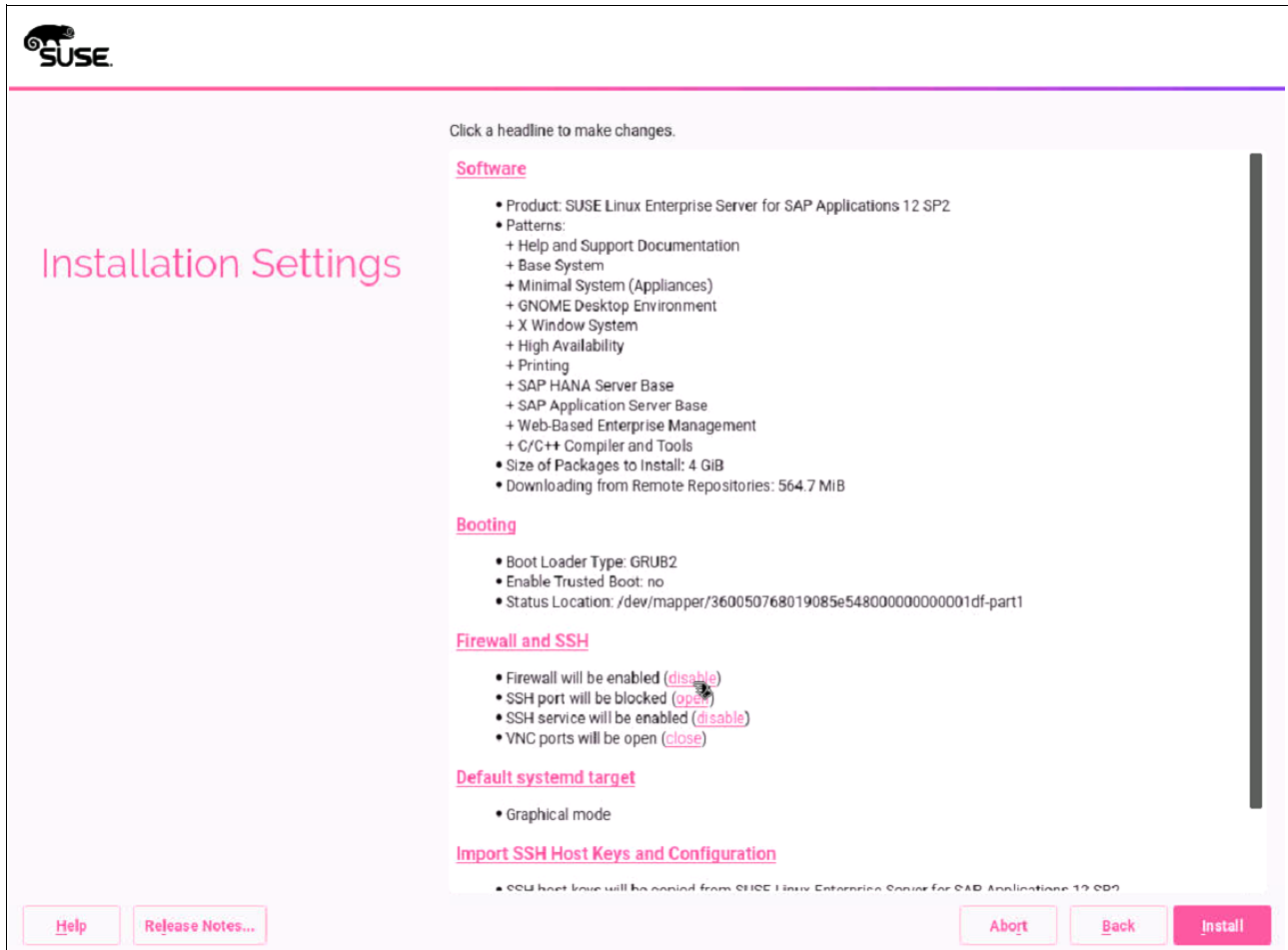
Figure 4-21   YaST2 Software with patterns selected window

19. Click **disable** for the Firewall and SSH menu. If your HANA installation requires hardening at the OS level, see *Operating System Security Hardening Guide for SAP HANA*.

After the firewall is disabled, you see the window that shown in Figure 4-22.
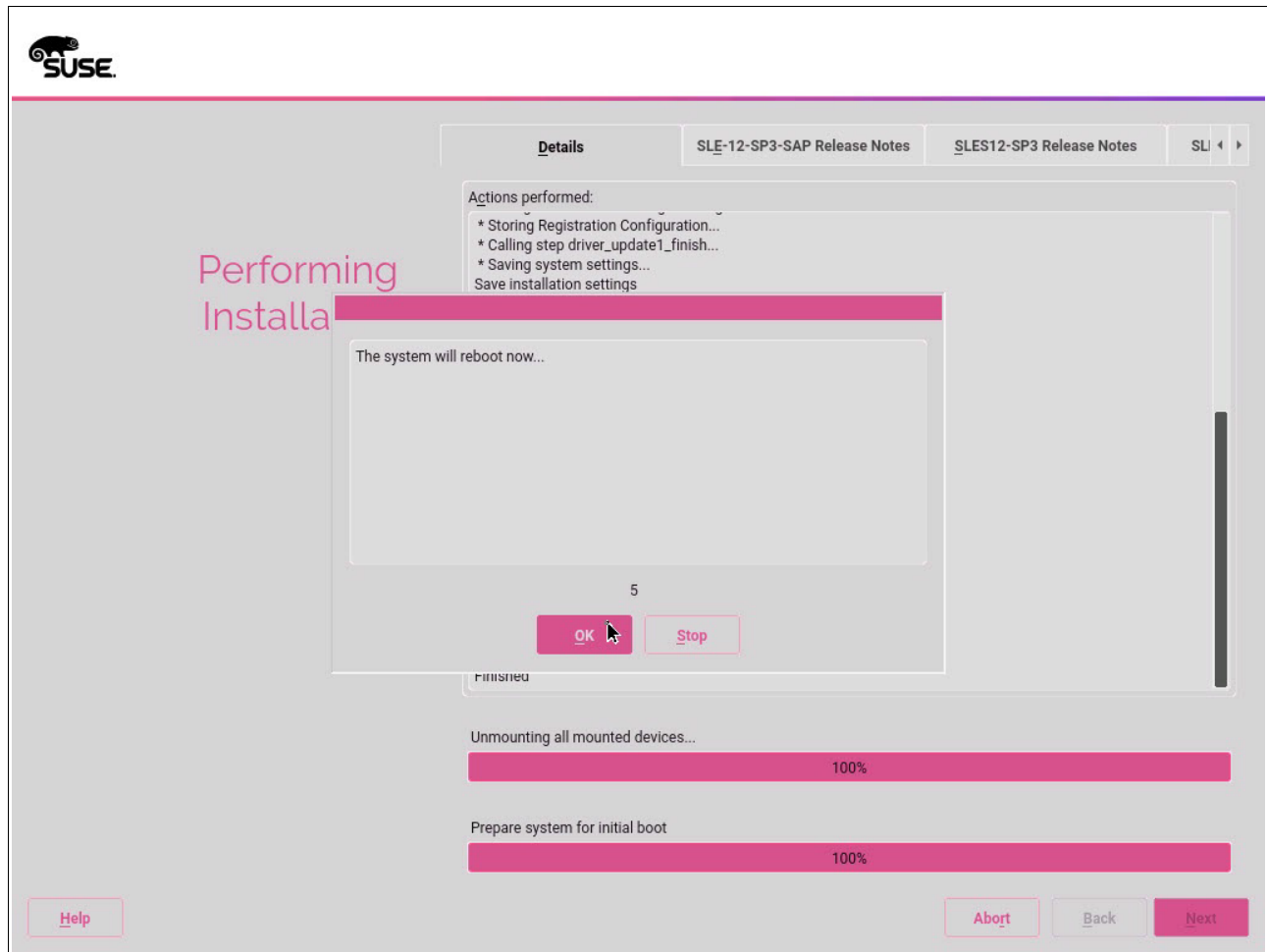
Figure 4-22   YaST2 Software window all selections for installation

20. Click **Install** to start the installation. On the confirmation window, click **Install** if you are ready to perform the installation.

    After several minutes, the installer completes and updates the system. The system also restarts so that you can then connect by using RDP or SSH when the installation completes.

Installing the service and productivity tools for SUSE Linux Enterprise Server

If you did not install the service and productivity tools by using the **Extension and Module Selection** selection that is shown in Figure 4-8 on page 42, then you must install the Service and productivity tools for Linux on Power Servers on the LPAR that you installed. To do so, you must download the binary files from the website and follow the instructions to install the repositories. In this scenario, we download the binary files directly to the LPAR and install them as shown in Example 4-14.

Example 4-14   Installing the service and productivity tools RPM

```
hana001:~ # rpm -vih
http://public.dhe.ibm.com/software/server/POWER/Linux/yum/download/ibm-power-repo-
latest.noarch.rpm
```

```
Retrieving
http://public.dhe.ibm.com/software/server/POWER/Linux/yum/download/ibm-power-repo-
latest.noarch.rpm
warning: /var/tmp/rpm-tmp.kejIvd: Header V4 DSA/SHA1 Signature, key ID 3e6e42be:
NOKEY
Preparing...                          ################################# [100%]
Updating / installing...
   1:ibm-power-repo-3.0.0-17          ################################# [100%]
```

After the file is installed in your system, you must run the **/opt/ibm/lop/configure** command and accept the license.

Regardless, you can add the files by using the CLI or during the BOS installation. After the IBM repositories are visible (run the **zypper lr** command), you can install the needed software. Follow the instructions for Tools Installation instructions for SUSE. In this scenario, because the LPAR is managed by an HMC, we install the tools that are shown in Example 4-15.

Example 4-15   Installing the service and productivity tools binary files

```
redhana001:~ # zypper install ibm-power-managed-sles12
Refreshing service
'SUSE_Linux_Enterprise_Server_for_SAP_Applications_12_SP3_ppc64le'.
Loading repository data...
Reading installed packages...
Resolving package dependencies...

The following 12 NEW packages are going to be installed:
  DynamicRM IBMinvscout devices.chrp.base.ServiceRM ibm-power-managed-sles12
ibmPMLinux librtas1 powerpc-utils-python rsct.basic rsct.core rsct.core.utils
  rsct.opt.storagerm src

The following 12 packages are not supported by their vendor:
  DynamicRM IBMinvscout devices.chrp.base.ServiceRM ibm-power-managed-sles12
ibmPMLinux librtas1 powerpc-utils-python rsct.basic rsct.core rsct.core.utils
  rsct.opt.storagerm src

12 new packages to install.
Overall download size: 26.3 MiB. Already cached: 0 B. After the operation,
additional 85.4 MiB will be used.
Continue? [y/n/...? shows all options] (y): y
Retrieving package src-3.2.2.1-17038.ppc64le
(1/12), 307.5 KiB (  1.1 MiB unpacked)
Retrieving: src-3.2.2.1-17038.ppc64le.rpm
.............................................................................
...........[done (16.3 KiB/s)]

[ snip ]

(12/12) Installing: ibm-power-managed-sles12-1.3.1-0.ppc64le
.............................................................................
.....[done]
```

Now, you can use dynamic LPAR operations on this LPAR for adding or removing devices, such as CPU or memory.

Network tuning for SUSE Linux Enterprise Server

For network tuning, follow the instructions in *Configuring platform largesend for SAP HANA on SUSE Linux Enterprise Server with VIOS*. You must complete the steps first on the VIOS, as described in 3.2, "Virtual I/O Server" on page 20, before doing the tuning at the SUSE LPAR level.

> **Note:** Enabling an MTU higher than 1500 can make some hosts unreachable or cause severe performance degradation due to network configurations across the LAN. Check that all the needed configurations are done before enabling jumbo frames. After you enable MTU 9000, check it by running the `ping -M do -s 8972 [destinationIP]` command. If you cannot reach the host with the `ping` command, it means that fragmentation happened, which means that the MTU change worsened the performance. MTU changes are not changes that are done only at the OS level; they must be incorporated with the settings in all participating devices on the flow.

For single root input/output virtualization (SR-IOV), vNIC is not supported. However, you can use SR-IOV-capable adapters. For more information about this topic and future developments, see SAP HANA on IBM Power Systems and IBM System Storage - Guides.

Configuring the Network Time Protocol client for SUSE Linux Enterprise Server

If you did not configure the NTP client at installation time (see Figure 4-16 on page 50), you must do so before installing the HANA software. In this scenario, we use two NTP servers that we configure manually. For more information and other configuration options, see SUSE Doc: Administration Guide - Manually Configuring NTP in the Network.

Complete the following steps:

1. Add the IP addresses of the servers to the `/etc/ntp.conf` file, as shown in Example 4-16.

Example 4-16   Adding Network Time Protocol IP servers to the ntp.conf file

```
hana001:~ # echo server 10.10.12.10 iburst >> /etc/ntp.conf
hana001:~ # echo server 10.10.12.9 iburst >> /etc/ntp.conf
```

2. Enable, start, and query the NTP service by running the `systemctl` command, as shown in Example 4-17.

Example 4-17   The systemcl enable, start, and query Network Time Protocol commands

```
hana001:~ # systemctl enable ntpd
Created symlink from /etc/systemd/system/multi-user.target.wants/ntpd.service
to /usr/lib/systemd/system/ntpd.service.
hana001:~ # systemctl start ntpd
hana001:~ # systemctl status ntpd
? ntpd.service - NTP Server Daemon
   Loaded: loaded (/usr/lib/systemd/system/ntpd.service; enabled; vendor
preset: disabled)
  Drop-In: /run/systemd/generator/ntpd.service.d
           ··50-insserv.conf-$time.conf
   Active: active (running) since Wed 2017-07-12 17:05:14 EDT; 1s ago
     Docs: man:ntpd(1)
  Process: 42347 ExecStart=/usr/sbin/start-ntpd start (code=exited,
status=0/SUCCESS)
 Main PID: 42354 (ntpd)
    Tasks: 2 (limit: 512)
```

```
      CGroup: /system.slice/ntpd.service
              ··42354 /usr/sbin/ntpd -p /var/run/ntp/ntpd.pid -g -u ntp:ntp -c
/etc/ntp.conf
              ··42355 ntpd: asynchronous dns resolver

Jul 12 17:05:14 hana001 systemd[1]: Starting NTP Server Daemon...
Jul 12 17:05:14 hana001 ntpd[42353]: ntpd 4.2.8p10@1.3728-o Thu May 18 14:00:11
UTC 2017 (1): Starting
Jul 12 17:05:14 hana001 ntpd[42353]: Command line: /usr/sbin/ntpd -p
/var/run/ntp/ntpd.pid -g -u ntp...conf
Jul 12 17:05:14 hana001 ntpd[42354]: proto: precision = 0.072 usec (-24)
Jul 12 17:05:14 hana001 ntpd[42354]: switching logging to file /var/log/ntp
Jul 12 17:05:14 hana001 start-ntpd[42347]: Starting network time protocol
daemon (NTPD)
Jul 12 17:05:14 hana001 systemd[1]: Started NTP Server Daemon.
Hint: Some lines were ellipsized, use -l to show in full.
```

3. Query the servers by running the **ntpq** command. You can see that both NTP servers were contacted, as shown in Example 4-18.

Example 4-18   Querying the Network Time Protocol servers

```
hana001:~ # ntpq
ntpq> host localhost
current host set to localhost
ntpq> peers
     remote           refid      st t when poll reach   delay   offset  jitter
==============================================================================
*stg-ad02.stg.fo 62.236.120.71    3 u    9   64    3   0.827   -2.343   1.773
+stg-ad03.stg.fo 91.207.136.50    3 u   11   64    3   1.564    5.978  15.375
ntpq> exit
```

5

Storage and file systems setup and configuration

This chapter describes the storage layout and file system setup for an SAP HANA on IBM Power Systems installation, and includes both scale-up and scale-out scenarios.

This chapter describes step-by-step implementation instructions for:

► Customizing multipathing with recommended settings

► Customizing multipathing with aliases

► Creating the logical volume manager (LVM) groups and volumes with recommended settings

► Creating the file systems with recommended settings

► Applying Linux I/O subsystem recommended settings

This chapter covers the following topics:

► Storage layout
► Linux multipath setup
► File system creation and setup
► More Linux I/O subsystem tuning

5.1 Storage layout

Before you start, review the following key aspects:

► In 2.1.1, "Storage and file system requirements" on page 15, you set up three file systems for HANA, including `/usr/sap`.

► For this publication, our lab environments are configured with 140 GB of memory, so each one of our logical partitions (LPARs) use, according to 2.1.1, "Storage and file system requirements" on page 15[1], the following specifications:

  – Four 100 GB LUNs for the HANA data area
  – Four 35 GB LUNs for the HANA log area
  – One 50 GB LUN for `/usr/sap`

> **Important:** The log LUNs must either be on solid-state drives (SSDs) or flash arrays, or the storage subsystem must provide a low-latency write cache area.

► The LUNs are locally attached to the node, and *each* node has the number of LUNs described before, for either a scale-up or scale-out scenario. For a scale-out scenario, you must share each data and log LUN to each of the participating scale-out cluster nodes, which means that you must zone and assign all data and log LUNs to all nodes. For more information, see 5.3.2, "File systems for scale-out systems" on page 76.

► Regarding the HANA shared area, there are different approaches depending on whether you are working with a scale-up or with a scale-out scenario.

5.1.1 HANA shared area storage layout for scale-up systems

The HANA shared area is basically where the HANA binary files are stored, along with HANA configuration files.

Scale-up systems do not need to share this area with any other nodes, so the simplest approach is to create a local file system by using locally attached storage disks. In our scale-up lab systems, in addition to the LUNs that are described in 5.1, "Storage layout" on page 62, we also add one 50 GB LUN for HANA that is shared to the LPAR.

If you are planning on implementing scale-up systems, see 5.2, "Linux multipath setup" on page 66.

5.1.2 HANA shared area storage layout for scale-out systems

HANA scale-out clusters do not use a local HANA shared area. Instead, they share a file system as a shared area, which means that the binary and configuration files are shared among the participating nodes. This sounds logical because all nodes must have the same software version for HANA and the same configuration.

> **Note:** The size of the HANA shared file system in a scale-out configuration is 1x the amount of RAM per every four HANA worker nodes. For example, if you have nodes with 140 GB of memory and you have up to four worker nodes, your shared file system size is 140 GB. If you have 5 - 8 working nodes, the size is 280 GB, and so on.

---

[1] Our systems use bigger data and log areas than the guidelines that are presented. Those rules are minimum values.

Scale-out systems have the following alternatives for sharing the HANA shared area:

- ► Place the `/hana/shared` file system on a highly available Network File System (NFS) server and have each scale-out cluster connect to it over the network.
- ► Create an IBM Spectrum Scale cluster with the scale-out nodes, and create an IBM Spectrum Scale file system for the `/hana/shared` area.

Both alternatives yield the same result, which ensures all scale-out nodes can access the contents of the `/hana/shared` file system. Be careful because this file system must be protected by high availability (HA) technologies, or it becomes a single point of failure (SPOF) for the entire scale-out cluster.

You might already have an existing NetWeaver shared file system in your landscape. If so, consider taking advantage of it.

HANA shared area that is managed with IBM Spectrum Scale

If you use IBM Spectrum Scale to create and manage the HANA shared file system, you can opt to create one LUN per node for the HANA shared area and map them to all scale-out nodes. Then, you create the IBM Spectrum Scale cluster and use these shared LUNs to create the Network Shared Disks (NSDs), and then create an IBM Spectrum Scale file system on them.

> **Tip:** The minimum IBM Spectrum Scale version is 5.0.0 because prior versions caused `hdblcm` to run into issues due to wrong reported subblock sizes.. For more information, see *SAP HANA and ESS: A Winning Combination*, REDP-5436.

Recall that an IBM Spectrum Scale cluster with shared LUNs on all nodes provides a highly available file system on those LUNs. Even if a node in the scale-out cluster fails, the `/hana/shared` content is still accessible by all the remaining nodes. This is a reliable architecture, not complex to implement, and it does not need for any environments to be set up externally to the HANA scale-out nodes themselves.

HANA shared area that is managed with Network File System

This approach is probably the easiest if you already have a highly available NFS server infrastructure in your enterprise environment. If you do not have one, you *must* ensure that you create an NFS server with an HA infrastructure.

> **Caution:** You *must* use an NFS server with HA capability or your overall HANA scale-out cluster will be in jeopardy because of a SPOF in your NFS server system.

Here are services that you must consider when implementing a highly available NFS server topology in your environment:

- ► Highly Available NFS service with DRBD and Pacemaker with SUSE Linux Enterprise High Availability Extension
- ► Highly Available NFS service on AIX with PowerHA SystemMirror
- ► High Availability NFS service with DRBD and IBM Tivoli System Automation for Multiplatform (SA MP) with SUSE Linux Enterprise High Availability Extension
- ► SAP HANA on NetApp Systems with NFS

For more information about how to set up the NFS server *export* parameters and the clients *mount* parameters, see 5.3.2, "File systems for scale-out systems" on page 76.

HANA shared area that is managed with Elastic Storage Server

Another alternative to host the HANA shared area is to use scale-out nodes on an IBM Elastic Storage™ Server. For more information, see *SAP HANA and ESS: A Winning Combination*, REDP-5436.

5.1.3 Probing for newly attached disks

If you have not yet attached the data, log, shared, and `/usr/sap` disks to your system, this section provides a quick and easy way to recognize them without requiring a system restart. Not all Linux distributions provide a system management tool, so in this scenario we provide an approach that works with all distributions.

> **Note:** If IBM XIV®, IBM Spectrum Accelerate™, IBM FlashSystem® A9000R or IBM FlashSystem A9000 storage systems are used to provide storage for the HANA environment, install the IBM Storage Host Attachment Kit. The IBM Storage Host Attachment Kit is a software pack that simplifies the task of connecting a host to supported IBM storage systems, and it provides a set of command-line interface (CLI) tools that help host administrators perform different host-side tasks.
>
> For more information, see IBM Storage Host Attachment Kit welcome page.

The `multipath -ll` (double lowercase L) command shows the storage disks that are attached to your system and the paths to access them. Example 5-1 shows the output of one of our lab systems that contains only the operating system (OS) installation disk that is attached to it.

Example 5-1   Output of the multipath -ll command

```
hanaonpower:~ # multipath -ll
2001738002ae12c88 dm-0 IBM,2810XIV 1
size=64G features='2 queue_if_no_path retain_attached_hw_handler' hwhandler='1
alua' wp=rw 2
`-+- policy='service-time 0' prio=50 status=active 3
  |- 2:0:0:1 sda 8:0   active ready running
  |- 2:0:1:1 sdb 8:16  active ready running
  |- 3:0:0:1 sdc 8:32  active ready running
  |- 3:0:1:1 sdd 8:48  active ready running
  |- 4:0:0:1 sde 8:64  active ready running
  |- 4:0:1:1 sdf 8:80  active ready running
  |- 5:0:0:1 sdg 8:96  active ready running
  `- 5:0:1:1 sdh 8:112 active ready running
```

First, check the line that is marked as **1** in the example. In that line, you can identify the LUN ID of the target disk, which in this case is `2001738002ae12c88`. Also, the disk device name is *dm-0*. The characters *dm* are the standard Linux nomenclature for devices that are managed by the *device mapper* service. The output tells you that this is an IBM XIV LUN. So, in summary, you know that the XIV LUN with ID `2001738002ae12c88` is mapped as `/dev/dm-0` in that system. As this is the only disk you have so far, you know that this is the OS installation disk.

Line **2** in the output shows some characteristics of that disk. The important information to notice is the disk size, which in this case is 64 GB.

Finally, line **3** onwards displays each one of the paths over which the disk is accessed. Our system has four Fibre Channel adapters, and the XIV storage has two controllers, so you see each disk through eight paths. Linux uses an ordinary disk device name for each path by which a disk is accessed, and then joins those disks under a device mapper entry. In our case, the `sda` through `sdh` devices are the same OS installation disk that is seen by each of the eight paths, and then joined under the `dm-0` device.

> **Note:** By design, all paths to the XIV storage devices are active. Other storage device types, such as IBM Spectrum Virtualize™ (formerly IBM SAN Volume Controller) work with active and passive paths.

You are most likely asking yourself why we have not yet mapped all of the other HANA LUNs to the system. The answer is based on experience. If you do so, the SUSE Linux Enterprise Server disk probing mechanisms during the OS installation probe all the disks multiple times: once during start, another during the initial disk partitioning layout suggestion, and another one each time you change the partitioning scheme. Probing multiple disks multiple times is a time-consuming task. So, to save time, attach the HANA LUNs after the OS is installed.

To probe for newly attached disks, complete the following steps:

1. Attach the other LUNs to your system by using your storage management configuration tool.

2. Run the `rescan-scsi-bus.sh` command. This command sends a loop initialization protocol (lip) signal to probe for new disks on each one of the Fibre Channel connections. This is essentially a bus scan operation that removes and adds devices to the scanned targets, as shown in Example 5-2.

Example 5-2   Issuing rescan-scsi-bus.sh to probe for newly added disks

```
hanaonpower:~ # rescan-scsi-bus.sh
```

3. Now, you can run the `multipath -ll` command again and verify that all of your disks appear in the listing. Send the output to the `grep` command to make the output shorter. If you want to see all of the output, run `multipath -ll` alone. Check that all the disks are there. Example 5-3 shows the output after attaching the remaining LUNs.

Example 5-3   Output of multipath -ll with all LUNs attached to the system

```
hanaonpower:~ # multipath -ll | grep dm-
2001738002ae12c8a dm-6 IBM,2810XIV
2001738002ae12c89 dm-5 IBM,2810XIV
2001738002ae12c88 dm-0 IBM,2810XIV
2001738002ae12c8f dm-10 IBM,2810XIV
2001738002ae12c8e dm-8 IBM,2810XIV
2001738002ae12c8d dm-9 IBM,2810XIV
2001738002ae12c8c dm-7 IBM,2810XIV
2001738002ae12c8b dm-4 IBM,2810XIV
2001738002ae12c92 dm-13 IBM,2810XIV
2001738002ae12c91 dm-12 IBM,2810XIV
2001738002ae12c90 dm-11 IBM,2810XIV
```

Notice in Example 5-3 that we have a total of 11 disks: `dm-0`, `dm-4`, `dm-5`, `dm-6`, `dm-7`, `dm-8`, `dm-9`, `dm-10`, `dm-11`, `dm-12`, and `dm-13`. Also, these names can change when the system restarts. It is *not* a best practice to rely on the device naming. A better approach is to use *aliases* for the disks. Section 5.2, "Linux multipath setup" on page 66 provides information about how to use aliases for the disks.

5.2 Linux multipath setup

Linux systems access disks over multiple paths by using its multipathing service. The configuration file that controls access is /etc/multipath.conf. This file is not created during the default Linux installation, so it must be created and populated based on your environment.

Each storage subsystem vendor has its own best practices for setting up the multipathing parameters for Linux. Check your storage vendor's documentation for their specific recommendations.

The first piece of information that you must know about the /etc/multipath.conf file is that it is composed of four sections:

► The defaults section: Contains the parameters values that are applied to the devices.

► The blacklist section: Excludes the devices in this list from having the default or specific device parameter values applied.

► The multipaths section: Defines aliases for the disks.

► The devices section: Overrides the default parameters section to enable specific parameter values according to the storage in use.

Example 5-4 is a fully functional multipath.conf file that is built for an IBM XIV Storage System. We reference this file throughout this chapter to explain the concepts behind it.

Example 5-4   A multipath.conf file for IBM XIV storage

```
hanaonpower:~ # cat /etc/multipath.conf
defaults {
        user_friendly_names     yes
}

blacklist {
    devnode "^(ram|raw|loop|fd|md|dm-|sr|scd|st)[0-9]*"
    devnode "^(hd|xvd|vd)[a-z]*"
}

multipaths {

    #ROOTVG
    multipath {
        wwid 2001738002ae12c88
        alias ROOTVG
    }

    #USRSAP
    multipath {
        wwid 2001738002ae12c89
        alias USRSAP
    }

    #HANA DATA
    multipath {
        wwid 2001738002ae12c8c
        alias HANA_DATA_1_1
    }

    multipath {
        wwid 2001738002ae12c8d
        alias HANA_DATA_1_2
```

```
        }

    multipath {
          wwid 2001738002ae12c8e
          alias HANA_DATA_1_3
    }

    multipath {
          wwid 2001738002ae12c8f
          alias HANA_DATA_1_4
    }

    #HANA LOG
    multipath {
          wwid 2001738002ae12c8a
          alias HANA_LOG_1_1
    }

    multipath {
          wwid 2001738002ae12c90
          alias HANA_LOG_1_2
    }

    multipath {
          wwid 2001738002ae12c91
          alias HANA_LOG_1_3
    }

    multipath {
          wwid 2001738002ae12c92
          alias HANA_LOG_1_4
    }


    #HANA SHARED
    multipath {
          wwid 2001738002ae12c8b
          alias HANA_SHARED01
    }
}

devices {
        device {
                vendor "IBM"
                product "2810XIV"
                path_selector "round-robin 0"
                path_grouping_policy multibus
                rr_min_io 15
                path_checker tur
                failback 15
                no_path_retry queue
        }
}
```

Note: The `mutlipath.conf` file that is shown here is for the IBM XIV Storage System. All storage system vendors have their own recommended settings for different storage systems and OSes that must be followed. Appendix B, "Example of a multipath.conf file for SAP HANA systems" on page 153 covers the IBM Spectrum Virtualize storage system recommended `multipath.conf` settings for Linux OS.

Let us analyze each one of the four sections of this `multipath.conf` file.

The defaults section

The defaults section in `multipath.conf` contains the parameters and values that are applied to all disks in the system. In this case, the only one that is used in our HANA environments is the **user_friendly_names** parameter, which is set to `yes`.

The **user_friendly_names** parameter instructs the system to either not use friendly names (`no`, which is the default option) and instead use the LUN WWID to name the devices under the `/dev/mapper` directory, or use some better naming, such as `/dev/mapper/mpath<n>` when using friendly names. Even though the form `mpath<n>` looks better than the LUN WWID for listing the disks, we still recommend overriding the disk names with custom aliases entries, as described in "The multipaths section" on page 68.

For more information about **user_friendly_names**, see SUSE Doc: Storage Administration Guide - Configuring User-Friendly Names or Alias Names.

The blacklist section

The blacklist section instructs the system to ignore the devices in the list when applying the multipath settings that are defined in the `multipath.conf` file. Usually, devices that are not multipathed are put into this list. Example 5-4 on page 66 has a standard blacklist that is used in implementations in the field.

For more information about blacklisting devices, see SUSE Doc: Storage Administration Guide - Blacklisting Non-Multipath Devices.

The multipaths section

This section is where you define aliases for your disks. Although aliasing does not bring any performance advantage to your setup, it brings ease of management and avoids problems with device renaming across system restarts. Consider using aliases in your implementation.

Example 5-4 on page 66 creates a `multipath { }` entry for each one of our 11 disks. Each entry contains a description for the LUN WWID and then the alias we want to assign to it. How do you know which LUN is supposed to be used with each alias? It was probably your storage admin who created the LUNs per your request, so ask what the LUN ID is for each one of the LUNs that are created for you. Verify this information by looking at the output of the **multipath -ll** command, as shown in Example 5-3 on page 65, and check that the LUN IDs and size match what you are expecting.

After you have this information, populate your `multipath.conf` file.

> **Important:** Check that you correctly assign the log LUN ID to your log disk alias, especially if you are directly specifying an SSD or flash disk to be used as the HANA log area.

Scale-out clusters that use the HANA shared area from a highly available NFS server infrastructure do not see a locally attached HANA shared LUN, and do not need to define an alias for the HANA shared disk. Scale-out clusters that use IBM Spectrum Scale for the HANA shared area can still create an alias for the shared LUN on the cluster nodes.

The following naming convention is not mandatory, but is a best practice. Consider naming your HANA data and log LUNs according to the following scheme:

► `HANA_DATA_<node number>_<data disk number>`

► `HANA_LOG_<node number>_<log disk number>`

This naming scheme is especially useful in scale-out clusters where all data and log disks must be mapped to all nodes. In this way, you can easily identify that disk HANA_DATA32 is the second data disk from node 3.

The devices section

The `devices` section is used to override the settings in the `default` section, which is useful when more than one storage type is connected to the system. Using more than one storage type in a configuration is rare, but the devices section provides multivendor support if required.

Regardless of how many storage units you use, the best practice is to isolate the settings for them inside a `device { }` definition, as shown in Example 5-4 on page 66. Our configuration has a `device { }` entry for an `IBM` (vendor) `2810XIV` (product) storage unit type. We then have definitions for a multitude of parameters, such as `path_selector`, `rr_min_io`, and others. These settings deliver better performance for a multipath configuration by using IBM XIV.

Each storage vendor has its own recommendations for which parameters to use in this section and how to tune them for performance. Check their respective documentation for their best practices.

5.2.1 Applying changes to the multipath configuration

After you populate the `multipath.conf` file for your environment, you must apply the changes. To apply the changes, restart and reload operations against the multipathd service, as shown in Example 5-5.

Example 5-5   Restarting the multipath service

```
hanaonpower:/etc # service multipathd restart
hanaonpower:/etc # service multipathd reload
```

Now, check what happens to the output of the **multipath -ll** command, as shown in Example 5-6. We use a combination of the **multipath -ll** command with the **grep** command to output only the important information we want to validate: alias, WWID, and disk size.

Example 5-6   Checking the changes that are applied to the multipath configuration

```
hanaonpower:~ # multipath -ll | grep IBM -A 1
USRSAP (2001738002ae12c89) dm-5 IBM,2810XIV
size=48G features='2 queue_if_no_path retain_attached_hw_handler' hwhandler='1 alua' wp=rw
--
HANA_LOG_1_1 (2001738002ae12c8a) dm-6 IBM,2810XIV
size=35G features='2 queue_if_no_path retain_attached_hw_handler' hwhandler='1 alua' wp=rw
--
HANA_LOG_1_2 (2001738002ae12c90) dm-11 IBM,2810XIV
size=35G features='2 queue_if_no_path retain_attached_hw_handler' hwhandler='1 alua' wp=rw
--
HANA_LOG_1_3 (2001738002ae12c91) dm-12 IBM,2810XIV
size=35G features='2 queue_if_no_path retain_attached_hw_handler' hwhandler='1 alua' wp=rw
--
HANA_LOG_1_4 (2001738002ae12c92) dm-13 IBM,2810XIV
size=35G features='2 queue_if_no_path retain_attached_hw_handler' hwhandler='1 alua' wp=rw
--
ROOTVG (2001738002ae12c88) dm-0 IBM,2810XIV
size=64G features='2 queue_if_no_path retain_attached_hw_handler' hwhandler='1 alua' wp=rw
--
HANA_SHARED01 (2001738002ae12c8b) dm-4 IBM,2810XIV
size=144G features='2 queue_if_no_path retain_attached_hw_handler' hwhandler='1 alua' wp=rw
--
```

```
HANA_DATA_1_4 (2001738002ae12c8f) dm-10 IBM,2810XIV
size=112G features='2 queue_if_no_path retain_attached_hw_handler' hwhandler='1 alua' wp=rw
--
HANA_DATA_1_3 (2001738002ae12c8e) dm-8 IBM,2810XIV
size=112G features='2 queue_if_no_path retain_attached_hw_handler' hwhandler='1 alua' wp=rw
--
HANA_DATA_1_2 (2001738002ae12c8d) dm-9 IBM,2810XIV
size=112G features='2 queue_if_no_path retain_attached_hw_handler' hwhandler='1 alua' wp=rw
--
HANA_DATA_1_1 (2001738002ae12c8c) dm-7 IBM,2810XIV
size=112G features='2 queue_if_no_path retain_attached_hw_handler' hwhandler='1 alua' wp=rw
```

All disks are now referenced by their aliases and are mapped as symbolic links in the
/dev/mapper/ directory. Example 5-7 lists our lab environment system's disk aliases in
/dev/mapper.

Example 5-7 Listing the disk aliases under /dev/mapper

```
hanaonpower:~ # ls -la /dev/mapper
total 0
drwxr-xr-x  2 root root     340 Jun 22 18:50 .
drwxr-xr-x 15 root root    5660 Jun 22 18:50 ..
lrwxrwxrwx  1 root root       7 Jun 22 18:50 HANA_DATA_1_1 -> ../dm-7
lrwxrwxrwx  1 root root       7 Jun 22 18:50 HANA_DATA_1_2 -> ../dm-9
lrwxrwxrwx  1 root root       7 Jun 22 18:50 HANA_DATA_1_3 -> ../dm-8
lrwxrwxrwx  1 root root       8 Jun 22 18:50 HANA_DATA_1_4 -> ../dm-10
lrwxrwxrwx  1 root root       7 Jun 22 18:50 HANA_LOG_1_1 -> ../dm-6
lrwxrwxrwx  1 root root       7 Jun 22 18:50 HANA_LOG_1_2 -> ../dm-11
lrwxrwxrwx  1 root root       7 Jun 22 18:50 HANA_LOG_1_3 -> ../dm-12
lrwxrwxrwx  1 root root       7 Jun 22 18:50 HANA_LOG_1_4 -> ../dm-13
lrwxrwxrwx  1 root root       7 Jun 22 18:50 HANA_SHARED01 -> ../dm-4
lrwxrwxrwx  1 root root       7 Jun 22 18:50 ROOTVG -> ../dm-0
lrwxrwxrwx  1 root root       7 Jun 22 18:50 ROOTVG1 -> ../dm-1
lrwxrwxrwx  1 root root       7 Jun 22 18:50 ROOTVG2 -> ../dm-2
lrwxrwxrwx  1 root root       7 Jun 22 18:50 ROOTVG3 -> ../dm-3
lrwxrwxrwx  1 root root       7 Jun 22 18:50 ROOTVG_part1 -> ../dm-1
lrwxrwxrwx  1 root root       7 Jun 22 18:50 ROOTVG_part2 -> ../dm-2
lrwxrwxrwx  1 root root       7 Jun 22 18:50 ROOTVG_part3 -> ../dm-3
lrwxrwxrwx  1 root root       7 Jun 22 18:50 USRSAP -> ../dm-5
crw-------  1 root root 10, 236 Jun 22 18:16 control
```

Now that you have all your LUNs available and using aliases and the multipath daemon is set
up according to your storage vendor specifications, it is time to create the HANA file systems.

5.3 File system creation and setup

This section explains how to create the following file systems, along with the setup that each
of them requires:

- ▶ HANA data file system, mounted on /hana/data/<SID>
- ▶ HANA log file system, mounted on /hana/log/<SID>
- ▶ HANA shared file system, mounted on /hana/shared
- ▶ The /usr/sap file system

Each file system is created on its own disks and its own LVM elements. The `<SID>` is the SAP ID (SID) of the instance that you are installing. For example, if your SID is `HP0`, then the mount points are `/hana/data/HP0` and `/hana/log/HP0`.

Also, notice that the mount point setup for scale-out is different and is treated as described in 5.3.2, "File systems for scale-out systems" on page 76.

As the implementation for scale-up systems and scale-out clusters differs slightly, there is a separate section for each scenario.

All tunings that are described in this section come from the *IBM System Storage Architecture and Configuration Guide for SAP HANA Tailored Datacenter Integration*. The values that are used in this publication are current at the time of writing. When you implement SAP HANA on IBM Power Systems, review the white paper to check for current tuning recommendations and guidance.

> **Note:** If you work with Multiple Component One System (MCOS) implementations where more than one HANA instance co-exists in the same OS, then use `/hana/data<SID>` and `/hana/log/<SID>` as the mount points for your file systems, and name each one of the instance's volume groups (VGs) and logical volumes (LVs) uniquely to avoid confusion.

5.3.1 File systems for scale-up systems

A scale-up system has only local disks, so all HANA file systems are created on these local disks. Also, we use the cosmetic disk aliasing that is described in 5.2, "Linux multipath setup" on page 66.

HANA data file system

To create a HANA data file system, complete the following steps:

1. Create the HANA data file system on the HANA data disks. Recall that in the alias configuration, the disks are named `/dev/mapper/HANA_DATA*`. So, use this wildcard to ease the creation of the LVM VG and LV for the HANA data area.

   Example 5-8 shows how to create the HANA data VG. As a best practice, call the VG `hanadata`. You can copy and paste the steps from all of our further examples if you choose do to so. Also, notice that we use some tuning parameters, such as:

   — `-s 1M`: Creates the VG with an extent size of 1 MB.

   — `--dataalignment 1M`: Aligns the start of the data to a multiple of 1 MB.

Example 5-8   Creating the HANA data logical volume manager volume group

```
hanaonpower:~ # vgcreate -s 1M --dataalignment 1M hanadata
/dev/mapper/HANA_DATA*
  Physical volume "/dev/mapper/HANA_DATA01" successfully created
  Physical volume "/dev/mapper/HANA_DATA02" successfully created
  Physical volume "/dev/mapper/HANA_DATA03" successfully created
  Physical volume "/dev/mapper/HANA_DATA04" successfully created
  Volume group "hanadata" successfully created
```

2. Create an LV inside that newly created VG. Name this LV as `datalv`. Example 5-9 shows this step. Notice that we use some tuning parameters, such as:

  – `-i 4`: Creates the LV by using four stripes across the VG. Because we use four data disks, I/O operations on this LV are spread to all four disks, thus maximizing read and write performance.

  – `-I 256K`: Uses a 256 KB stripe size. Recall from Example 5-8 on page 71 where we create the VG with an extent size and data alignment of 1 MB. So, having four stripes of 256 KB matches the extent size nicely.

Example 5-9   Creating the HANA data logical volume manager volume group

```
hanaonpower:~ # lvcreate -i 4 -I 256K -l 100%FREE -n datalv hanadata
  Logical volume "datalv" created.
```

3. Create the file system on the newly created LV. The SAP supported file system that we use in our examples is Extents File System (XFS). Example 5-10 shows this step. Notice that we use some tuning parameters, such as:

  – `-b size=4096`: Sets the block size to 4096 bytes.

  – `-s size=4096`: Sets the sector size to 4096 bytes.

Example 5-10   Creating the HANA data file system

```
hanaonpower:~ # mkfs.xfs -b size=4096 -s size=4096 /dev/mapper/hanadata-datalv
meta-data=/dev/mapper/hanadata-datalv isize=256    agcount=16, agsize=7339968 blks
         =                            sectsz=4096  attr=2, projid32bit=1
         =                            crc=0        finobt=0, sparse=0
data     =                            bsize=4096   blocks=117439488, imaxpct=25
         =                            sunit=64     swidth=256 blks
naming   =version 2                   bsize=4096   ascii-ci=0 ftype=1
log      =internal log                bsize=4096   blocks=57343, version=2
         =                            sectsz=4096  sunit=1 blks, lazy-count=1
realtime =none                        extsz=4096   blocks=0, rtextents=0
```

HANA log file system

The creation and setup for the HANA log file system in a scale-up configuration is similar to the HANA data file system. Complete the following steps:

1. Example 5-11 shows the creation of the HANA log VG. As a best practice, name the VG `hanalog`.

Example 5-11   Creating the HANA log volume group

```
hanaonpower:~ # vgcreate -s 1M --dataalignment 1M hanalog /dev/mapper/HANA_LOG*
  Physical volume "/dev/mapper/HANA_LOG01" successfully created
  Volume group "hanalog" successfully created
```

2. Create the LV. The tuning parameters are the same ones that are used for the data file system, as shown in Example 5-12.

Example 5-12   Creating the HANA log logical volume

```
hanaonpower:~ # lvcreate -i 4 -I 256K -l 100%FREE -n loglv hanalog
  Logical volume "loglv" created.
```

3. Create the XFS file system on the newly created LV, as shown in Example 5-13. The HANA log file system has the same settings that are used for the HANA data area.

Example 5-13   Creating the HANA log file system

```
hanaonpower:~ # mkfs.xfs -b size=4096 -s size=4096 /dev/mapper/hanalog-loglv
meta-data=/dev/mapper/hanalog-loglv isize=256    agcount=4, agsize=9437120 blks
         =                          sectsz=4096  attr=2, projid32bit=1
         =                          crc=0        finobt=0, sparse=0
data     =                          bsize=4096   blocks=37748480, imaxpct=25
         =                          sunit=0      swidth=0 blks
naming   =version 2                 bsize=4096   ascii-ci=0 ftype=1
log      =internal log             bsize=4096   blocks=18431, version=2
         =                          sectsz=4096  sunit=1 blks, lazy-count=1
realtime =none                      extsz=4096   blocks=0, rtextents=0
```

HANA shared file system

The creation and setup of the HANA shared file system in a scale-up configuration is similar to the HANA log file system. Complete the following steps:

1. Example 5-14 shows the creation of the HANA shared VG. As a best practice, name the VG hanashared.

Example 5-14   Creating the HANA shared volume group

```
hanaonpower:~ # vgcreate hanashared /dev/mapper/HANA_SHARED*
  Physical volume "/dev/mapper/HANA_SHARED01" successfully created
  Volume group "hanashared" successfully created
```

2. When you create the LV, there is no need to apply any striping parameters because the shared area is mapped onto one disk only. Example 5-15 shows the creation of the LV.

Example 5-15   Creating the HANA shared logical volume

```
hanaonpower:~ # lvcreate -l 100%FREE -n sharedlv hanashared
  Logical volume "sharedlv" created.
```

3. Create the XFS file system on the newly created LV. For the HANA shared area, there is no need to apply any file system tuning flags, as shown in Example 5-16.

Example 5-16   Creating the HANA shared file system

```
hanaonpower:~ # mkfs.xfs /dev/mapper/hanashared-sharedlv
meta-data=/dev/mapper/hanashared-sharedlv isize=256    agcount=4,
agsize=9437120 blks
         =                          sectsz=512   attr=2, projid32bit=1
         =                          crc=0        finobt=0, sparse=0
data     =                          bsize=4096   blocks=37748480, imaxpct=25
         =                          sunit=0      swidth=0 blks
naming   =version 2                 bsize=4096   ascii-ci=0 ftype=1
log      =internal log             bsize=4096   blocks=18431, version=2
         =                          sectsz=512   sunit=0 blks, lazy-count=1
realtime =none                      extsz=4096   blocks=0, rtextents=0
```

The /usr/sap file system

Finally, you must create the /usr/sap file system. We follow the same procedure that was used for creating VGs, LVs, and the file systems, except that there is no tuning needed at all for /usr/sap. Complete the following steps:

1. Example 5-17 illustrates the creation of the VG. As a best practice, name this VG usrsap.

Example 5-17   Creating the /usr/sap volume group

```
hanaonpower:~ # vgcreate usrsap /dev/mapper/USRSAP
  Physical volume "/dev/mapper/USRSAP" successfully created
  Volume group "usrsap" successfully created
```

2. Create the LV for /usr/sap, as shown in Example 5-18.

Example 5-18   Creating the /usr/sap logical volume

```
hanaonpower:~ # lvcreate -l 100%FREE -n saplv usrsap
  Logical volume "saplv" created.
```

3. Create the /usr/sap file system, as shown in Example 5-19.

Example 5-19   Creating the /usr/sap file system

```
hanaonpower:~ # mkfs.xfs /dev/mapper/usrsap-saplv
meta-data=/dev/mapper/usrsap-saplv isize=256    agcount=4, agsize=3150848 blks
         =                       sectsz=512   attr=2, projid32bit=1
         =                       crc=0        finobt=0, sparse=0
data     =                       bsize=4096   blocks=12603392, imaxpct=25
         =                       sunit=0      swidth=0 blks
naming   =version 2              bsize=4096   ascii-ci=0 ftype=1
log      =internal log           bsize=4096   blocks=6154, version=2
         =                       sectsz=512   sunit=0 blks, lazy-count=1
realtime =none                   extsz=4096   blocks=0, rtextents=0
```

Creating the mount points and configuring /etc/fstab

Now that you are done creating the HANA file systems, you must create the mount points and create the entries in the /etc/fstab file in order for the file systems to be mounted by the system.

The HANA mount points are standardized by SAP, and must follow the guidelines that are stated in 5.3, "File system creation and setup" on page 70. Complete the following steps:

1. Example 5-20 depicts the creation of those mount points according to the proper SAP nomenclature.

Example 5-20   Creating the HANA file systems mount points

```
hanaonpower:~ # mkdir -p /hana/data/<SID> /hana/log/<SID> /hana/shared /usr/sap
hanaonpower:~ #
```

2. Append a section to the `/etc/fstab` file with the definitions of the HANA file systems. Example 5-21contains the entries that you must append, in **bold**. *Do not* change any existing entries and remember that your UUIDs are different from the ones in Example 5-21.

Example 5-21   Creating entries in /etc/fstab for the HANA file systems

```
hanaonpower:~ # cat /etc/fstab
UUID=58e3f523-f065-4bea-beb5-5ac44313ad30 swap swap defaults 0 0
UUID=72f1be2c-022d-49f8-a6e6-a9c02122375a / btrfs defaults 0 0
UUID=72f1be2c-022d-49f8-a6e6-a9c02122375a /home btrfs subvol=@/home 0 0
UUID=72f1be2c-022d-49f8-a6e6-a9c02122375a /opt btrfs subvol=@/opt 0 0
UUID=72f1be2c-022d-49f8-a6e6-a9c02122375a /srv btrfs subvol=@/srv 0 0
UUID=72f1be2c-022d-49f8-a6e6-a9c02122375a /usr/local btrfs subvol=@/usr/local 0 0
UUID=72f1be2c-022d-49f8-a6e6-a9c02122375a /var/log btrfs subvol=@/var/log 0 0
UUID=72f1be2c-022d-49f8-a6e6-a9c02122375a /var/opt btrfs subvol=@/var/opt 0 0
UUID=72f1be2c-022d-49f8-a6e6-a9c02122375a /var/spool btrfs subvol=@/var/spool 0 0
UUID=72f1be2c-022d-49f8-a6e6-a9c02122375a /.snapshots btrfs subvol=@/.snapshots 0 0

#hana
/dev/mapper/hanadata-datalv     /hana/data/<SID> xfs    defaults   0   0
/dev/mapper/hanalog-loglv       /hana/log/<SID> xfs     defaults   0   0
/dev/mapper/hanashared-sharedlv /hana/shared    xfs     defaults   0   0
/dev/mapper/usrsap-saplv        /usr/sap        xfs     defaults   0   0
```

3. Finally, as depicted in Example 5-22, mount all the HANA file systems and use **df -h** to check that they are all mounted. Take the time to check the file system sizes as well.

Example 5-22   Mounting the HANA file systems

```
hanaonpower:~ # mount -a
hanaonpower:~ # df -h
Filesystem                      Size  Used Avail Use% Mounted on
devtmpfs                        3.5G     0  3.5G   0% /dev
tmpfs                           8.0G  128K  8.0G   1% /dev/shm
tmpfs                           3.5G   30M  3.5G   1% /run
tmpfs                           3.5G     0  3.5G   0% /sys/fs/cgroup
/dev/mapper/ROOTVG3              62G   11G   50G  19% /
/dev/mapper/ROOTVG3              62G   11G   50G  19% /var/opt
/dev/mapper/ROOTVG3              62G   11G   50G  19% /var/log
/dev/mapper/ROOTVG3              62G   11G   50G  19% /.snapshots
/dev/mapper/ROOTVG3              62G   11G   50G  19% /srv
/dev/mapper/ROOTVG3              62G   11G   50G  19% /opt
/dev/mapper/ROOTVG3              62G   11G   50G  19% /home
/dev/mapper/ROOTVG3              62G   11G   50G  19% /usr/local
/dev/mapper/ROOTVG3              62G   11G   50G  19% /var/spool
tmpfs                           708M     0  708M   0% /run/user/0
/dev/mapper/hanadata-datalv     448G   33M  448G   1% /hana/data
/dev/mapper/hanalog-loglv       144G   33M  144G   1% /hana/log
/dev/mapper/hanashared-sharedlv 144G   33M  144G   1% /hana/shared
/dev/mapper/usrsap-saplv         50G   33M   50G   1% /usr/sap
```

5.3.2 File systems for scale-out systems

This paper covers the scenarios that are commonly used by customers in the field. The next sections explain and point to existing reference documentation when appropriate.

> **Note:** The `/usr/sap` file system is still local to each node, so follow the procedures that are described in "The /usr/sap file system" on page 74 to create it on each scale-out cluster node.

Regardless of which implementation is used, your nodes must pass the SAP HANA Hardware Configuration Check Tool (HWCCT) File System I/O benchmark tests.

IBM Elastic Storage Server and IBM Spectrum Scale based environments

The IBM Elastic Storage Server solution is fully compliant with SAP requirements and makes it simple to manage file systems for HANA scale-out clusters. Using an Elastic Storage Server and IBM Spectrum Scale is supported for all three HANA file systems: data, log, and shared.

In this implementation, all three file systems are created in your Elastic Storage Server or IBM Spectrum Scale infrastructure (data, log, and shared), and all scale-out nodes are connected to them. For more information, see *SAP HANA and ESS: A Winning Combination*, REDP-5436.

After reviewing that paper, confirm that all of your nodes can see the bolded file systems, as shown in Example 5-23. Our example is from a four-node scale-out cluster, so the file systems are sized.

Example 5-23   HANA scale-out cluster file systems that use Elastic Storage Server

```
saphana005:~ # mount
[ ... snip ...]
hanadata on /hana/data type gpfs (rw,relatime)
hanalog on /hana/log type gpfs (rw,relatime)
hanashared on /hana/shared type gpfs (rw,relatime)
saphana005:~ #
saphana005:~ # df -h
Filesystem              Size  Used Avail Use% Mounted on
[... snip ...]
hanadata                1.0T  257M  1.0T   1% /hana/data
hanalog                 512G  257M  512G   1% /hana/log
hanashared              1.0T  257M  1.0T   1% /hana/shared
```

HANA shared file system by using NFS

The NFS is another alternative to use with HANA scale-out clusters file systems. It is fully supported by SAP. The settings are governed by SAP Note 2099253. In addition to that note, we add some best practice configurations from our experience in the field. Using NFS is supported by all three HANA file systems: data, log, and shared.

Basically, you must ensure that your NFS server infrastructure has HA built in, as described in "HANA shared area that is managed with Network File System" on page 63. Then, you must create an export entry on the NFS server to be shared to the HANA clients.

In this book, we illustrate a single, overall export entry on the NFS server to share the entire /hana folder instantly instead of having an entry for data (/hana/data), log (/hana/log), and shared (/hana/shared). If you want to use NFS only for the shared area, which is a more commonly used scenario, then adjust your settings. Even though we use a single export point from the NFS server, each area (data, log, and shared) has its own file system on the NFS server and uses a compliant file system type according to SAP Note 2055470, for example, XFS. Apply all the tunings that are pertinent to each file system, as described in 5.3.1, "File systems for scale-up systems" on page 71.

The NFS server export entry in the /etc/exports file for the /hana directory uses the parameters that are outlined in Example 5-24.

Example 5-24   HANA export on the NFS server

```
/hana node1(fsid=0, crossmnt, rw, no_root_squash, sync, no_subtree_check)
node2(fsid=0, crossmnt, rw, no_root_squash, sync, no_subtree_check) node3(fsid=0,
crossmnt, rw, no_root_squash, sync, no_subtree_check) node4(fsid=0, crossmnt, rw,
no_root_squash, sync, no_subtree_check)
```

The export entry is granted permission to be mounted only by the participating scale-out nodes. This is a best practice that is followed to ensure that other systems do not have any access to the HANA file systems of the scale-out cluster. Example 5-24 shows the /hana exported to four hosts: node1, node2, node3, and node4. This is a single, long line that contains all hosts for which the entry is exported, with each description node in the following format:

```
<node_hostname>(fsid=0, crossmnt, rw, no_root_squash, sync, no_subtree_check)
```

After completing the required configuration on your NFS server with that export entry, each of the scale-out nodes must mount it by using the correct parameters to ensure optimal performance.

Create a /hana mount point on your scale-out cluster nodes, and add the following line to the /etc/fstab file, where <nfsserver> is the resolvable host name of the NFS server:

```
<nfsserver>:/hana   /hana   nfs   rw,soft,intr,rsize=8192,wsize=8192   0  0
```

Run the **mount /hana** command on all nodes and test whether they can all access that NFS share with read/write permissions.

To obtain the current NFS best practices from SAP, see *SAP HANA Storage Requirements*, especially if you are planning to use NFS for HANA data and log file systems.

HANA shared file system that uses IBM Spectrum Scale

IBM Spectrum Scale is an alternative to sharing the HANA shared file system among the scale-out nodes. IBM Spectrum Scale is also supported for sharing the data and log areas, as described in "IBM Elastic Storage Server and IBM Spectrum Scale based environments" on page 76.

Installing IBM Spectrum Scale and creating a file system is out of the scope of this publication. For more information about implementing IBM Spectrum Scale, see *Implementing IBM Spectrum Scale*, REDP-5254. We provide guidance about how to design your IBM Spectrum Scale cluster to ensure HA. No special tuning is required for the file system.

If you decide to use IBM Spectrum Scale for the `/hana/shared` file system, here are some guidelines that you want to follow:

► Use an odd number of quorum nodes.

► If your scale-out cluster is composed of only two nodes, use IBM Spectrum Scale tiebreaker disks, or add a third IBM Spectrum Scale only node for quorum.

► You can choose to create the file system over a single LUN or multiple ones. As a best practice, use a number of LUNs equal to the number of IBM Spectrum Scale nodes for better performance.

► Share the LUNs among the scale-out nodes by using the direct-attached methodology, where all nodes have access to them over local Fibre Channel adapters, whether physical or virtual (N_Port ID Virtualization (NPIV)) adapters.

Storage Connector API for the data and log areas

If you do not plan to work with truly shared file systems for the data and log areas, your scale-out HANA cluster can use the Storage Connector API to have a standby node take over the file systems from a failed node and mount it. For more information, see *SAP HANA Fibre Channel Storage Connector Admin Guide*.

You use the Storage Connector API to create as many individual data and log file systems equal to the master and worker nodes in the scale-out solution. Figure 5-1 illustrates such scenario by using a four-node cluster that is composed of one master node, two worker nodes, and one standby node. The standby node takes over if one of the other nodes fail.
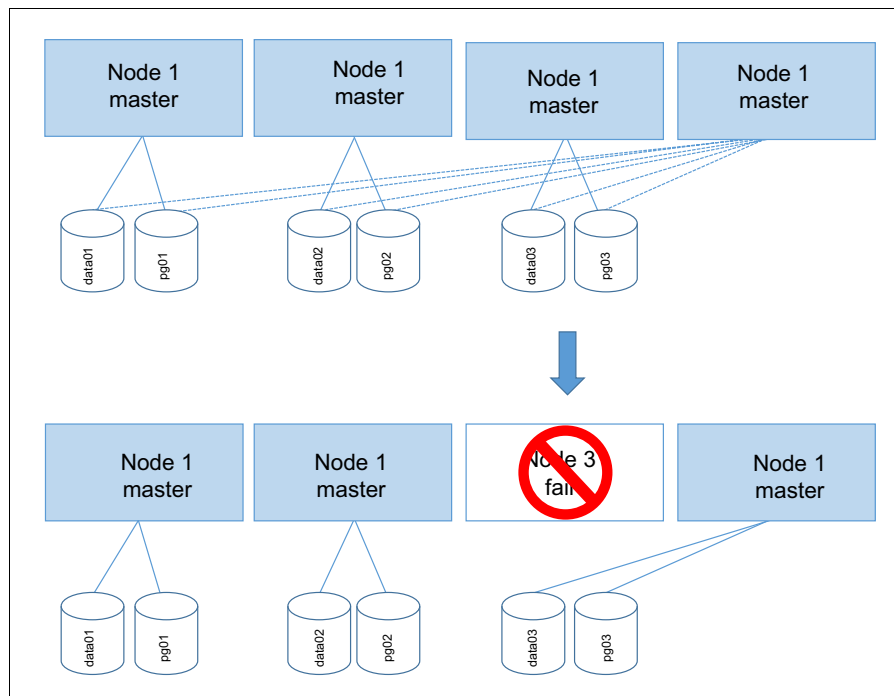

Figure 5-1   Scale-out cluster: Data and log file systems that are under the Storage Connector API

To get started, create each of the data and log file systems individually, each in its own VG and LV, according to the explanations in "HANA data file system" on page 71 and "HANA log file system" on page 72. Recall from "The multipaths section" on page 68 that we recommend naming the LUNs after the standard `HANA_DATA_<node number>_<data disk number>` and `HANA_LOG_<node number>_<log disk number>`. So, in a four-node cluster with one master node, two worker nodes, and one standby node, we have the following (in the format *<VG name>-<LV name>: participating disks*):

▶ Master (represented as node 1):

 – `hanadata01-datalv01: /dev/mapper/HANA_DATA_1_*`

 – `hanalog01-loglv01: /dev/mapper/HANA_LOG_1_*`

▶ Worker (represented as node 2):

 – `hanadata02-datalv02: /dev/mapper/HANA_DATA_2_*`

 – `hanalog02-loglv02: /dev/mapper/HANA_LOG_2_*`

▶ Worker (represented as node 3):

 – `hanadata03-datalv03: /dev/mapper/HANA_DATA_3_*`

 – `hanalog03-loglv03: /dev/mapper/HANA_LOG_3_*`

▶ Standby (represented as node 4): Does not have associated data or log file systems. but takes over any of the other three nodes file systems if any of these nodes fail.

> **Important:** All of the data and log LUNs are attached to all of the scale-out cluster nodes, so any of the nodes can access and mount the file systems.

You can create all the file systems on just one node because all of them have access to all the LUNs. You do not need to include the information of the file systems in the `/etc/fstab` file because HANA in a scale-out cluster handles the mounting of them automatically when you use the storage connector API. When you are done creating the file systems, run a **vgscan** command on all nodes and check that they can all see the VGs that you created.

Now, from a disks and file systems perspective, everything is set up to trigger a HANA scale-out installation.

5.4 More Linux I/O subsystem tuning

In addition to *IBM System Storage Architecture and Configuration Guide for SAP HANA Tailored Datacenter Integration*, there is *Linux I/O Performance Tuning for IBM System Storage, V1.5*. This guide provides more tuning for the I/O stack of the Linux OS on IBM Power Systems servers. The information that is provided in the next section was correct at the time of writing. Check for updates to the papers for current details.

5.4.1 I/O device tuning

These changes to the I/O device tuning are applied only if you encounter performance issues with your disks.

The first change that you can make to improve performance is to add the **rr_min_io_rq** parameter to your `/etc/multipath.conf` **"device { }"** and set it to 1.

The second change is to increase the queue depth of your disk devices. Contact your storage vendor about which increment to use and how to make it persistent across restarts. To change the queue depth dynamically without making the changes permanent, run the following command on all your disk devices, and replace *<NN>* with the chosen queue depth value:

```
echo <NN> > cat /sys/bus/scsi/devices/<device>/queue_depth
```

5.4.2 I/O scheduler tuning

The preferred I/O scheduler for HANA disks is NOOP. You can change it in the boot loader configuration by using the **elevator** boot parameter, in which case the changes are permanent and applied in the next system restart. Example 5-25 shows how to edit your /etc/default/grub file to make this change. You append elevator=noop to the line that starts with GRUB_CMDLINE_LINUX_DEFAULT, as shown in **bold**.

Example 5-25   Choosing the NOOP I/O scheduler at start time with the elevator parameter

```
hanaonpower:~ # cat /etc/default/grub
# If you change this file, run 'grub2-mkconfig -o /boot/grub2/grub.cfg' afterward
to update /boot/grub2/grub.cfg.

# Uncomment to set your own custom distributor. If you leave it unset or empty,
the default
# policy is to determine the value from /etc/os-release
GRUB_DISTRIBUTOR=
GRUB_DEFAULT=saved
GRUB_HIDDEN_TIMEOUT=0
GRUB_HIDDEN_TIMEOUT_QUIET=true
GRUB_TIMEOUT=8
GRUB_CMDLINE_LINUX_DEFAULT="splash=silent quiet showopts elevator=noop"
GRUB_CMDLINE_LINUX=""

[... cropped ...]
```

After making the changes to /etc/default/grub, run **grub2-mkconfig** to apply the changes to the grub2 boot loader, as shown in Example 5-26.

Example 5-26   Applying the changes to the /etc/default/grub file

```
hanaonpower:~ # grub2-mkconfig -o /boot/grub2/grub.cfg
Generating grub configuration file ...
Found linux image: /boot/vmlinux-4.4.21-69-default
Found initrd image: /boot/initrd-4.4.21-69-default
done
```

Afterward, if you also want to change the scheduler algorithm dynamically without restarting your system, you can do so by using the /sys interface. Every device mapper disk that represents your LUNs has an I/O scheduler interface that is accessed at /sys/block/<dm-X>/queue/scheduler. Example 5-27 shows how to change the I/O scheduler of one of our device mapper disks from cfq to noop.

Example 5-27   Changing the disk I/O scheduler dynamically

```
To check current (default is cfq):

for BLOCKDEV in `ls -1 /sys/block/| grep sd`; do echo "Device: $BLOCKDEV"; cat
/sys/block/$BLOCKDEV/queue/scheduler; done

To set to NOOP on this boot:

for BLOCKDEV in `ls -1 /sys/block/| grep sd`; do echo "Device: $BLOCKDEV"; echo
noop > /sys/block/$BLOCKDEV/queue/scheduler; done
```

You must do the same task for each disk that you have in your system by using their dm-*X* form. To get a list of such disks, use the commands from Example 5-7 on page 70.

Your environment is ready for HANA to be installed on it.

6

SAP HANA software stack installation for a scale-up scenario

This chapter provides the instructions about how to use the SAP HANA on IBM Power Systems installer for a scale-up installation. This chapter covers both the GUI and the text interface installations.

Additionally, this chapter provides a quick guide about how to use SAP HANA Studio to connect to the HANA instance to manage it.

This chapter covers the following topics:

- ► SAP HANA installation overview
- ► Installation methods
- ► Postinstallation notes

6.1 SAP HANA installation overview

The SAP HANA installer offers a good degree of flexibility in terms of installation interfaces and installation options.

The available installation interfaces are:

► A GUI interface
► A browser-based interface
► A text interface

This chapter demonstrates how to perform the installation by using both the GUI and the text interfaces. The results are the same for either approach.

The GUI installation requires the installation of the X11 packages and a Virtual Network Computing (VNC) server to which to connect. If you followed the operating system (OS) installation guidelines that are described in SAP Note 2235581 (select the OS Configuration guide for your OS), you already have an X11 capable with VNC-enabled environment, and all you need to do is to connect to it by using a VNC client.

If you prefer to perform the installation in text mode, all you need is an SSH connection to your system. It is simple because this approach uses a text-based preinstallation wizard to create a response file and then uses it to drive the installation in unattended mode.

From an installation options point of view, you use HANA to select which components to install. These components are the server, the client, the Application Function Library (AFL) component, the Smart Data Access component, and so on. In this publication, we install only the server and the client components because we are interested in providing you with the guidelines for installing a highly available HANA environment.

> **Note:** As documented in SAP Note 2423367, starting with HANA 2.0 SPS1, all databases (DBs) are configured in multitenant mode only. There is no option to install a single-container instance.

Before you start the installation, you need the following information as input for the installer:

► The SAP ID (SID) of the instance that you plan to install.

► The instance number of the instance that you plan to install.

► The passwords that you plan to use for the `<sid>adm` user, the `SYSTEM` user, and the SAP Host Agent user (`sapadm`).

The following sections provide information about how the instance number works. This information is used later to access the instance through HANA Studio, and is also used to assign the port number that is used for communicating with HANA. The port number has the form of *3<instance number>15*. If you use an instance number of `00`, then the port number that HANA listens to is `30015`.

Also, if you plan to use SAP HANA System Replication (HSR) between two HANA instances, the replication itself uses the next instance number to assign the port number for replication services. For example, if the instances you want to replicate have a SID / instance number of `RB1 / 00`, then connecting to those instances happens over port 30015 and the replication between them over port 30115.

6.2 Installation methods

This section guides you through the installation process of HANA.

> **Note:** Throughout the following chapters, we show the installation of HANA by using SUSE Linux. The same steps apply for Red Hat Enterprise Linux. We point out differences between HANA for the SUSE and Red Hat Enterprise Linux installations where applicable.

If you have not yet done so, download the installer files for HANA, as described in 2.2.3, "Getting the SAP HANA on IBM Power Systems installation files" on page 17. At the time of writing, the HANA 2.0 on Power Systems installer is composed of four compressed files as follows (the first one has a `.exe` extension, and the others a `.rar` extension that is appended to them):

- ▶ 51052480_part1.exe
- ▶ 51052480_part2.rar
- ▶ 51052480_part3.rar
- ▶ 51052480_part4.rar

In SUSE, you can decompress those files by running the **unrar** command. As a best practice, place these installation files in a directory inside `/tmp` to avoid having issues with file permissions during the installation. Example 6-1 shows the command to decompress the files, along with some of the expected output.

Example 6-1 Decompressing the HANA installation files

```
hana002:/tmp/hana20_sps03 # unrar x 51052480_part1.exe

UNRAR 5.01 freeware     Copyright (c) 1993-2013 Alexander Roshal

Extracting from 51052480_part1.exe

Creating    51052480                                              OK
Creating    51052480/DATA_UNITS                                   OK
Creating    51052480/DATA_UNITS/HDB_CLIENT_LINUX_S390X_64         OK
Extracting  51052480/DATA_UNITS/HDB_CLIENT_LINUX_S390X_64/hdbsetup  OK
[…]
Extracting  51052480/COPY_TM.TXT                                  OK
Extracting  51052480/COPY_TM.HTM                                  OK
Extracting  51052480/MD5FILE.DAT                                  OK
Extracting  51052480/SHAFILE.DAT                                  OK
All OK
```

When you complete the decompression of the files, you see that the contents are extracted into a folder that is named after an SAP product number, and that it contains a data structure similar to Example 6-2. The HANA installer scripts and components are inside the `DATA_UNITS` folder.

Example 6-2 SAP HANA installation media directory structure

```
hana002:/tmp/hana20_sps03 # ls
51052480 51052480_part1.exe 51052480_part2.rar 51052480_part3.rar
51052480_part4.rar
hana002:/tmp/hana20_sps03 # cd 51052480/
hana002:/tmp/hana20_sps03/51052480 # ls -a
.          CDLABEL.EBC  DATA_UNITS  LABELIDX.ASC  SHAFILE.DAT
```

```
..              COPY_TM.HTM  LABEL.ASC   MD5FILE.DAT   VERSION.ASC
CDLABEL.ASC  COPY_TM.TXT  LABEL.EBC   MID.XML       VERSION.EBC
```

If you choose to perform an installation by using a GUI, see 6.2.1, "GUI installation" on page 86. If you plan to perform a text-mode installation, see 6.2.2, "Text-mode installation" on page 97.

6.2.1 GUI installation

To start a GUI installation, complete the following steps:

1. Use a VNC client to connect to the HANA system. The VNC server is available on display: 1 of your system. Access it by using its IP address and display: 1, as shown in Figure 6-1.
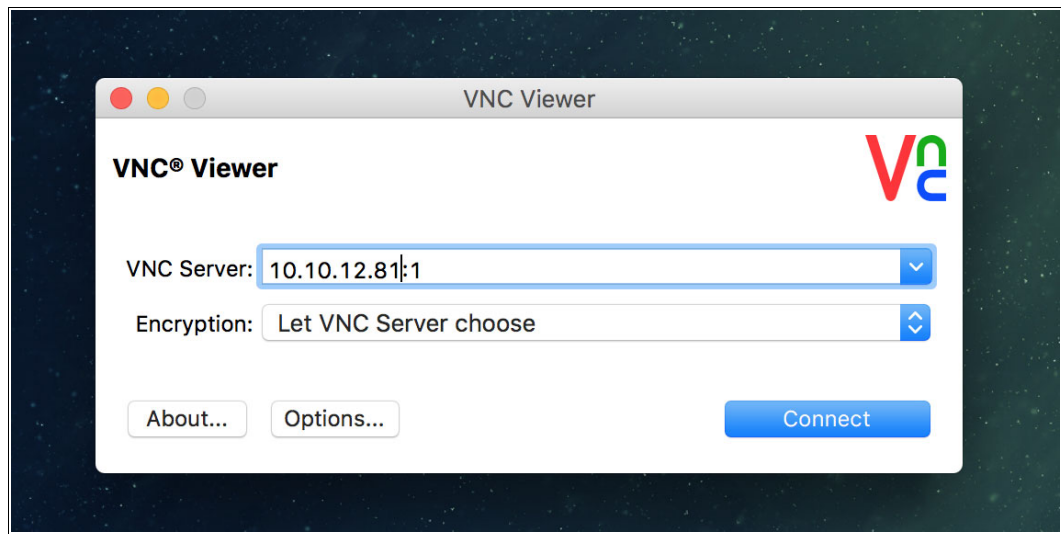
Figure 6-1   Connecting to SUSE by using a Virtual Network Computing client

2. After connecting to the system, log in by using the root user and password.

3. After you are logged in, open a terminal and go to the `DATA_UNITS/HDB_LCM_LINUX_PPC64LE` directory where you decompressed the HANA installer files. Then, run the **hdblcmgui** command to begin the installation.

4. After the GUI starts, provide the inputs that are required by the GUI installer as it progresses. The first window shows a list of the available components, as shown in Figure 6-2. Click **Next** and proceed with the installation.
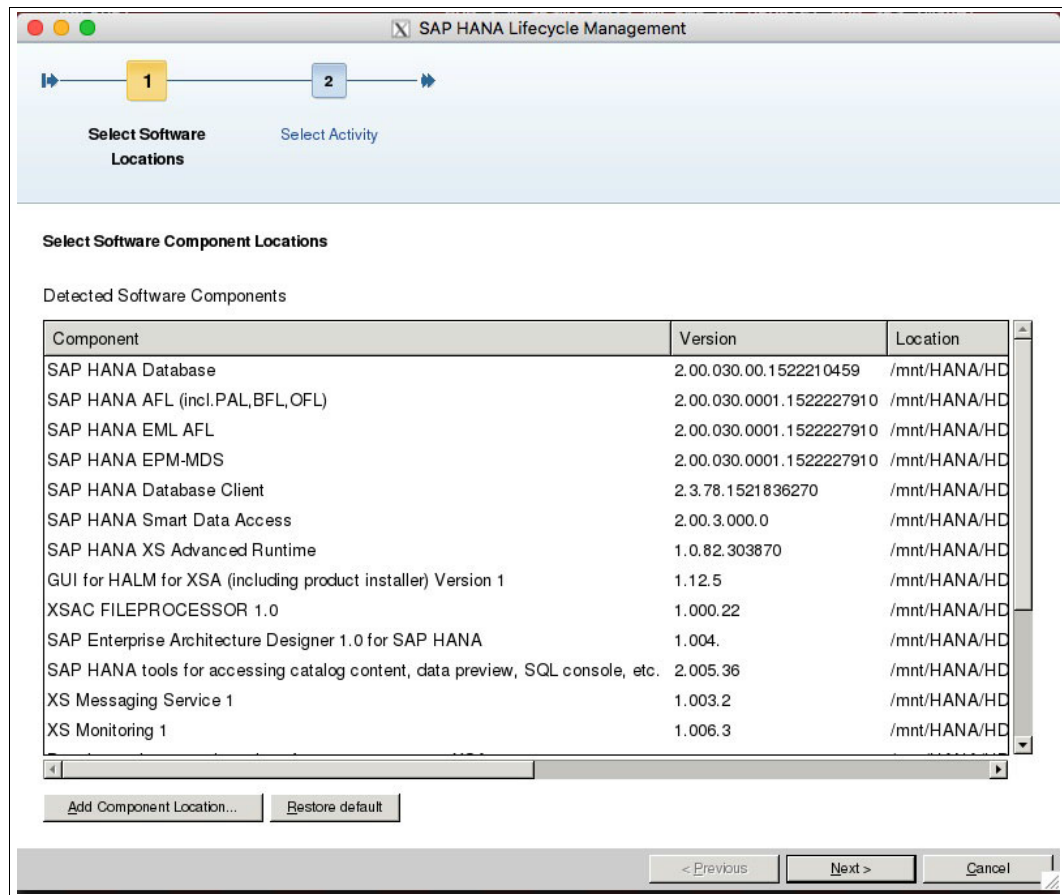

Figure 6-2   SAP HANA available components

5. The next step prompts an update of an existing system. That option is disabled because you have no HANA environments that are installed yet or are installing a new system. Select **Install New System** and click **Next** to proceed, as shown in Figure 6-3.
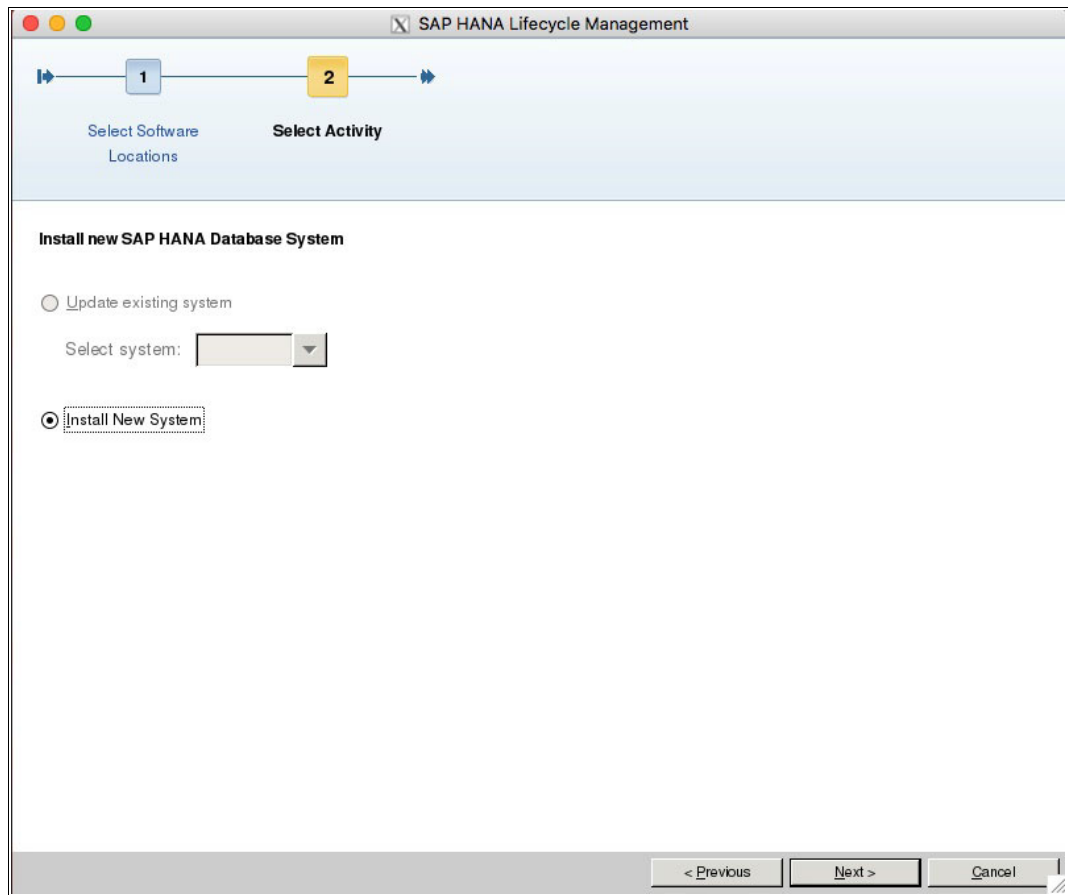

Figure 6-3   Installing a new HANA system

6. The next window prompts you for the components that you want to install. In our lab environment, we installed only the server and client options, as shown in Figure 6-4. The client is not required for a HANA installation because it is optional, so you can leave it out if you do not plan to use it. Click **Next** to proceed.
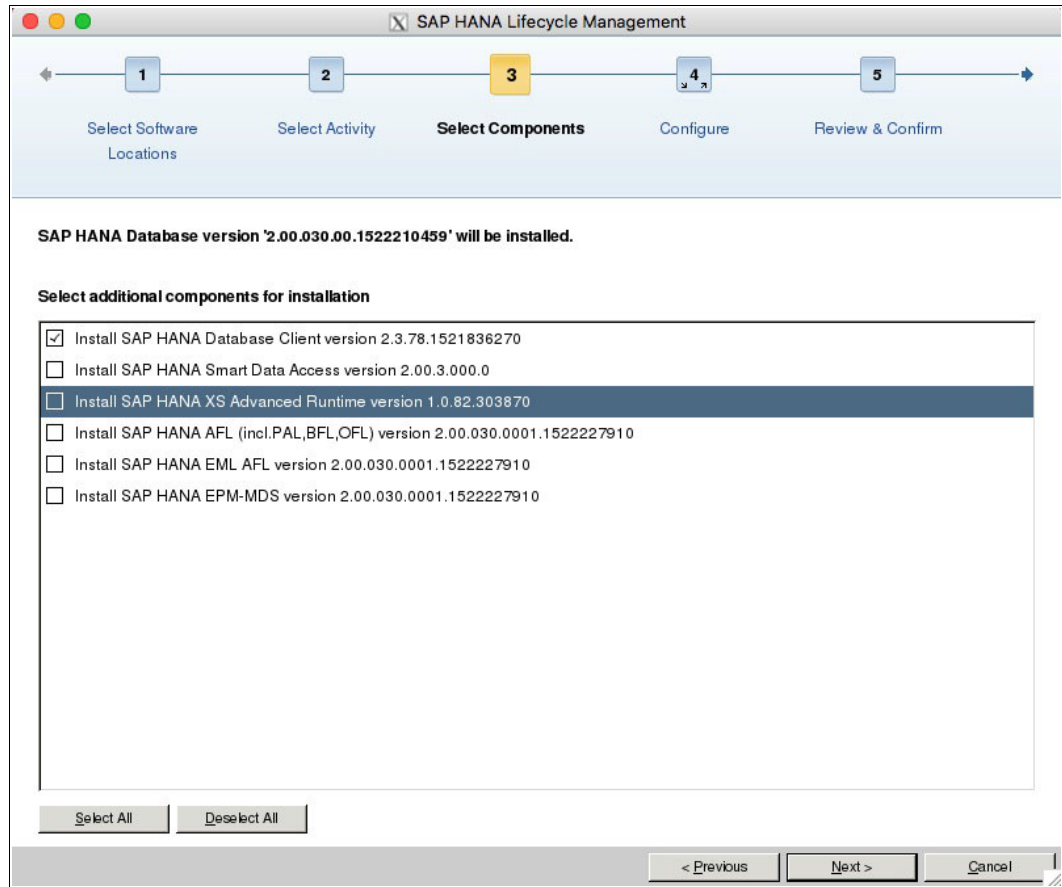

Figure 6-4   Selecting the components to install

7. Because this is a scale-up system installation, select **Single-Host System**, as shown in Figure 6-5. Click **Next**.
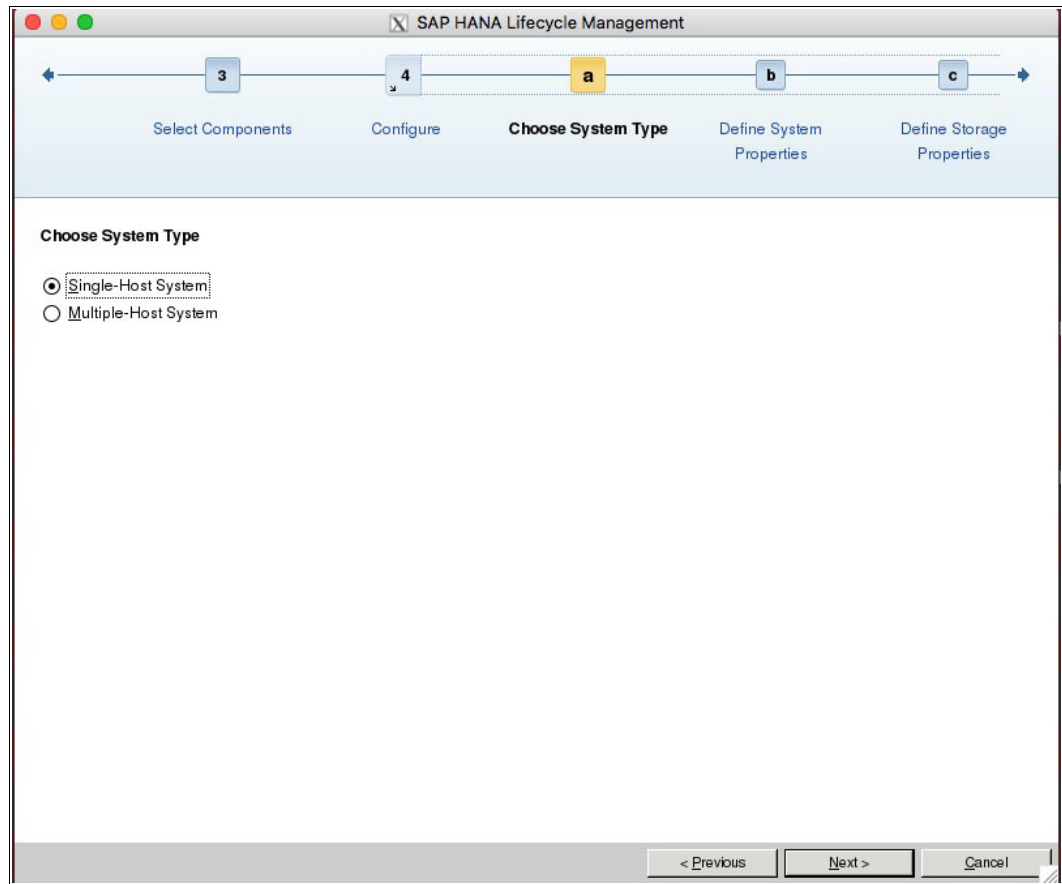

Figure 6-5   Single-host (scale-up) system installation

8. Input the information for the SID of your system. Most of the other parameters already are predefined, such as the host name and the installation path. As a best practice, keep the default values except for the SID and instance number, which you input according to your planning in 6.1, "SAP HANA installation overview" on page 84. You can also change the **System Usage** parameter according to what the system is used for. In our example (Figure 6-6), we selected **Custom - System usage is neither production, test nor development**. Your choice of System Usage relaxes some of the landscape test checks against your environment. After completion, click **Next**.
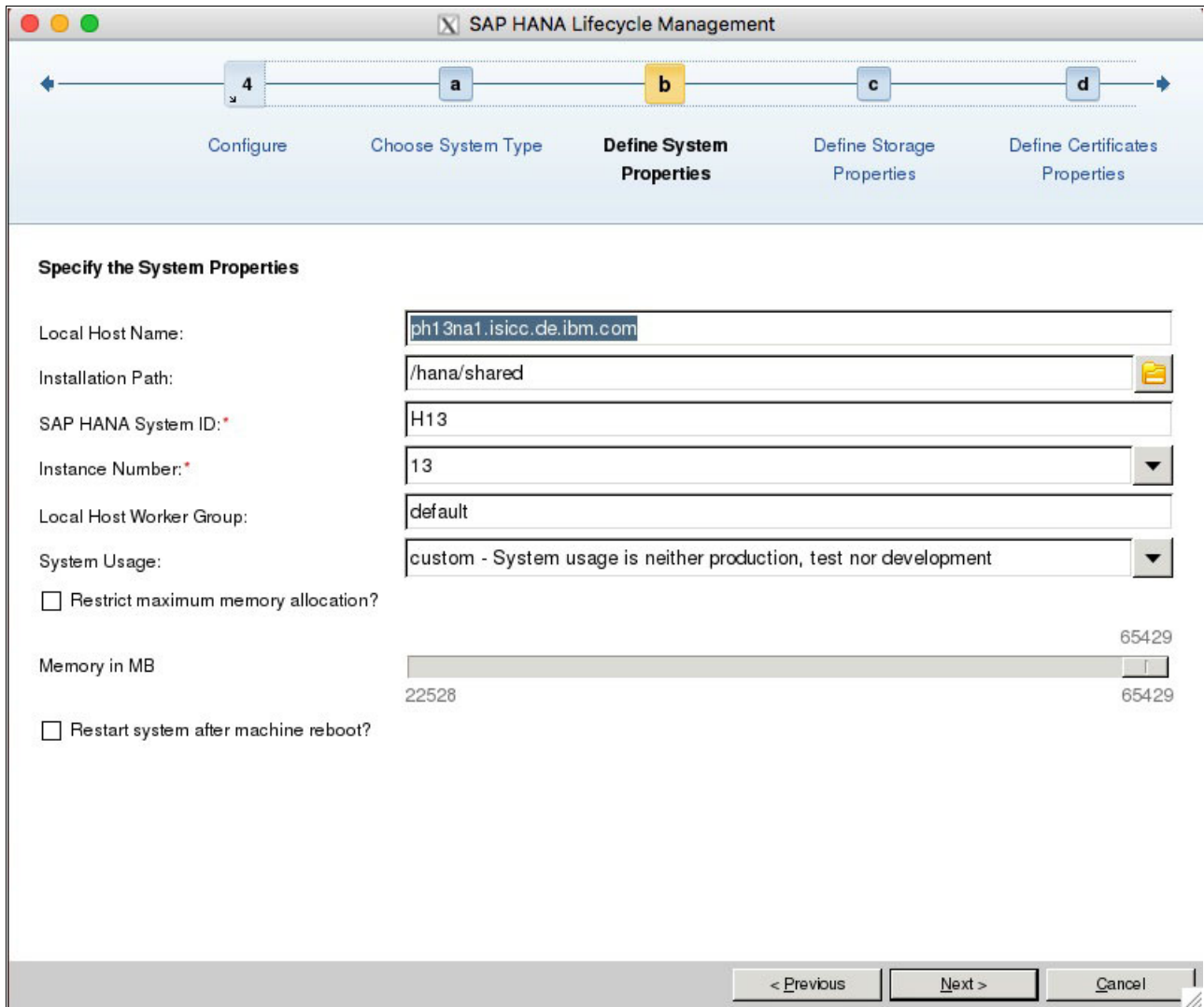

Figure 6-6   HANA installation parameters

9. Provide input for the location of the data and log files. The values are automatically completed based on your SID choice. As a best practice, do not change those values, as shown in Figure 6-7. Click **Next**.
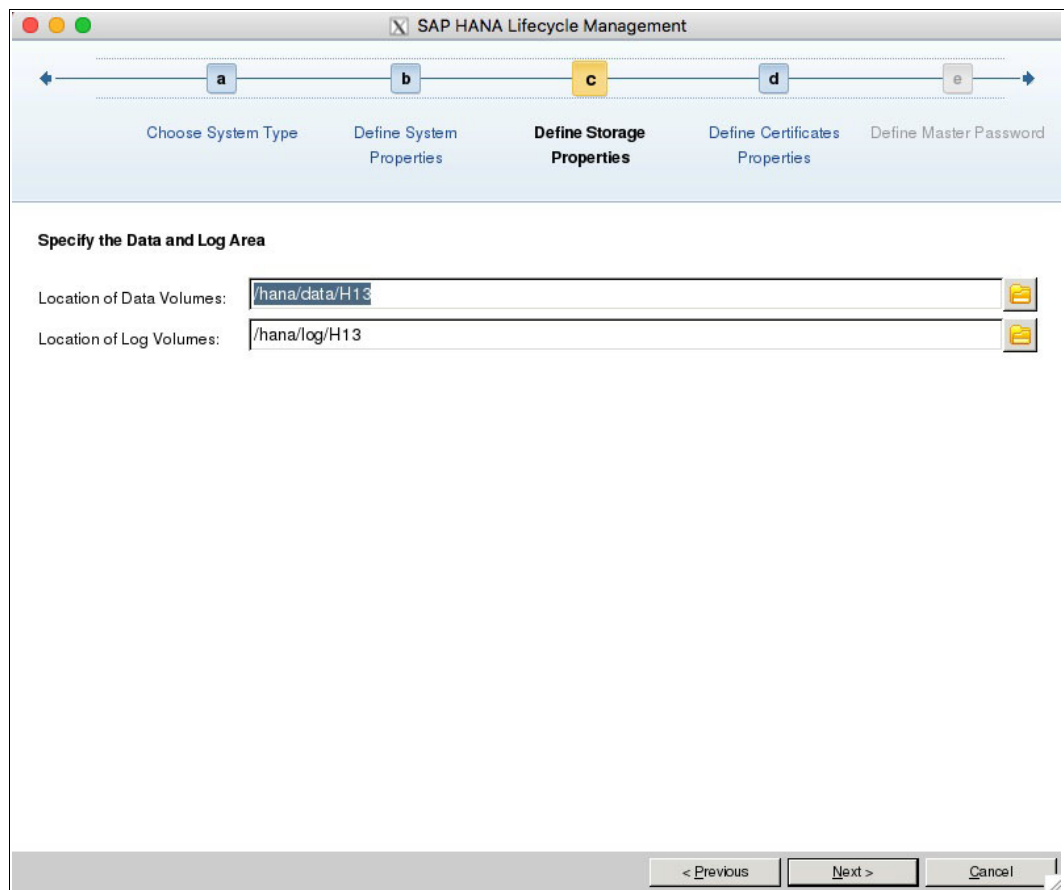

Figure 6-7   HANA storage parameters

10. The window shows information about the host certificate. As a best practice, do not change this information; accept the defaults, as shown in Figure 6-8. Click **Next**.
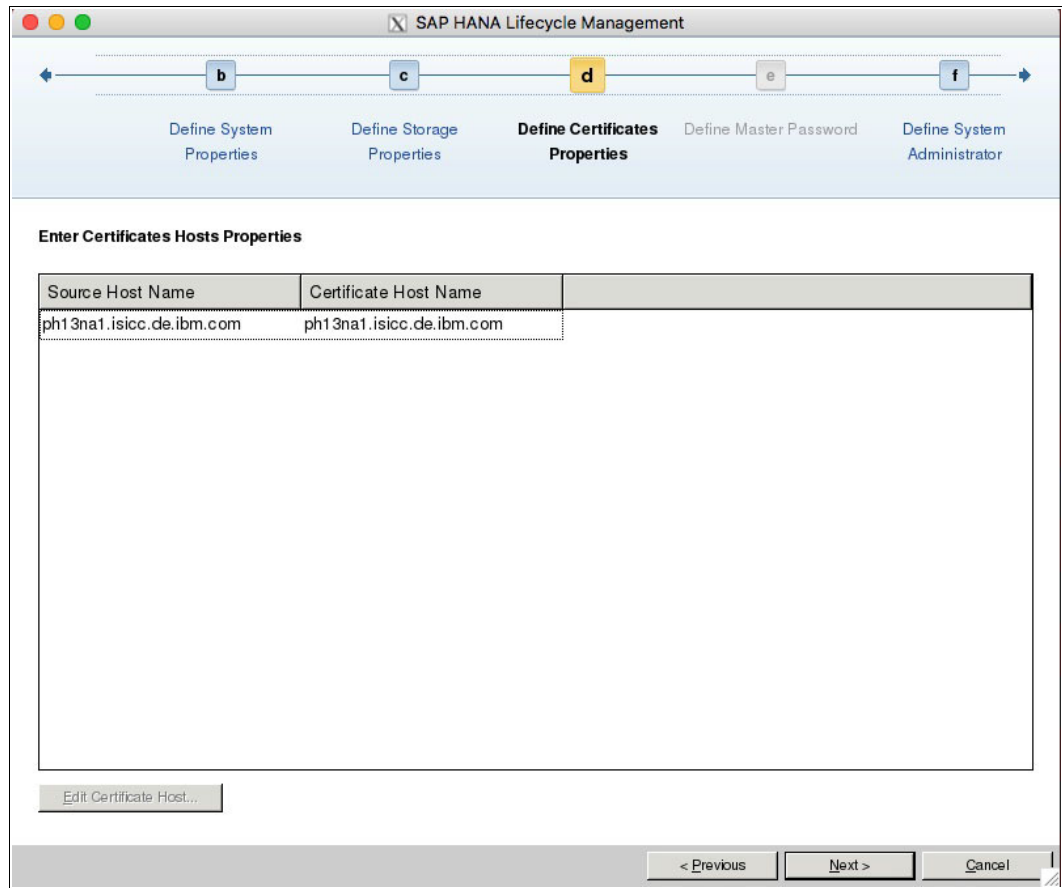

Figure 6-8   HANA installation: Host certificate

11.Set up the `<sid>adm` user password. In our installation scenario, the SID is *h13*, so our user is called `h13adm`. Enter a password, as illustrated in Figure 6-9. We do not make any changes to any of the other parameters. Click **Next**.
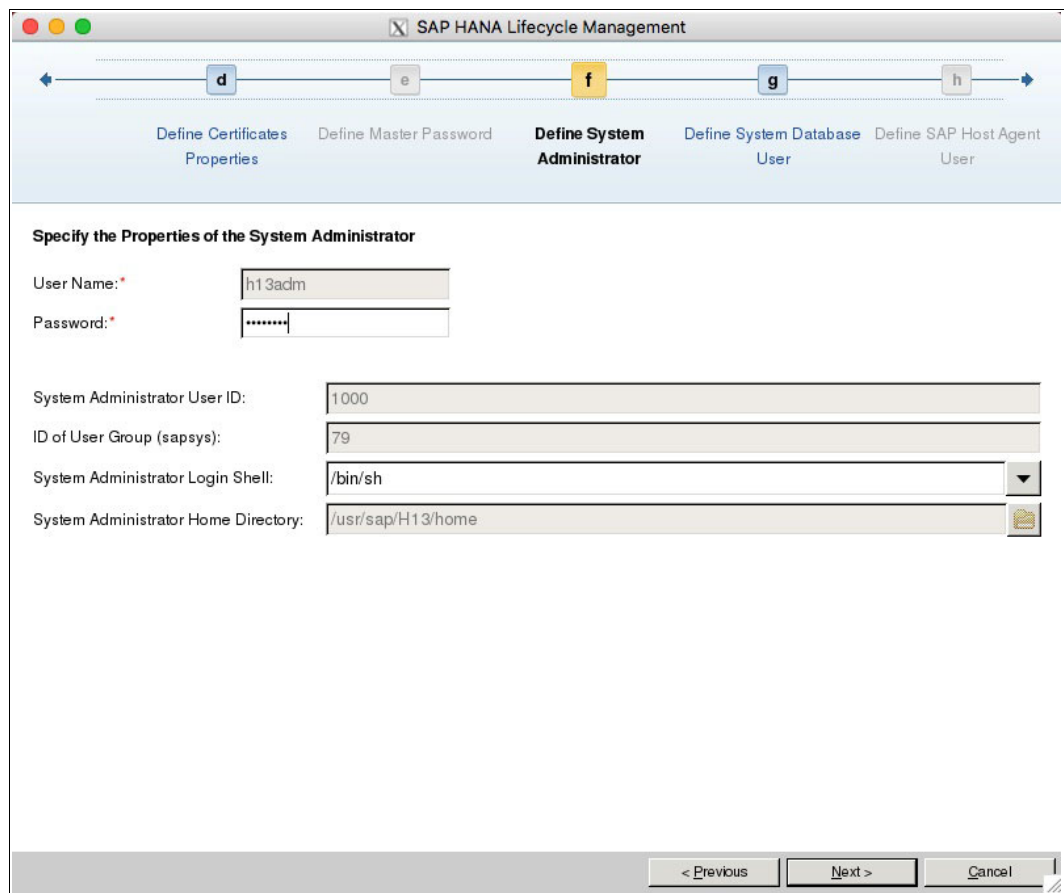
Figure 6-9   HANA <sid>adm user password

> **Note:** For environments that employ central user management, such as Microsoft Active Directory or Lightweight Directory Access Protocol (LDAP), you must create the HANA `<SID>adm` user before running the HANA installation process. This action ensures that the `<SID>adm` user has the proper system user ID of your choice, especially when you already have an SAP landscape with an existing `<SID>adm` user in your Microsoft Active Directory or LDAP user base.

12.Define the password for the DB SYSTEM user, as shown in Figure 6-10.

Figure 6-10   Defining the SYSTEM user password

13.The installation wizard summary opens, as shown in Figure 6-11. Validate that the selections reflect your choices, and then click **Install**.
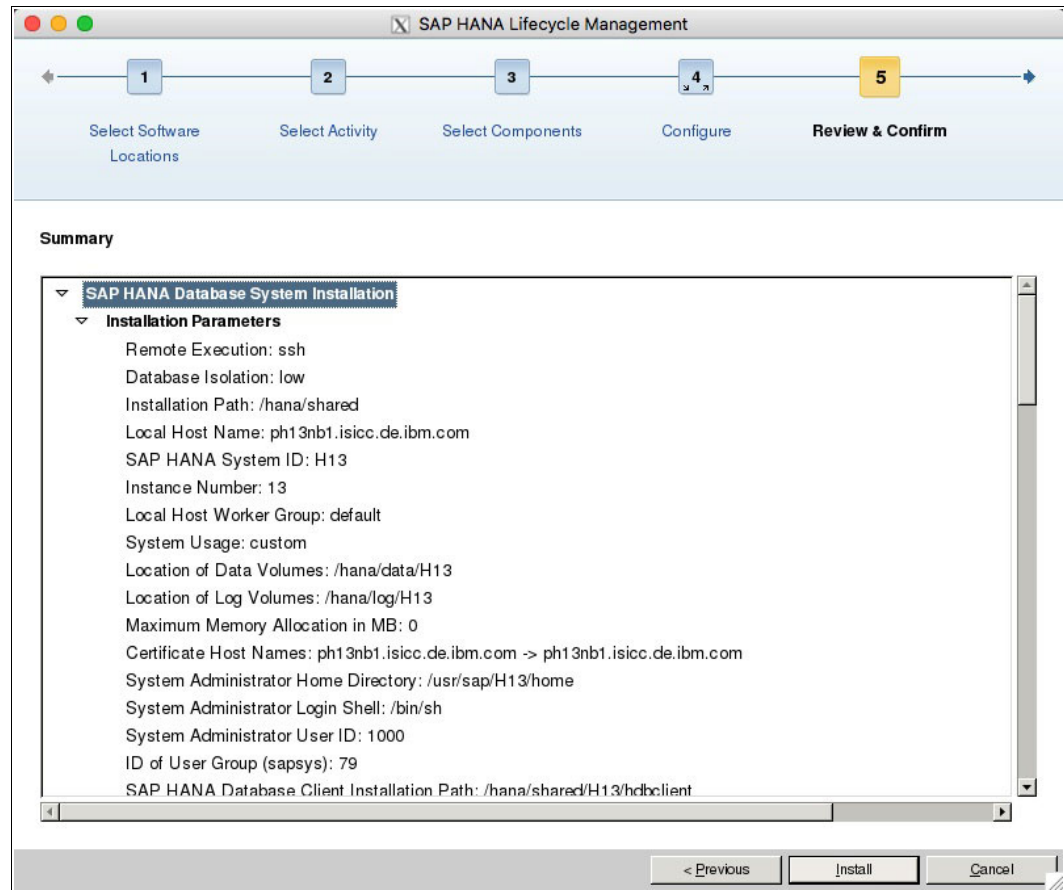
Figure 6-11   HANA installation summary

At the same time, the HANA DB is being installed, you see a window and installation details similar to Figure 6-12 on page 97.
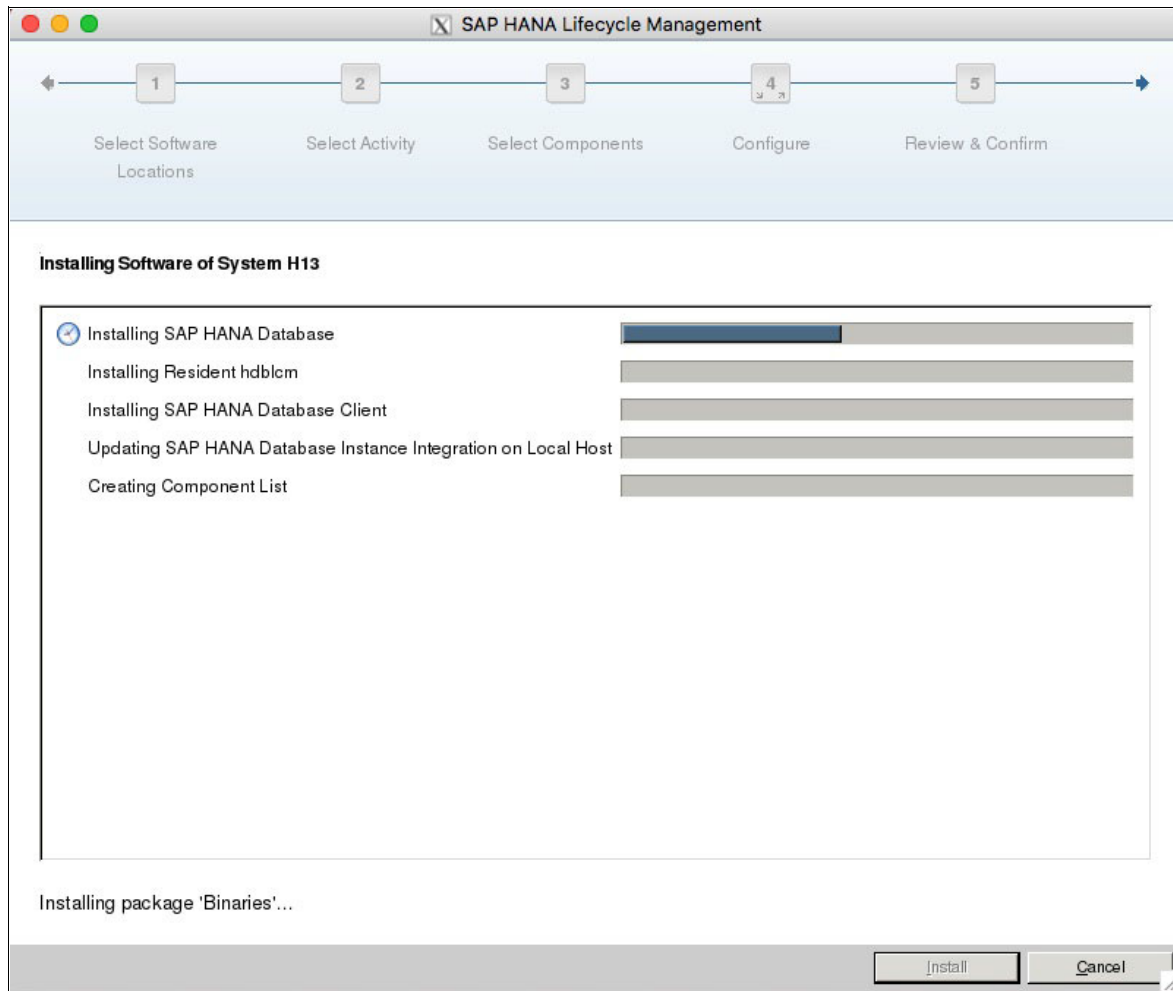
Figure 6-12   HANA system installation progress

When the installation finishes, your HANA DB is installed, running, and fully functional. To set up your HANA Studio connection to your newly installed HANA system, go to 6.3, "Postinstallation notes" on page 101.

6.2.2 Text-mode installation

To start a text-mode installation, you must run the **hdblcm** command in the `DATA_UNITS/HDB_LCM_LINUX_PPC64LE` directory of the folder where you decompressed the HANA installer. This command starts the Life Cycle Manager that is installed as a text-mode wizard that collects your input to drive the installation. All that you need to do is to provide the required information. Except for the SID, instance number, and passwords, all other provided information is standard. The default inputs are between brackets [ ].

Example 6-3 shows the entire text mode installation process, where all of the user inputs are displayed in **bold**. Ensure that you select the server, client components, and any other options that your environment requires. The installer can be used to add more components if required.

Also, because this is a scale-up installation, answer no to the question where it prompts you if you want to add more hosts to the system.

Example 6-3   HANA text mode installation

```
hana002:/tmp/hana20_sps1/51052031/DATA_UNITS/HDB_LCM_LINUX_PPC64LE # ./hdblcm

SAP HANA Lifecycle Management - SAP HANA Database 2.00.010.00.1491294693
**********************************************************************

Scanning Software Locations...
Detected components:
    SAP HANA Database (2.00.010.00.1491294693) in
/tmp/51052031/DATA_UNITS/HDB_SERVER_LINUX_PPC64LE/server
    SAP HANA AFL (incl.PAL,BFL,OFL,HIE) (2.00.010.0000.1491308763) in
/tmp/51052031/DATA_UNITS/HDB_AFL_LINUX_PPC64LE/packages
    SAP HANA EPM-MDS (2.00.010.0000.1491308763) in
/tmp/51052031/DATA_UNITS/SAP_HANA_EPM-MDS_10/packages
    SAP HANA Database Client (2.1.37.1490890836) in
/tmp/51052031/DATA_UNITS/HDB_CLIENT_LINUX_PPC64LE/client
    SAP HANA Smart Data Access (2.00.0.000.0) in
/tmp/51052031/DATA_UNITS/SAP_HANA_SDA_20_LINUX_PPC64LE/packages
    SAP HANA XS Advanced Runtime (1.0.55.288028) in
/tmp/51052031/DATA_UNITS/XSA_RT_10_LINUX_PPC64/packages
    GUI for HALM for XSA (including product installer) Version 1 (1.11.3) in
/tmp/51052031/DATA_UNITS/XSA_CONTENT_10/XSACALMPIUI11_3.zip
    XSAC FILEPROCESSOR 1.0 (1.000.1) in
/tmp/51052031/DATA_UNITS/XSA_CONTENT_10/XSACFILEPROC00_1.zip
    SAP HANA tools for accessing catalog content, data preview, SQL console, etc.
(2.001.3) in /tmp/51052031/DATA_UNITS/XSAC_HRTT_20/XSACHRTT01_3.zip
    XS Monitoring 1 (1.004.0) in
/tmp/51052031/DATA_UNITS/XSA_CONTENT_10/XSACMONITORING04_0.ZIP
    Develop and run portal services for customer apps on XSA (1.001.0) in
/tmp/51052031/DATA_UNITS/XSA_CONTENT_10/XSACPORTALSERV01_0.zip
    SAP Web IDE Web Client (4.001.0) in
/tmp/51052031/DATA_UNITS/XSAC_SAP_WEB_IDE_20/XSACSAPWEBIDE01_0.zip
    XS Services 1 (1.004.2) in
/tmp/51052031/DATA_UNITS/XSA_CONTENT_10/XSACSERVICES04_2.ZIP
    SAPUI5 FESV3 XSA 1 - SAPUI5 SDK 1.44 (1.044.10) in
/tmp/51052031/DATA_UNITS/XSA_CONTENT_10/XSACUI5FESV344_10.zip
    XSAC XMLA Interface For Hana 1 (1.000.1) in
/tmp/51052031/DATA_UNITS/XSA_CONTENT_10/XSACXMLAINT00_1.zip
    Xsa Cockpit 1 (1.000.0) in
/tmp/51052031/DATA_UNITS/XSA_CONTENT_10/XSACXSACOCKPIT00_0.zip


Choose an action

  Index | Action            | Description
  -----------------------------------------------
  1     | install           | Install new system
  2     | extract_components | Extract components
```

```
   3    | Exit (do nothing)  |
```

Enter selected action index [3]: 1

SAP HANA Database version '2.00.010.00.1491294693' will be installed.

Select additional components for installation:

```
  Index | Components | Description

--------------------------------------------------------------------------------
   1    | server     | No additional components
   2    | all        | All components
   3    | afl        | Install SAP HANA AFL (incl.PAL,BFL,OFL,HIE) version
2.00.010.0000.1491308763
   4    | client     | Install SAP HANA Database Client version 2.1.37.1490890836
   5    | smartda    | Install SAP HANA Smart Data Access version 2.00.0.000.0
   6    | xs         | Install SAP HANA XS Advanced Runtime version 1.0.55.288028
   7    | epmmds     | Install SAP HANA EPM-MDS version 2.00.010.0000.1491308763
```

Enter comma-separated list of the selected indices [4]: 1,4
Enter Installation Path [/hana/shared]:
Enter Local Host Name [hana002]:
Do you want to add hosts to the system? (y/n) [n]:
Enter SAP HANA System ID: RB1
Enter Instance Number [00]: 00
Enter Local Host Worker Group [default]:

```
  Index | System Usage | Description
  ----------------------------------------------------------------------------
   1    | production   | System is used in a production environment
   2    | test         | System is used for testing, not production
   3    | development  | System is used for development, not production
   4    | custom       | System usage is neither production, test nor development
```

Select System Usage / Enter Index [4]: 1
Enter Location of Data Volumes [/hana/data/RB1]:
Enter Location of Log Volumes [/hana/log/RB1]:
Enter Certificate Host Name For Host 'hana002' [hana002]:
Enter SAP Host Agent User (sapadm) Password: ********
Confirm SAP Host Agent User (sapadm) Password: ********
Enter System Administrator (rb1adm) Password: ********
Confirm System Administrator (rb1adm) Password: ********
Enter System Administrator Home Directory [/usr/sap/RB1/home]:
Enter System Administrator Login Shell [/bin/sh]:
Enter System Administrator User ID [1001]:
Enter ID of User Group (sapsys) [79]:
Enter Database User (SYSTEM) Password: ********
Confirm Database User (SYSTEM) Password: ********
Restart system after machine reboot? [n]:

Summary before execution:
=========================

SAP HANA Database System Installation

```
    Installation Parameters
       Remote Execution: ssh
       Database Isolation: low
       Installation Path: /hana/shared
       Local Host Name: hana002
       SAP HANA System ID: RB1
       Instance Number: 00
       Local Host Worker Group: default
       System Usage: production
       Location of Data Volumes: /hana/data/RB1
       Location of Log Volumes: /hana/log/RB1
       Certificate Host Names: hana002 -> hana002
       System Administrator Home Directory: /usr/sap/RB1/home
       System Administrator Login Shell: /bin/sh
       System Administrator User ID: 1001
       ID of User Group (sapsys): 79
       SAP HANA Database Client Installation Path: /hana/shared/RB1/hdbclient
    Software Components
       SAP HANA Database
          Install version 2.00.010.00.1491294693
          Location: /tmp/51052031/DATA_UNITS/HDB_SERVER_LINUX_PPC64LE/server
       SAP HANA AFL (incl.PAL,BFL,OFL,HIE)
          Do not install
       SAP HANA EPM-MDS
          Do not install
       SAP HANA Database Client
          Install version 2.1.37.1490890836
          Location: /tmp/51052031/DATA_UNITS/HDB_CLIENT_LINUX_PPC64LE/client
       SAP HANA Smart Data Access
          Do not install
       SAP HANA XS Advanced Runtime
          Do not install


Do you want to continue? (y/n): y


Installing components...
Installing SAP HANA Database...

[ ... Cropped ... ]

Creating Component List...
SAP HANA Database System installed
You can send feedback to SAP with this form:
https://hana002:1129/lmsl/HDBLCM/RB1/feedback/feedback.html
Log file written to
'/var/tmp/hdb_RB1_hdblcm_install_2017-07-04_14.13.04/hdblcm.log' on host
'hana002'.
```

At this point, your HANA DB is installed, running, and fully functional.

6.3 Postinstallation notes

After you have installed HANA on your system, this section shows useful information to help you get started accessing and managing your DB. Complete the following steps:

1. First, you must install HANA Studio on your workstation. The installer is inside the `DATA_UNITS/HDB_STUDIO_<platform>` folder in your HANA system. Use an SCP client to copy the folder corresponding to your workstation architecture and install the software on your workstation. In our case, we use the MacOS 64-bit version of HANA Studio.

2. After installation, when you open HANA Studio on your workstation, you see a window similar to Figure 6-13. Add a connection to your newly installed HANA instance by using the **Add System** button on the left side of the Systems navigation tab, as shown by the red arrow in Figure 6-13.
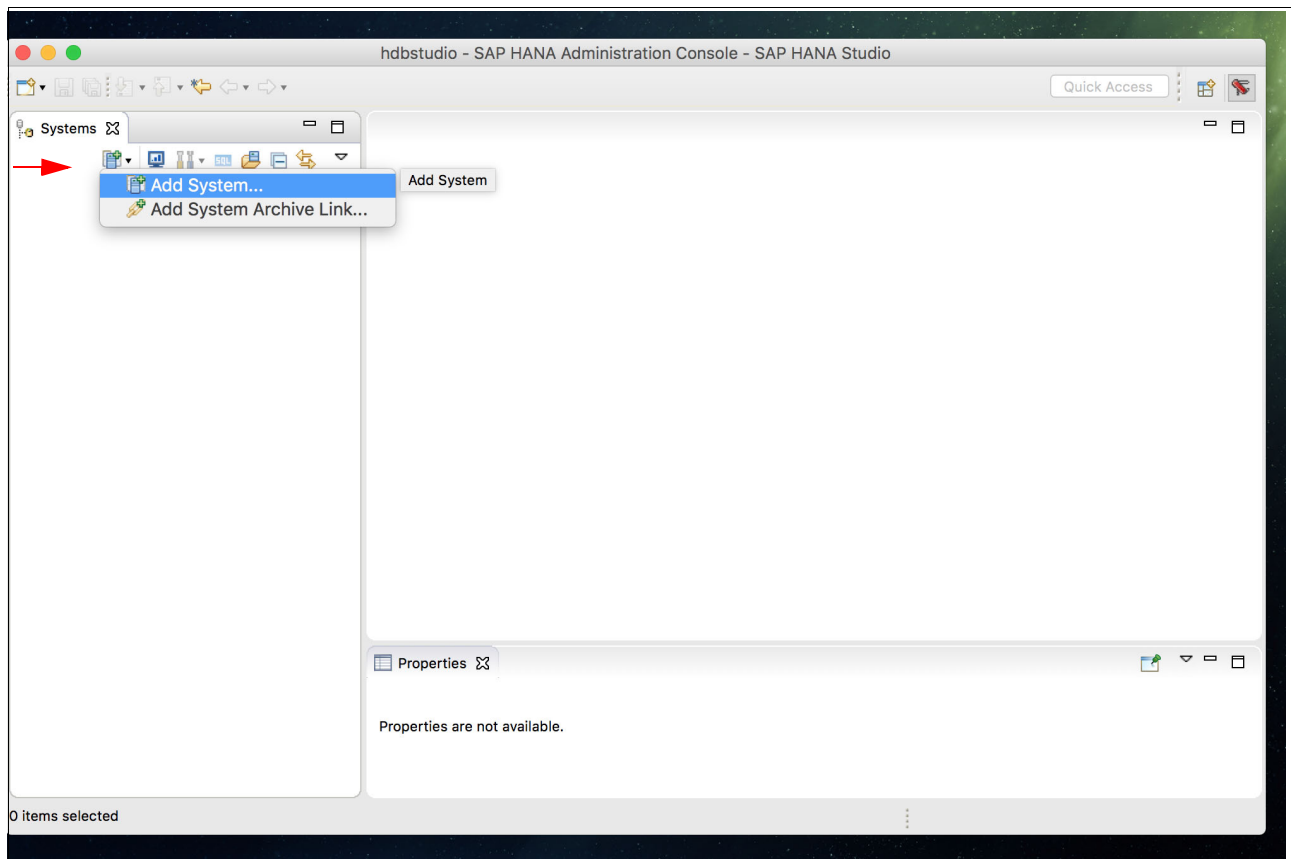

Figure 6-13   HANA Studio console: Adding a system

3. Input the required information, as shown in Figure 6-14. If your HANA environment is already registered in your DNS server, use the host name to complete the information that is depicted as **1** in the figure; otherwise, use the IP address. Complete the instance number information for **2**. Starting with HANA 2.0 SPS1, all systems are multi-tenant, so make the proper selection as described by **3** and use the `SYSTEM` DB user to connect to the DB, as shown by **4**. Give it a proper description, as outlined by **5**, and click **Next**.

> **Note:** When you add `System` to HANA, you must select **Multiple containers** because HANA V2.0 sps01 uses Multiple containers DB mode. Otherwise, you see the following error message.

Specify System

Specify the host name and instance number of the system.

Host Name:	ℹ️ 10.10.12.81 **1**
Instance Number:	00 **2**
Mode:	⚪ Single container
	🔘 Multiple containers **3**
	⚪ Tenant database
	Name:
	🔘 System database **4**
Description:	RB1 database **5**
Locale:	English (United States)
Folder:	/ Browse...

?   < Back   Next >   Cancel   Finish

Figure 6-14   HANA Studio: Information to add a HANA instance

4. Configure your connection to the DB by using the `SYSTEM` user and its password, as shown in Figure 6-15. Click **Finish** to complete this setup.
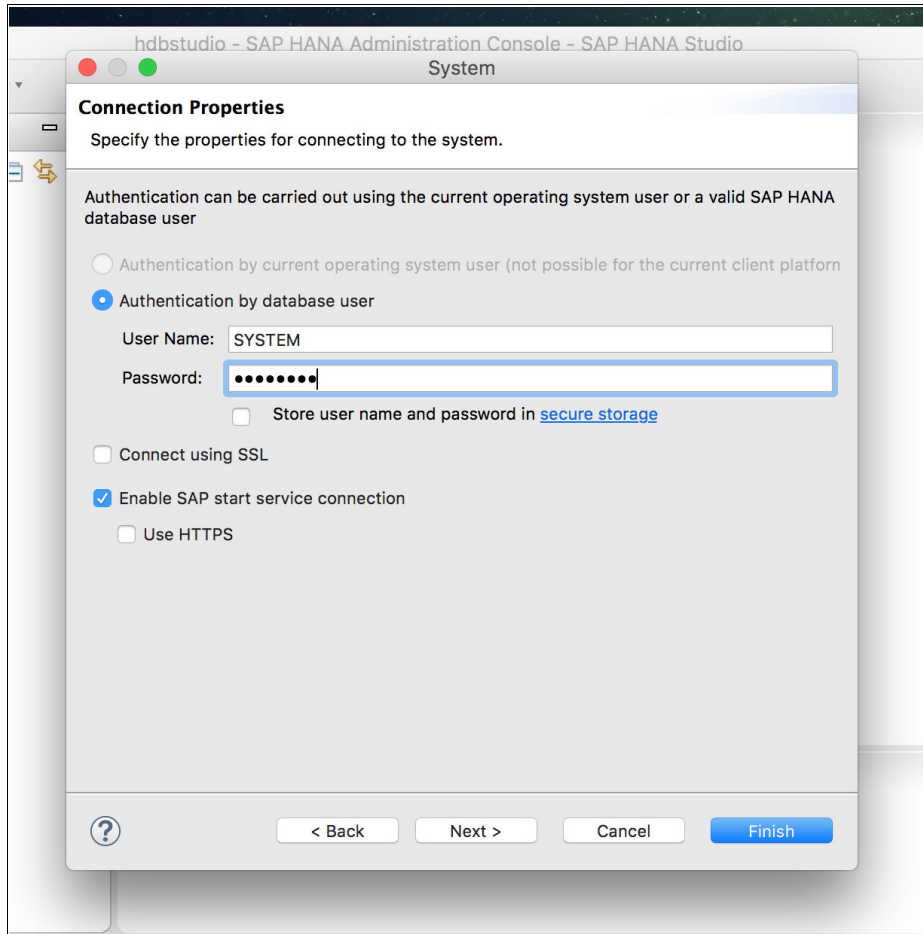

Figure 6-15   HANA Studio: Connecting to the HANA instance

After completion, double-click your instance entry on the left side of the **Systems** navigation menu to open its properties, as shown in Figure 6-16. Go to the **Landscape** tab of your instance to validate that all services are running.
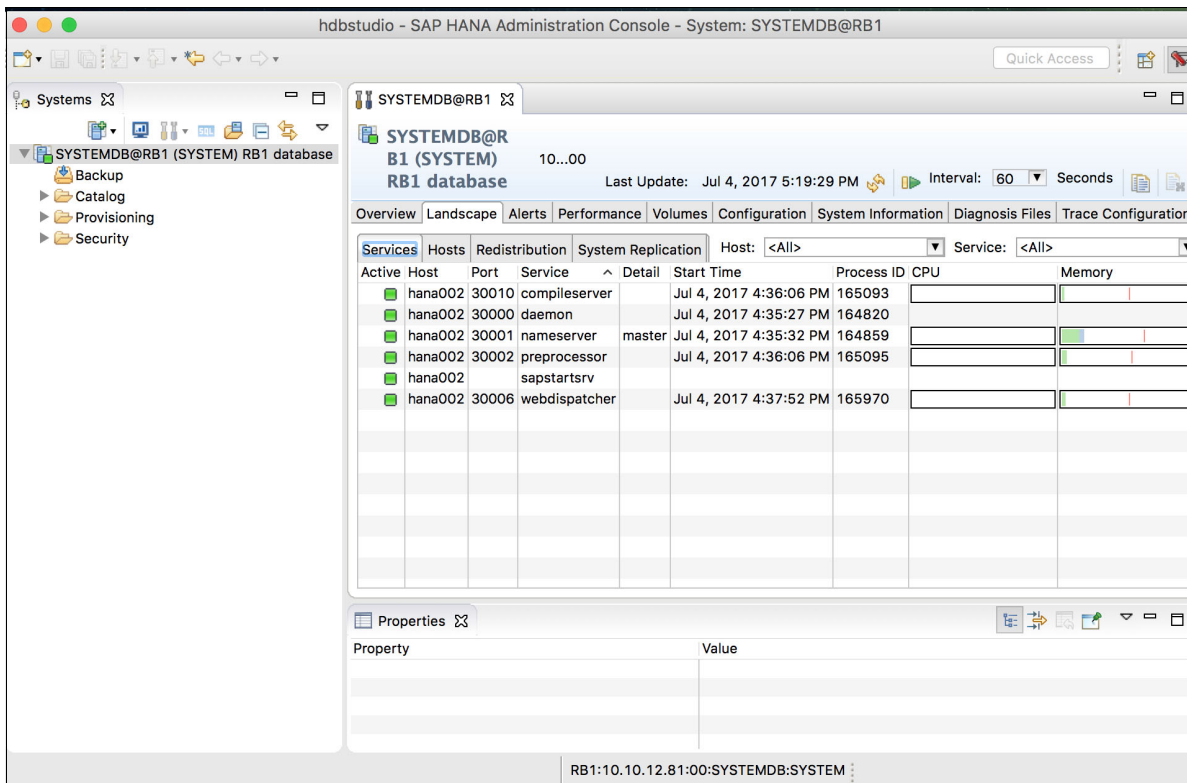

Figure 6-16   HANA Studio: Checking the Landscape tab of your HANA system

You now have full control of your HANA system. From within the HANA Studio interface, you can take backups of your instance, configure HSR, change the configuration of the DB parameters, and much more.

Note: It is a best practice to configure the backup strategy of your HANA system.

For a complete guide about how to manage a HANA DB, see *SAP HANA Administration Guide*.

If you are planning to configure HSR between two scale-up systems, or configuring high availability (HA) with the SUSE HA Extension, see Chapter 7, "SAP HANA System Replication for high availability and disaster recovery scenarios" on page 105.

7

SAP HANA System Replication for high availability and disaster recovery scenarios

This chapter introduces the concepts of SAP HANA System Replication (HSR), which is a feature that is included in HANA, and can be used for replicating data between HANA systems.

This chapter also provides a step-by-step description of how to set up HSR and perform some simple tests before you place HSR into production.

This chapter covers the following topics:

- ► SAP HANA System Replication methods
- ► Implementing SAP HANA System Replication
- ► SAP HANA System Replication and takeover tests

7.1 SAP HANA System Replication methods

HSR has two fundamental concepts that you must understand:

► The replication mode: Either *synchronous to memory*, *synchronous*, *synchronous with full sync option*, or *asynchronous*.

► The operation mode: Either *delta shipping*, *log replay*, or *log replay with read access*.

This section explains these two concepts in more detail.

SAP HANA System Replication replication mode

The HSR replication mode determines the behavior of the secondary system in terms of acquiring the data from the primary system and having it ready for use.

One of the four replication modes is the *asynchronous* mode, in which the primary system sends the data to the secondary one but does not wait for the secondary to acknowledge it. In this manner, commit operations on the primary node do not require an acknowledgment from the secondary one, which does not delay the overall process for creating or modifying data on the primary system. However, an asynchronous operation means that your recovery point objective (RPO) is nonzero. To achieve zero data loss, you must use synchronous replication, which is usually done between systems at the same site. Nevertheless, the asynchronous replication mode is the only alternative for replicating data between distant data centers because making it synchronous in this case adds too much latency to commit operations in the primary site.

For the *synchronous* mode, all data that is created or modified on the primary system is sent to the secondary system, and commit operations must wait until the secondary system replies that the data is safely stored on its disks. This process ensures an RPO of zero, but adds latency to all write operations. If you choose to replicate data by using this mode, your systems must adhere to the file system key performance indicators (KPIs) in which the maximum latency for log writing is 1000 μs. In essence, the network connection between your two systems must provide a low latency to enable this setup.

Another mode is the *synchronous in-memory*, where replication also happens synchronously, but the commit is granted after the replicated data is stored onto the secondary node's memory. After that, the secondary node makes this data persistent on its disks. As the memory of the secondary system is already loaded with data, the takeover is much faster.

The last mode is the *synchronous with full sync* option, which requires the secondary node to be online and receiving the replicated data. If the secondary node goes offline, all transactions on the primary node are suspended until the secondary node comes back online.

SAP HANA System Replication operation mode

The HSR operation mode determines how the data is sent and processed by the secondary system.

The *delta shipping* mode (`delta_shipping`) is the classic mode, and it works by shipping all of the modified data to the secondary system every 10 minutes by default. During these 10-minute intervals, the redo-logs are shipped to the secondary system. So, when a failure of the primary node occurs, the secondary processes the redo-logs since the last delta shipping point.

In the *log replay* mode (`logreplay`), only the redo-logs are sent over to the secondary system, and they are immediately replayed. This mode makes the transfer lighter in terms of network consumption, and also makes the takeover operations faster.

The *log replay with read access* mode (`logreplay_readaccess`) works in a similar fashion to the log replay mode, except that it also enables the receiver to operate in read-only mode, so you can query the database (DB) with proper SQL statements.

7.1.1 SAP HANA System Replication requirements

There are a few systems requirements that you must meet before you start the configuration of HSR:

► All participating hosts must be able to ping each other by name, either by having the DNS server resolve the information or by using local `/etc/hosts` entries.

► The date and time of all participating hosts must be synchronized with a Network Time Protocol (NTP) server. If you do not have any suitable NTP servers in your network, choose one of the hosts to act as one, and have all the other hosts synchronize their time with it.

► Both HANA systems must have the same SAP ID (SID) and instance number. Also, the next instance number must be available on both systems, that is, if you use instance number `00`, then `01` is also available. Host names can be different (a virtual IP and host name are defined when using HSR with SUSE High Availability Extension (HAE)).

► Both HANA instances that take part in the replication must have a current backup. If you have not implemented a backup strategy, take a file-based backup from HANA Studio as a temporary solution.

► The *systemPKI SSFS data* and *key* must be exchanged between the two systems, or the secondary system cannot be registered for the replication. This is a new requirement for HSR starting with HANA V2.0.

► The secondary HANA system must be turned off before configuring it to replicate data from the primary system.

You can use an isolated network or VLAN with HSR. It is not mandatory, but it is a best practice to avoid having the replication network traffic compete with the data traffic on the same logical network interface. If you decide to follow this best practice, see "Using a dedicated network for SAP HANA System Replication" because there are some extra configuration steps that you must make to your HANA instances `global.ini` profile.

Using a dedicated network for SAP HANA System Replication

The primary network for HANA is the data network through which the application servers access the DB. When you use a secondary network to handle HSR, your network layout looks similar to Figure 7-1, where HSR is set up between two scale-up systems on the same site.
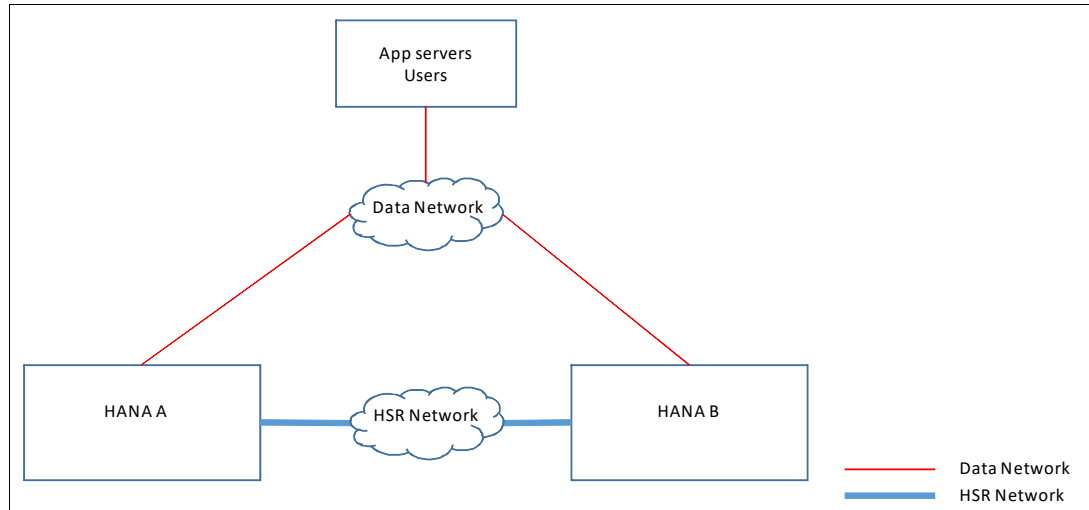

Figure 7-1   HANA environment with a separate network for SAP HANA System Replication

In our implementation, the `10.10.12.0/24` network is used by our hosts for data traffic, and for communicating with the external world. The `192.168.1.0/24` network is defined for HSR. Example 7-1 illustrates how our `/etc/hosts` file is configured to meet the second requirement in 7.1.1, "SAP HANA System Replication requirements" on page 107. Note the sections in **bold** with the definitions for the IP addresses for the data and replication networks.

Example 7-1   The /etc/hosts file from host hana002

```
# IP-Address  Full-Qualified-Hostname  Short-Hostname

127.0.0.1       localhost

# special IPv6 addresses
::1             localhost ipv6-localhost ipv6-loopback

fe00::0         ipv6-localnet

ff00::0         ipv6-mcastprefix
ff02::1         ipv6-allnodes
ff02::2         ipv6-allrouters
ff02::3         ipv6-allhosts

#Data network - RB1 hosts
10.10.12.81     hana002.pok.stglabs.ibm.com hana002
10.10.12.82     hana003.pok.stglabs.ibm.com hana003

#HSR network - RB1 hosts

192.168.1.81    hana002-rep.pok.stglabs.ibm.com  hana002-rep
192.168.1.82    hana003-rep.pok.stglabs.ibm.com  hana003-rep
```

When configuring HSR, you must use the host name of the system to create the replication scheme. However, the host names of the systems are always bound to the primary interface that is used for data traffic. So, if you do not explicitly instruct HANA to use the replication interface with HSR, it ends up using the data network.

All HANA instances have a `global.ini` configuration file. There are many parameters that can be changed within this file. To instruct HANA to use the replication interface with HSR, you must edit the following file:

`/hana/shared/<SID>/exe/linuxppc64le/HDB_<HANA version>/config/global.ini`

Open this file by using your favorite text editor and look for a section that is named "`[system_replication_hostname_resolution]`", as shown in Example 7-2. It is empty when you first view it, and then you add the lines for your hosts, as shown by the highlights **1** and **2**. These are the entries that instruct HANA to use the IP addresses in the `192.168.1.0/24` network for HSR. You are still able to use the systems host name for replication.

Example 7-2   Configuring the global.ini file to use a dedicated replication network for SAP HANA System Replication

```
[... cropped ...]
# .short_desc
# specifies the resolution of remote site hostnames to addresses for system replication
# .full_desc
# specifies the resolution of remote site hostnames to addresses for system replication.
# for multi-tier replication only the direct neighbors must be specified
# Format:   ipaddress = internal hostname
# e. g.   192.168.100.1 = hanahost01
[system_replication_hostname_resolution]
192.168.1.81=hana002 1
192.168.1.82=hana003 2
#
# .short_desc
# Configuration of system replication communication settings
# .full_desc
# This section contains parameters that are related to configuration
# of various system replication communication settings.
[system_replication_communication]

# .short_desc
# the network interface the processes shall listen on, if system replication is enabled.
# .full_desc
#
# Possible values are:
#  - .global: all interfaces
listeninterface = .global 3
[... cropped ...]
```

Also, check that the **listeninterface** parameter is set to `.global` inside the section `[system_replication_communication]`, as shown in Example 7-2 and depicted by **3**.

> **Special consideration:** You must be careful when you change the value of the **listeninterface** parameter that is under the `[system_replication_communication]` section because there are two different **listeninterface** parameters in that file, and the other one has nothing to do with HSR.

7.2 Implementing SAP HANA System Replication

This section illustrates how to replicate data between two scale-up systems. Both systems already meet the three requirements that are described in 7.1.1, "SAP HANA System Replication requirements" on page 107. Complete the following steps:

1. You must take a backup. Considering that we are using an ordinary lab environment that contains no data inside the DB, we take a file-based backup. To do so, right-click your HANA instance inside HANA Studio and go through the menu to trigger a backup of the SYSTEM DB, as shown in Figure 7-2. This action opens a wizard that defaults to taking a file-based backup if you do not change any of its predefined parameters. Create the backup and proceed. Then, back up your SID tenant DB by repeating the process, but instead choose **Back Up Tenant Database**, as shown in Figure 7-2. Back up both the source and destination systems, which take part in the replication setup.
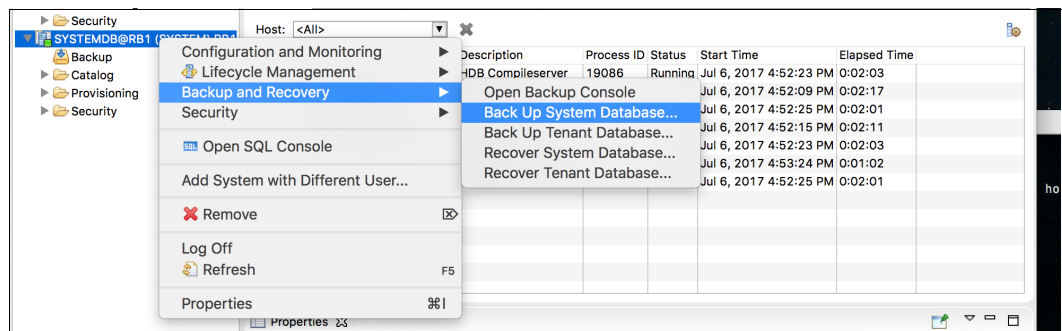

Figure 7-2   Backing up the HANA instance

2. Get the systemPKI, the SSFS data, and key from the source system and copy them to the secondary system. The locations where these files are kept are:

   – /hana/shared/<SID>/global/security/rsecssfs/data/SSFS_<SID>.DAT
   – /hana/shared/<SID>/global/security/rsecssfs/key/SSFS_<SID>.KEY

   Example 7-3 shows how to perform these operations. In our example, we are logged on one of the servers (hana002) and copy the files over to the other server (hana003).

Example 7-3   Exchanging the systemPKI SSFS data, and key files

```
hana002:/ # cd /hana/shared/RB1/global/security/rsecssfs
hana002:/hana/shared/RB1/global/security/rsecssfs # ls
data  key
hana002:/hana/shared/RB1/global/security/rsecssfs # scp data/SSFS_RB1.DAT \
root@hana003:/hana/shared/RB1/global/security/rsecssfs/data/
Password:
SSFS_RB1.DAT                                      100% 2960     2.9KB/s
00:00
hana002:/hana/shared/RB1/global/security/rsecssfs # scp key/SSFS_RB1.KEY \
root@hana003:/hana/shared/RB1/global/security/rsecssfs/key
Password:
SSFS_RB1.KEY                                      100%  187     0.2KB/s
00:00
```

3. Enable replication on the source system. Avoid calling the systems *primary* and *secondary* because their roles are interchangeable as you perform takeovers. This is why we described them as *source* and *destination*, which is really how SAP refers to them in its replication wizard.

4. Enable HSR on the source system by right-clicking the source system in the HANA Studio and go to the menu that is shown in Figure 7-3.
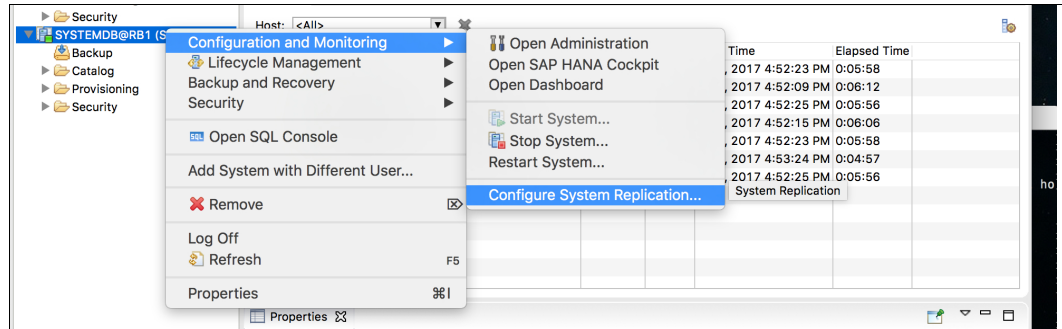

Figure 7-3   Enabling SAP HANA System Replication on the source system

5. Enable **System Replication**, as shown in Figure 7-4. All other options are disabled because they are not applicable to a source system that is being set up for replication. Click **Next**.
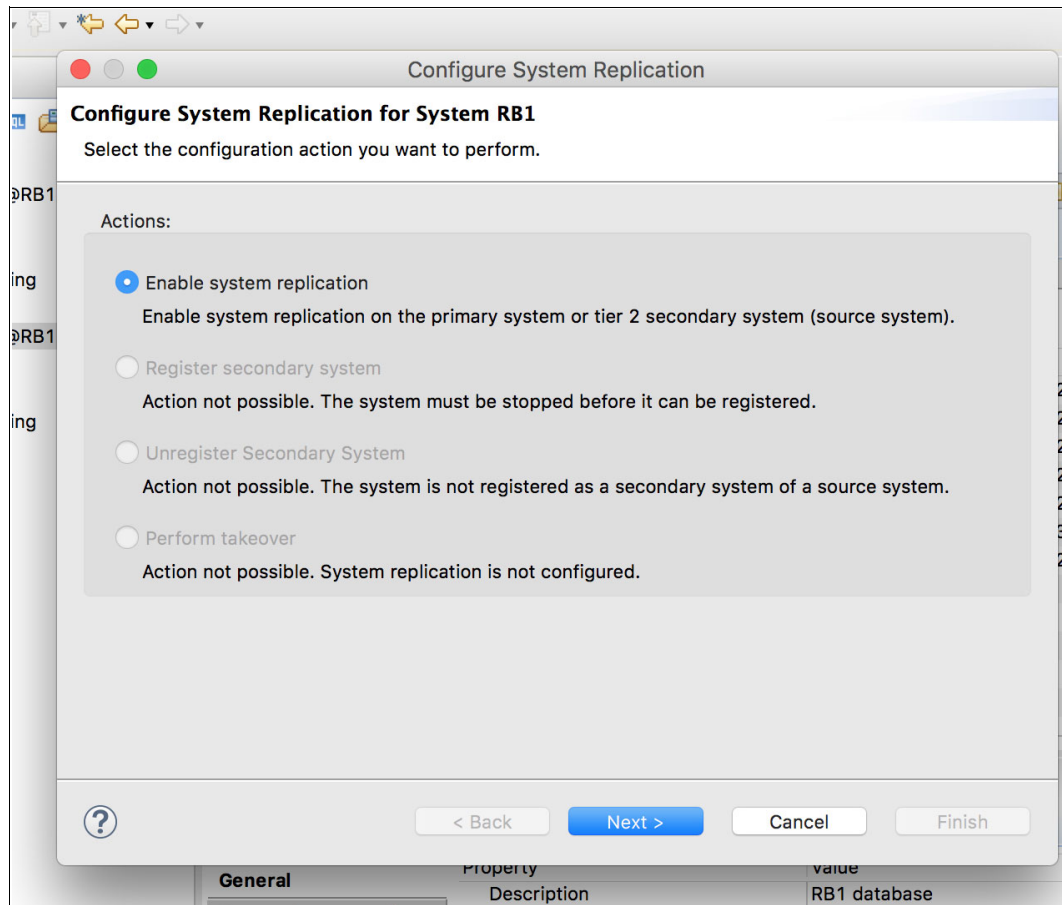
Figure 7-4   Enabling System Replication

6. Assign a Primary System Logical Name to your source system, as shown in Figure 7-5. This parameter is also known as the *site name*, and the HANA Studio HSR status information under the **Landscape →System Replication** tab refers to it as site name. In our example, we decide to name this first system MONTANA. Good choices are names that relate to the physical location of your system, such as data center locations or building names.
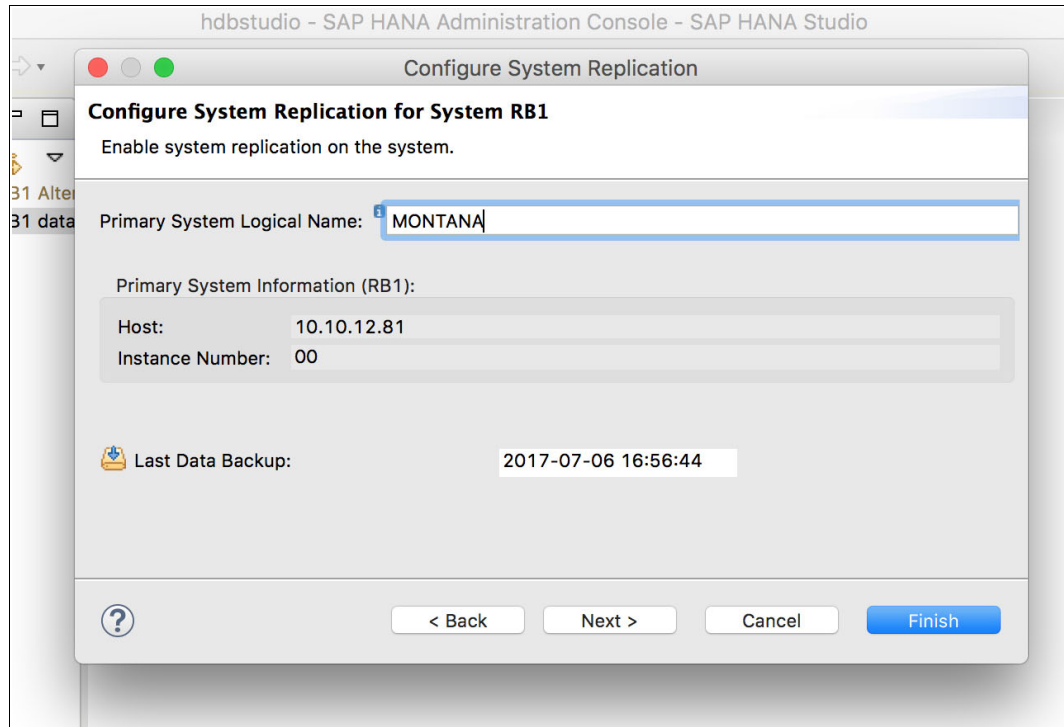
Figure 7-5   Assigning a primary system logical name

That is all there is to setting up the source system for HSR. Let us now set up the destination system. The destination instance must be shut down.

7. After the destination instance is shut down, right-click your secondary instance in the HANA Studio and open the replication wizard by going to the menu that is shown in Figure 7-3 on page 111. After the system is shut down, then the options for replication are different, as shown in Figure 7-6. The only action that you can take is to register the system as the secondary system for replication, that is, the one that receives the replica. Select **Register secondary system** and click **Next**.
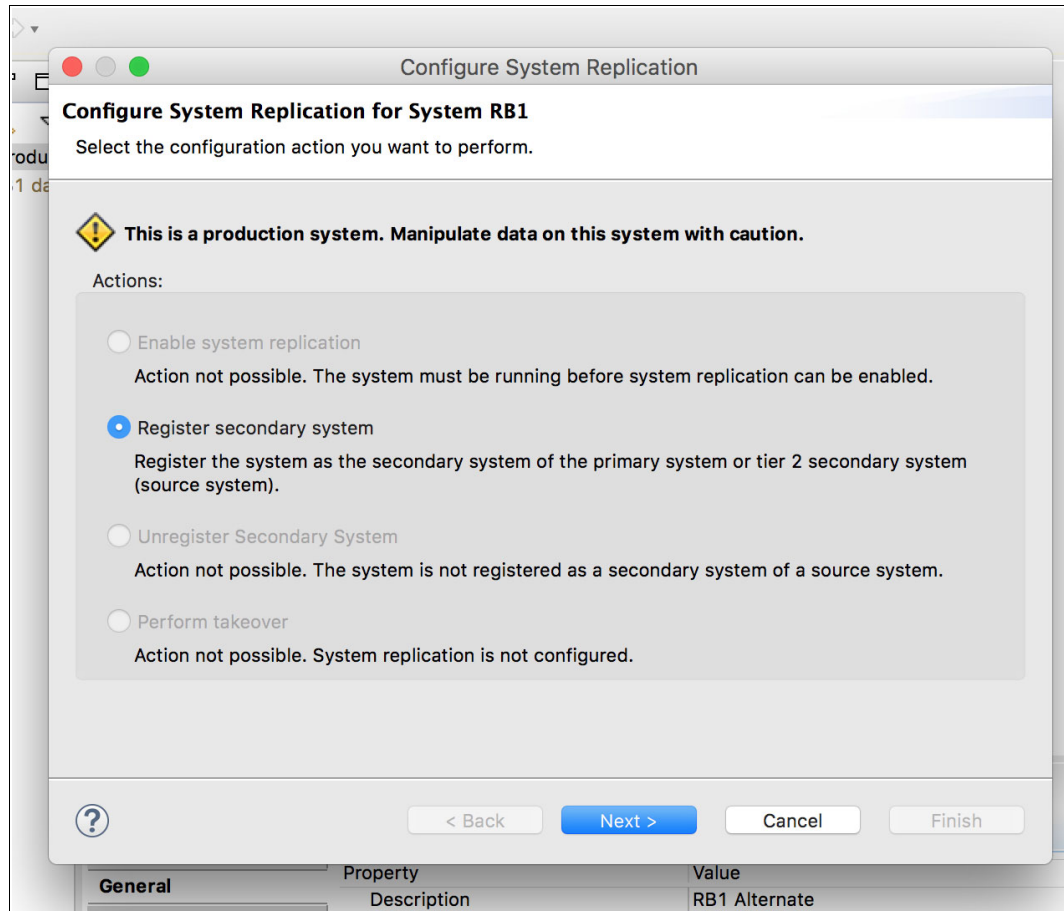
Figure 7-6   Registering the secondary system for replication

8. Set the **Secondary System Logical** name and the replication parameters for replication mode and operation mode. In our example, we use `CYRUS` as the secondary system logical name (or site name), as shown by **1** in Figure 7-7. We choose to replicate it *synchronously*, as shown by **2**. The operation mode that we choose is log replay, as shown by **3**. Also, we make this destination system point to the source one to get the data, as shown by **4**, by using the host name and HANA DB instance number of the source system. Click **Finish** to complete the wizard.

Figure 7-7   Choosing the replication parameters

9. The destination system starts in replication mode. You can check that the replication is running by double-clicking the source instance in the HANA Studio to open its properties, and then selecting **Landscape** →**System Replication**, as shown in Figure 7-8. Check all entries that are marked as `ACTIVE` in the column `REPLICATION_STATUS`.

Figure 7-8   Synchronized SAP HANA System Replication

7.3 SAP HANA System Replication and takeover tests

This section explains how to perform takeovers between two HANA systems by using System Replication. The next sections instruct you about how to create a simple table, add it, and then list an entry in it.

The following sections assume that your HSR is configured and synchronized.

7.3.1 Creating a test table and populating it

To create a table in your HANA DB, complete the following steps:

1. Open the SQL console inside HANA Studio, as shown in Figure 7-9, by right-clicking your HANA instance and selecting **Open SQL Console**.
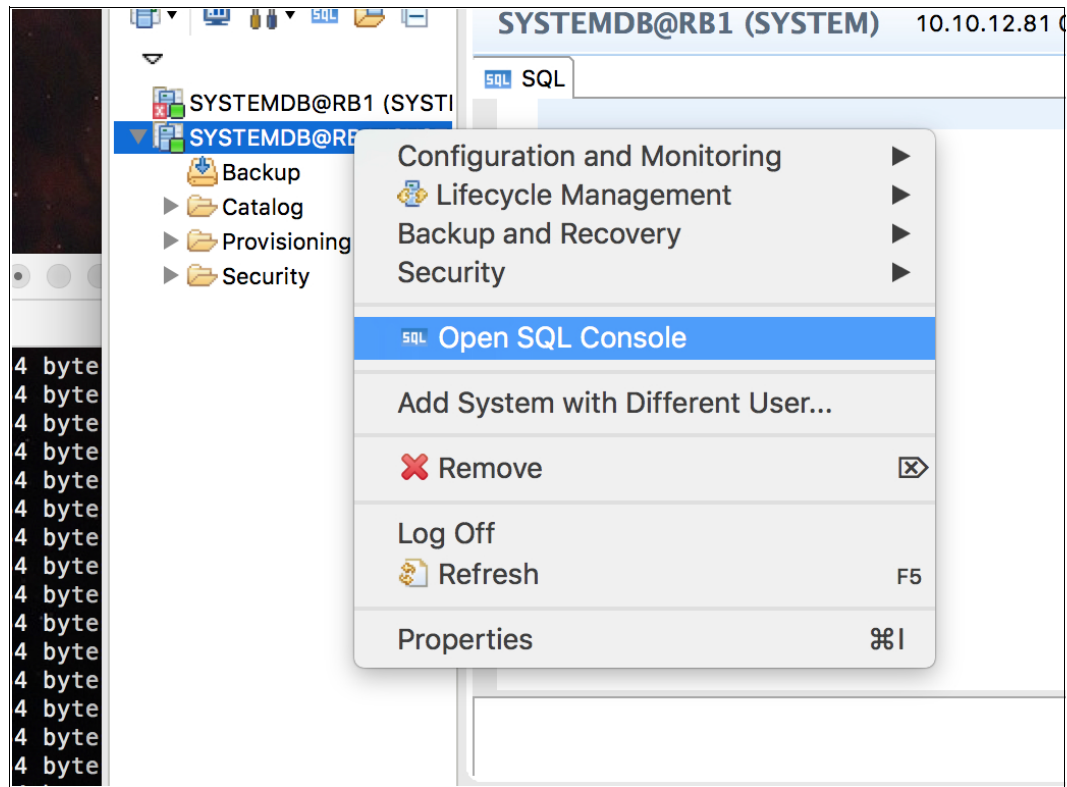
Figure 7-9   Opening the HANA instance SQL console

2. After the SQL console opens, use the SQL statements in Example 7-4 to create a table, populate it with one entry, and then query its contents. You can run one command at a time, or paste them all, one per line, and run them together. To run an SQL statement in the SQL console, press the F8 key or the `Execute` button at the upper right of the console.

Example 7-4   Creating some data with SQL statements

```
CREATE TABLE RESIDENCY (NAME CHAR(20), CITY CHAR (20) )

INSERT INTO RESIDENCY (NAME, CITY) VALUES ('hana', 'montana')

SELECT * FROM RESIDENCY
```

After running the three commands, the result of the **SELECT** statement displays an output similar to the one that is shown in Figure 7-10.
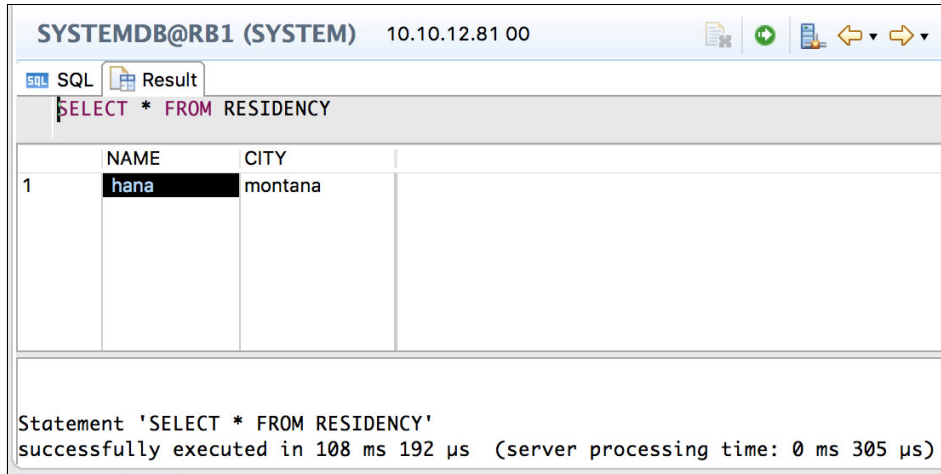
Figure 7-10   Querying a created entry by using SQL statements

7.3.2 Performing a takeover

Now that you have some data to validate a takeover, shut down the source system and initiate the takeover by right-clicking the *destination* system and selecting **Configure System Replication**, as shown in Figure 7-3 on page 111. Click **Perform takeover**, as shown in Figure 7-4 on page 111.

You do not have to change any of the parameters. The host and instance numbers are automatically input. Click **Finish** to proceed, as shown in Figure 7-11.
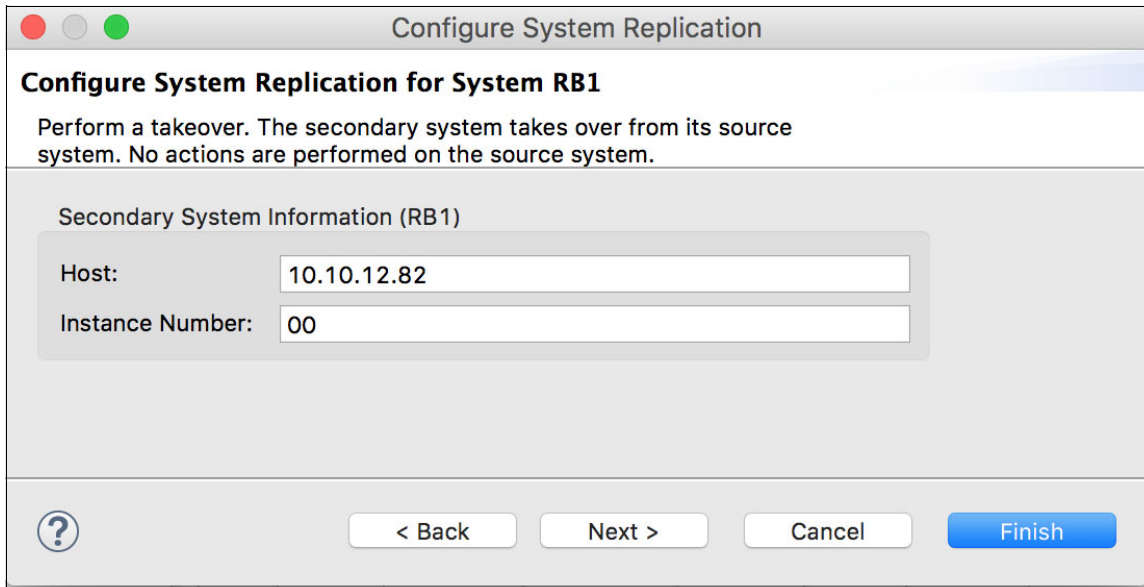
Figure 7-11   SAP HANA System Replication: Takeover to the destination HANA instance

The takeover operation brings the destination system to its full active state, and you can now connect to its DB and perform queries. A takeover does not shut down the DB on the source system in case of a controlled test, so applications can still send operations to it. You can take precautions to avoid applications from continuing to use the DB still running on the source system.

We encourage you to open the destination system SQL console, as explained in 7.3.1, "Creating a test table and populating it" on page 116, and run the **SELECT** statement from Example 7-4 on page 116. If you can see the created entry in the test table in your destination system, this means that the replication is working properly.

8

SAP HANA and IBM PowerHA SystemMirror

This chapter describes the integration of SAP HANA System Replication (HSR) with IBM PowerHA SystemMirror in a Red Hat Enterprise Linux V7.4 system.

This chapter covers only scenarios with two systems with HSR.

This chapter covers the following topics:

► Introduction
► Installing PowerHA SystemMirror
► Creating the PowerHA SystemMirror cluster
► Starting PowerHA SystemMirror
► Moving resources between nodes
► Closing notes

8.1 Introduction

HSR provides disaster recovery (DR) capabilities by having a copy of the active HANA database (DB) on a remote system. DR is a complex topic and is not covered by HSR or the HANA systems alone. DR includes external dependencies that cannot be ignored if an end-to-end DR solution is to be achieved. The following dependencies are not an exhaustive list:

► Client networks
► Application servers networks, storage
► Domain name servers (DNS)
► Routing to and from all clients on all sites
► Storage location and logical setup
► Internet Protocol (IP) domains
► Physical location of storage
► Internet Service Provider (ISP) connectivity on multiple sites
► Data center facilities

Enabling DR is a business decision and not a technical one. However, if you want to automate the HSR part to obtain a certain degree of high availability (HA) on scale-up systems, you can achieve this HA by adding an operating system (OS) HA solution. In this chapter, we cover the Red Hat OS with PowerHA SystemMirror as the HA layer to manage HSR on two HANA servers.

This section guides you on how to set up PowerHA SystemMirror on two HANA systems that are physically at two different failure domains. Those systems are running Red Hat Enterprise Linux for SAP Solutions Version V7.4, PowerHA SystemMirror for Linux V7.2.2.2, and SAP HANA V2.0.

This publication uses Red Hat Enterprise Linux V7.4. To check that you have a supported combination of Red Hat Enterprise Linux and HANA, see SAP Note 2235581 and Red Hat Customer Portal.

We recommend to always check the PowerHA SystemMirror for Linux IBM Knowledge Center before planning and setting up the cluster. Also, apply all the appropriate fixes to the OS and PowerHA SystemMirror before proceeding.

8.2 Installing PowerHA SystemMirror

Before installing PowerHA SystemMirror, a few tasks must be completed:

► HSR is configured.

► HSR can manually fail over and fall back, and you tested it.

► IP name resolution is working.

► There is one virtual IP address (VIPA), which is the IP that holds the active HANA service.

► IBM Linux on Power Tools is installed.

► You obtained the PowerHA SystemMirror software and all applicable fixes and e-fixes, if any.

> **Note:** Because HA is a complex topic and software clustering is not trivial, this chapter is intended to be a starting point to show you how PowerHA SystemMirror and HANA can interact together. It is not the intention of the authors that users in the field use this chapter alone without knowing the HA infrastructures that are part of the whole solution, or have expertise with PowerHA SystemMirror and Linux.

Complete the following steps:

1. Check whether all the requirements are fulfilled for PowerHA SystemMirror by running **`installPHA --onlyprereqcheck`**, as shown in Example 8-1. This parameter must be run on both Linux nodes, so repeat this operation on both nodes.

   Example 8-1   Checking the PowerHA SystemMirror prerequisites

   ```
   # ./installPHA --onlyprereqcheck
    checking :Prerequisites packages for PowerHA SystemMirror
   installPHA: Error: Prerequisite checking for the PHA installation failed:  RHEL
   7.4 ppc64le
   installPHA: One or more required packages are not installed: perl-Sys-Syslog
   (), perl-Pod-Parser (), sg3_utils (ppc64le)
   ```

2. To fix any missing prerequisites, run a **`yum`** command and rerun the precheck command in Example 8-1 to see whether there are any missing prerequisites, as shown in Example 8-2.

   Example 8-2   Installing the missing prerequisites of PowerHA SystemMirror and checking them

   ```
   # yum -y install perl-Sys-Syslog perl-Pod-Parser sg3_utils

   # ./installPHA --onlyprereqcheck
    checking :Prerequisites packages for PowerHA SystemMirror
    Success: All prerequisites of PowerHA SystemMirror installed
   installPHA: No installation only prerequisite check was performed .
   ```

3. After you successfully pass the prerequisite check on both nodes, continue with the installation. Then, on each node, run the **`installPHA`** command, as shown on Example 8-3.

   Example 8-3   Installing PowerHA SystemMirror

   ```
   # ./installPHA

   ...
   SNIP
   ...

   installPHA: Status of PHA after installation:

   Subsystem         Group         PID     Status
    ctrmc            rsct          10710   active
    IBM.ServiceRM    rsct_rm       10845   active
    IBM.DRM          rsct_rm       10851   active
    IBM.HostRM       rsct_rm       10912   active
    IBM.MgmtDomainRM rsct_rm       10944   active
    clcomd           clcomd        10978   active


   installPHA: All packages were installed successfully.
   ```

If you do not see the message `All packages were installed successfully` on both nodes, stop and contact support.

The PowerHA SystemMirror software is now installed and you are ready to configure the resources.

8.3 Creating the PowerHA SystemMirror cluster

Complete the following steps:

1. Set up the primary IP addresses of each node so that they can communicate with each other by running **Berkeley r-**. To do so, add the IP addresses to each node's `rhost` file, as shown in Example 8-4 (add your own IP addresses).

Example 8-4 Adding IP addresses to the rhost file

```
# echo 10.153.164.131 >>/etc/cluster/rhosts
# echo 10.153.164.136 >>/etc/cluster/rhosts
# refresh -s clcomd
0513-095 The request for subsystem refresh was completed successfully.
```

2. After you add the IP addresses and refresh the **clcomd** service on both nodes, you can create the cluster. To create the cluster, run the commands on only one node, which is the primary node that holds the HANA DB in normal operations. The PowerHA command can be run on any node of the cluster. In this example, run the **clmgr** command to create the cluster with nodes `ph13na1` and `ph13nb1`, as shown in Example 8-5.

Example 8-5 Creating a PowerHA SystemMirror cluster

```
# clmgr add cluster HDBH13 NODES=ph13na1,ph13nb1

"ph13na1" discovered a new node. Hostname is ph13na1.isicc.de.ibm.com. Adding it to the
configuration with Nodename "ph13na1".

"ph13nb1" discovered a new node. Hostname is ph13nb1.isicc.de.ibm.com. Adding it to the
configuration with Nodename "ph13nb1".


Attempting to create the cluster with following nodes:
ph13na1
ph13nb1

Successfully created the cluster: HDBH13

Creating Default Network.............


Successfully created default network: net_ether_01 :: eth0:ph13na1,eth0:ph13nb1


Configuring Cluster...

Setting the Split Policy...

Successfully set the split policy to "None"
```

3. Run the **clmgr** command to check the information about the cluster, as shown in Example 8-6.

Example 8-6   Listing the PowerHA SystemMirror cluster information

```
# clmgr list cluster
CLUSTER_NAME="HDBH13"
CLUSTER_ID="1538129092"
STATE="OFFLINE"
VERSION="7.2.2.2"
OSNAME="Linux"
SPLIT_POLICY="None"
TIE_BREAKER=""
NFS_SERVER=""
NFS_LOCAL_MOUNT_POINT=""
NFS_SERVER_MOUNT_POINT=""
NFS_FILE_NAME=""
DISK_WWID=""
```

> **Note:** You can also see that there are some properties that sound interesting, such as `NFS_SERVER`, `DISK_WWID`, and `TIE_BREAKER`. We describe a few of them in this section. For more information, see IBM Knowledge Center.

4. Now that the cluster is created, create the resources that this cluster manages. In our case, we are interested in PowerHA SystemMirror managing HSR replication flows, and VIPA. We use PowerHA SmartAssist to configure the HANA resource in the PowerHA SystemMirror cluster. The information that is needed to set it up must be the same that was used when configuring HSR. To set up the HANA resources with SmartAssist, run **clmgr** on one node only, as shown in Example on page 124.

Example 8-7   Starting SmartAssist to manage SAP HANA System Replication on PowerHA SystemMirror

```
# clmgr setup smart_assist APPLICATION=SAP_HANA SID=H13 INSTANCE=HDB13
```

> **Note:** When using PowerHA SmartAssist, use the primary host names, *not the aliases or other names* that do not match the primary host name on the nodes.
>
> ► Primary names:
>
>   10.153.164.131 ph13na1.isicc.de.ibm.com ph13na1
>   10.153.164.136 ph13na2.isicc.de.ibm.com ph13na2
>
> ► Other names on the block:
>
>   10.153.164.151 hanapr.isicc.de.ibm.com hanapr
>   10.153.164.156 hanadr.isicc.de.ibm.com hanadr
>
> Because the host names at the OS level are `ph13na1` and `ph13na2`, you must use them when configuring the nodes on the PowerHA SmartAssist, not `hanapr` and `hanadr`.
>
> The PowerHA SystemMirror support plan is for a scenario where both the nodes have separate host names. The primary host name (for example, `ph13na1` or `ph13na2`) is used to create the PowerHA SystemMirror cluster, but a different host name (for example, `hanapr` or `hanadr`) is used to install and configure HANA on both nodes. This support by PowerHA SystemMirror is intended to be delivered in the 7.2.3 release.

For more information about the parameters' format and values, see IBM Knowledge Center.

5. The command in Example produces a table that must be completed before you create the resources. The menu entries change version to version. For the version in this example, we populated them manually with the help of SmartAssist, which tries to obtain most of the values. You can see the input table that we had before running the deploy command, as shown in Example 8-8.

Example 8-8   HANA SmartAssist input data

```
PHA System Mirror Policy Setup Wizard

Policy: SAP HANA System Replication HA policy

Overall parameter status: OK

--------------------


                        Parameter overview
 # Parameter                                              Value

--------------------

1   Enter the name of your peer domain cluster.           OK ( HDBH13 )
2   Enter the hostname of nodes where you want to automate OK ( ph13na1,ph13nb1 )
    SAP HANA.
3   Specify the virtual IPv4 address that clients will use OK ( 10.153.164.138)
    to connect to SAP HANA.
4   Specify the netmask for the SAP HANA virtual IP address. OK ( 255.255.255.0 )
5   Enter the network interface for the SAP HANA IP address. OK ( eth0 )
6   Specify all site names of your SAP HANA nodes.        OK ( MONTANA,CYRUS )
7   Select the log replication mode for SAP HANA System   OK ( async )
    Replication.

Select the number of the parameter to start with
or one of the following options:

 0    Finish
 X    Cancel
```

When you concur that the values are correct, select 0 (the number) to continue to deploy the resources in the cluster. The output is similar to the one that is shown in Example 8-9.

Example 8-9   PowerHA SystemMirror deployment

```
Do you want to activate the policy now?

 1 Yes, activate as new policy
 2 Yes, activate by updating currently active policy
 3 No, save modifications and exit
 4 No, return to parameter overview


-Policy has been verified.


PHA1001I: The specified policy is valid.
```

EXPLANATION: The policy is valid and can be activated.
USER ACTION: No action required.
Are you sure you want to activate a new automation policy?

Yes(y)or No(n)?

```
/********************************Deleting Resources ******************************************/
Network:[net_ether_01]                                                    [SUCCESS]
Network:[powerha_hb]                                                      [SUCCESS]
/*******************************************************************************************/
/********************************Creating Resources ******************************************/
Application:[SAP_HDB_HDB13_sapstartsrv]                                   [SUCCESS]
Application:[SAP_HDB_HDB13_sr_primary_hdb]                                [SUCCESS]
Application:[SAP_HDB_HDB13_sr_secondary_hdb]                              [SUCCESS]
Equivalency:[SAP_HDB_HDB13_NETIF:ph13na1]                                 [SUCCESS]
Equivalency:[SAP_HDB_HDB13_NETIF:ph13nb1]                                 [SUCCESS]
Service-IP:[SAP_HDB_HDB13_sr_primary_ip]                                  [SUCCESS]
ResourceGroup:[SAP_HDB_HDB13_sapstartsrv_rg]                              [SUCCESS]
ResourceGroup:[SAP_HDB_HDB13_sr_primary_rg]                              [SUCCESS]
ResourceGroup:[SAP_HDB_HDB13_sr_secondary_rg]                            [SUCCESS]
IBM.Test:[SAP_HDB_HDB13_sr_hdb:ph13na1]                                   [SUCCESS]
IBM.Test:[SAP_HDB_HDB13_sr_hdb:ph13nb1]                                   [SUCCESS]
IBM.Test:[SAP_HDB_HDB13_sr_primary_configured:ph13na1]                    [SUCCESS]
IBM.Test:[SAP_HDB_HDB13_sr_primary_configured:ph13nb1]                    [SUCCESS]
IBM.Test:[SAP_HDB_HDB13_sr_primary_last_online:ph13na1]                   [SUCCESS]
IBM.Test:[SAP_HDB_HDB13_sr_primary_last_online:ph13nb1]                   [SUCCESS]
IBM.Test:[SAP_HDB_HDB13_sr_primary_replication_active:ph13na1]            [SUCCESS]
IBM.Test:[SAP_HDB_HDB13_sr_primary_replication_active:ph13nb1]            [SUCCESS]
IBM.Test:[SAP_HDB_HDB13_sr_primary_replication_syncing:ph13na1]           [SUCCESS]
IBM.Test:[SAP_HDB_HDB13_sr_primary_replication_syncing:ph13nb1]           [SUCCESS]
IBM.Test:[SAP_HDB_HDB13_sr_secondary_configured:ph13na1]                  [SUCCESS]
IBM.Test:[SAP_HDB_HDB13_sr_secondary_configured:ph13nb1]                  [SUCCESS]
IBM.Test:[SAP_HDB_HDB13_sr_secondary_register_in_progress:ph13na1]        [SUCCESS]
IBM.Test:[SAP_HDB_HDB13_sr_secondary_register_in_progress:ph13nb1]        [SUCCESS]
IBM.Test:[SAP_HDB_HDB13_sr_secondary_takeover_in_progress:ph13na1]        [SUCCESS]
IBM.Test:[SAP_HDB_HDB13_sr_secondary_takeover_in_progress:ph13nb1]        [SUCCESS]
Equivalency Dashboard                                                     [SUCCESS]
Dependency:[StartAfter:SAP_HDB_HDB13_s_StartAfter_p]                      [SUCCESS]
Dependency:[StartAfter:SAP_HDB_HDB13_primary_hdb_StartAfter_sapstartsrv]  [SUCCESS]
Dependency:[StartAfter:SAP_HDB_HDB13_secondary_hdb_StartAfter_sapstartsrv] [SUCCESS]
Dependency:[StopAfter:SAP_HDB_HDB13_sapstartsrv_StopAfter_hdb_primary]    [SUCCESS]
Dependency:[StopAfter:SAP_HDB_HDB13_sapstartsrv_StopAfter_hdb_secondary]  [SUCCESS]
Dependency:[StartAfter:SAP_HDB_HDB13_ip_StartAfter_primary]               [SUCCESS]
Dependency:[ForcedDownBy:SAP_HDB_HDB13_ip_ForcedDownBy_primary]           [SUCCESS]
/*******************************************************************************************/
```

> **Note:** You should understand the difference between options 1 and 2:
>
> ► `1 Yes, activate as new policy`
> ► `2 Yes, activate by updating currently active policy`
>
> If there are some existing PowerHA SystemMirror resources (for example, Persistent IP or disk heartbeat) and the `Activate` option is used, then those existing resources are deleted by SmartAssist.
>
> Option 1 is the same as the **clmgr add smart_assist** command, and option 2 is the same as the **clmgr update smart_assist** command.
>
> To learn more about the meanings of the corresponding options of the **clmgr** command, see IBM Knowledge Center.

You now have a simple PowerHA SystemMirror cluster that manages HSR.

> **Note:** If you must start over, you can clean up all the resources that were created by PowerHA SmartAssist by running **clmgr** (**clmgr delete smart_assist** with the relevant values of the SAP ID (SID), instance, and so on).

All communication to decide which node survives is done by using Ethernet. How resilient is this communication network, and what is the chance of failing depends on many factors, including the OS or virtualization configuration, switches, cabling, ISP, and many other items. However, for a more robust solution you must add other communication technologies. This section shows how you can add one shared disk to help decide which site will be the surviving site to avoid split-brain scenarios.

It is important to know how the shared disk is configured in the storage subsystem. If it is not resilient at either site or does not have a proper storage quorum, the situation can be worst. Detailed planning for end-to-end robust HA is needed.

In our case, the LUN that is used for the tiebreaker is provided by an IBM Spectrum Virtualize Enhanced Stretched cluster. The LUN is mirrored on both sites by an IBM Storwize® 7000 system and has a quorum on a third site that both sites can reach independently. For details about this setup, see *Implementing the IBM System Storage SAN Volume Controller with IBM Spectrum Virtualize V8.2.1*, SG24-7933.

The shared disk must be visible on both systems. You can make it visible to one node and give it a physical volume UUID, as shown in Example 8-10.

Example 8-10   Formatting a physical volume UUID on to a shared LUN

```
# pvcreate /dev/mapper/360050768019085e54800000000000227
  Physical volume "/dev/mapper/360050768019085e54800000000000227" successfully
created.

# pvdisplay /dev/mapper/360050768019085e54800000000000227
  "/dev/mapper/360050768019085e54800000000000227" is a new physical volume of
"4.00 GiB"
  --- NEW Physical volume ---
  PV Name               /dev/mapper/360050768019085e54800000000000227
  VG Name
  PV Size               4.00 GiB
  Allocatable           NO
```

```
PE Size                    0
Total PE                   0
Free PE                    0
Allocated PE               0
PV UUID                    FU2ZSI-LROe-S72J-Ji2N-sTlw-WiW2-fcPNds
```

Now you can present the LUN to the other node and ensure that the data is the same when querying with the **pvdisplay** command, as shown in Example 8-10 on page 126.

When you have the same physical volume UUID on both nodes, you can define it as a tiebreaker disk and query the data on PowerHA SystemMirror on one node only, as shown on Example 8-11.

> **Note:** The command that is shown in Example 8-11 creates a disk-based heartbeat network in addition to the network heartbeat that is enabled by default. There is a different command to create a tiebreaker.

Example 8-11   Defining a tiebreaker LUN in the PowerHA SystemMirror system

```
# clmgr add network powerha_hb TYPE=disk
pvid=FU2ZSI-LROe-S72J-Ji2N-sTlw-WiW2-fcPNds nodes=ph13na1,ph13nb1

 Successfully created DISK heartbeat network powerha_hb.

# clmgr query network powerha_hb
NAME="powerha_hb"
TYPE="disk"
INTERFACE_NODES="PVID=FU2ZSI-LROe-S72J-Ji2N-sTlw-WiW2-fcPNds:ph13na1
PVID=FU2ZSI-LROe-S72J-Ji2N-sTlw-WiW2-fcPNds:ph13nb1"
NETMASK=""
```

Now, you have a cluster that manages HSR and can decide the tiebreaker within the Ethernet and storage area network (SAN) network. The logical infrastructure overview is shown in Figure 8-1.
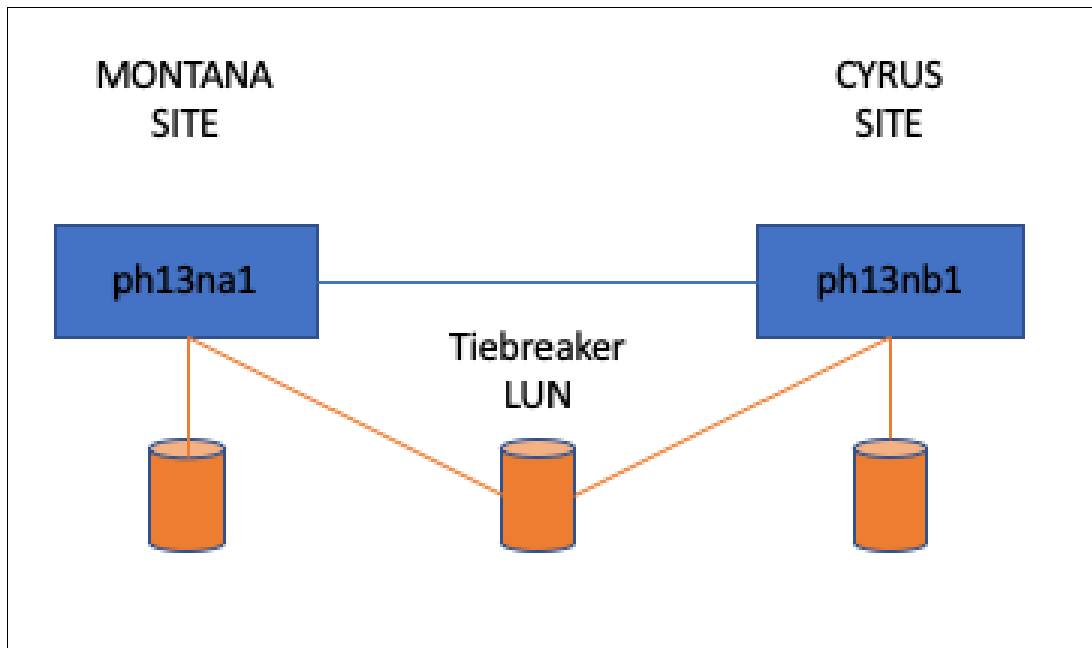

Figure 8-1   PowerHA SystemMirror logical infrastructure overview

With PowerHA SystemMirror, you can use a Network File System (NFS) as a tiebreaker, which is efficient for HANA because NFS is used to serve SAP interfaces. As with the LUN for tiebreaker setup, the NFS service must be highly available itself, and both sites must be able to access the NFS service independently.

Note: If there is an *even* number of nodes in a cluster, the user `must` configure a tiebreaker. When a heartbeat cannot be determined by both the network and disk, only one node continues. Otherwise, both nodes become *active*. As a best practice, configure the NFS tiebreaker. For more information, see the YouTube video PowerHA on Linux NFS Tie Breaker Operations.

8.4 Starting PowerHA SystemMirror

Start the cluster by running the `clmgr` command as shown in Example 8-12.

Example 8-12   Starting PowerHA SystemMirror

```
# clmgr online cluster

WARNING: MANAGE must be specified. Since it was not, a default of 'auto' will be used.


Cluster HDBH13 is running. We will try to bring the resource groups 'online' now, if any
exist.


Cluster services successfully started.
```

There are three applications that created by SmartAssist. You can show the applications and details about them by running the **clmgr** command as shown in Example 8-13.

Example 8-13   Listing applications and details about PowerHA SystemMirror

```
# clmgr list application
  SAP_HDB_H13_HDB13_sapstartsrv
  SAP_HDB_H13_HDB13_sr_primary_hdb
  SAP_HDB_H13_HDB13_sr_secondary_hdb

# clmgr view resource_group SAP_HDB_H13_HDB13_sr_primary_rg
NAME="SAP_HDB_H13_HDB13_sr_primary_rg"
CURRENT_NODE="ph13na1"
NODES="ph13na1 ph13nb1"
STATE="ONLINE"
TYPE="non-concurrent"
PRIORITY="0"
STARTUP=""
FALLOVER=""
FALLBACK=""
APPLICATIONS="SAP_HDB_H13_HDB13_sr_primary_hdb"
SERVICE_LABEL="SAP_HDB_H13_HDB13_sr_primary_ip"
FILESYSTEM=""
NESTED_RGS=""
```

To list all the resources and their status, run the **clRGinfo** command, as shown in Example 8-14.

Example 8-14   The clRGinfo command output

```
# clRGinfo


SAP_HDB_H13_HDB13_sapstartsrv_rg          ONLINE              ph13na1
                                ONLINE              ph13nb1


SAP_HDB_H13_HDB13_sr_primary_rg           ONLINE              ph13na1
                                INELIGIBLE          ph13nb1


SAP_HDB_H13_HDB13_sr_secondary_rg         INELIGIBLE          ph13na1
                                ONLINE              ph13nb1


```

> **Note:** You can view the detailed output of the resources status by running the following command:
>
> clRGinfo -e

8.5 Moving resources between nodes

Now, you can move the active HANA replica between nodes. PowerHA SystemMirror reverses the HSR flow from active to primary each time you move the resources, which also happens when the node crashes and the node comes back online again.

To move gracefully an application (resource), run the **clmgr** command, as shown in Example 8-15.

Example 8-15   Moving the primary SAP HANA System Replication replica to a second node

```
# clmgr move resource_group SAP_HDB_H13_HDB13_sr_primary_rg NODE=ph13nb1
```

The command that is shown in Example 8-15 stops HSR in `ph13na1`, which holds the primary role at this moment. The command makes the HSR replica the primary on `ph13nb1` and moves the VIPA from `ph13na1` to `ph13nb1`. If `ph13na1` remains active, the cluster configures the HSR replica with `ph13nb1` as the primary and `ph13na1` as the secondary.

Now, HSR is reversed. For more information about how to check the HSR status, see Chapter 7, "SAP HANA System Replication for high availability and disaster recovery scenarios" on page 105.

8.6 Closing notes

This chapter introduced PowerHA SystemMirror for Linux and HSR management by PowerHA SystemMirror. HA is a journey and this is the beginning of setting up a basic cluster. Each system is different because it has differences in the infrastructure (logical and physical) that make it different. To learn about and implement HA, see IBM Knowledge Center.

IBM Systems Lab Services has a program that is called *Power to Cloud Rewards* that specifically covers HANA and PowerHA SystemMirror. The program has no cost and includes an onsite workshop with the customer. To learn more on how to qualify for this program, see IBM Power to Cloud Rewards Program IBM Power to Cloud Rewards Program.

Even if you cannot qualify for the Power to Cloud Reward program, you can still obtain the expertise to design, deploy, and test your HANA and PowerHA SystemMirror with IBM Systems Lab Services. For more information about the details of the offers and contact information, see IBM Systems Lab Services.

SAP HANA and IBM VM Recovery Manager high availability and disaster recovery

This chapter describes IBM VM Recovery Manager high availability (HA) and disaster recovery (DR) availability solutions for SAP HANA running on IBM Power Systems servers.

This chapter covers the following topics:

► Business continuity and recovery orchestrator
► Power Systems HA and DR solutions for SAP HANA

9.1 Business continuity and recovery orchestrator

Business continuity is a part of any business operations. Many businesses have recovery plans if there are failures or disasters because downtime and disruptions can cause financial losses, bad public relations, and trust in the business.

A *recovery orchestrator* is the HA and DR software component (solution) that enables and manages the recovery of an IT infrastructure if there is an outage. A recovery orchestrator is easy to deploy and manage, and can perform repeatable recovery without disruptions.

IBM Power Systems offers a rich set of HA and DR solutions. A few more options were added recently, and this chapter provides an overview of the various solutions and which ones can be used to manage HA and DR for SAP HANA.

Clustering technologies play an important role regarding HA and DR operations. Cluster-based HA or DR solutions rely on redundant standby nodes in the cluster to take over the workloads and start them when the primary node fails. Each node in the cluster monitors the health of various elements, such as network interfaces, storage, and partner nodes, and to act when any of these elements fail. Clustering technologies are the closest to fault-tolerant environments regarding HA or DR support that is based on redundant software and hardware components. Clustering solutions are often operating system- (OS) or platform-specific, they provide detailed error monitoring, and they require effort to deploy and maintain.

A cluster DR model is shown in Figure 9-1 (the cluster DR model displays the DR solution, but it also applies to HA except that in the case of HA no replication is involved), which is contrasted with the virtual machine (VM) Restart DR model. Figure 9-1 shows that the entire VM (including system disk, rootvg, data disks, and so on) is replicated by using storage replication methods. These copies of VMs are used during a disaster to start the VMs on the DR site. OSes in these VMs start, and then the workload is started to return to normal operations. This model is more suited for cloud deployments and can scale to allow for DR management of the entire data center.

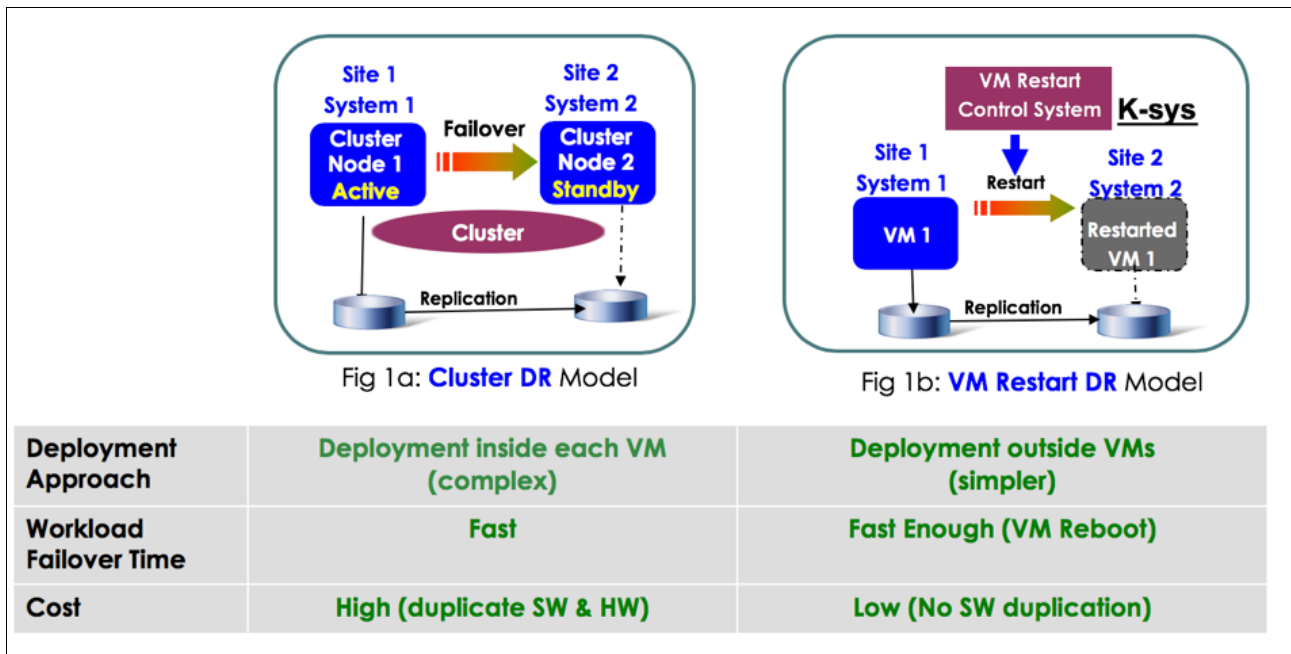|	Fig 1a: **Cluster DR** Model	Fig 1b: **VM Restart DR** Model
Deployment Approach	Deployment inside each VM (complex)	Deployment outside VMs (simpler)
Workload Failover Time	Fast	Fast Enough (VM Reboot)
Cost	High (duplicate SW & HW)	Low (No SW duplication)

Figure 9-1   Popular HA and DR models

VM Restart Manager -based HA involves a similar concept of restarting the VM or logical partition (LPAR) on some other host within a data center, and relies on a shared storage- (between hosts) based image to start the VM. Also, in the case of HA, VMs are restarted automatically as compared to manual restart in DR cases.

IBM Power Systems servers now offer both types of HA and DR solutions for the PowerVM platform, as shown in Figure 9-2.
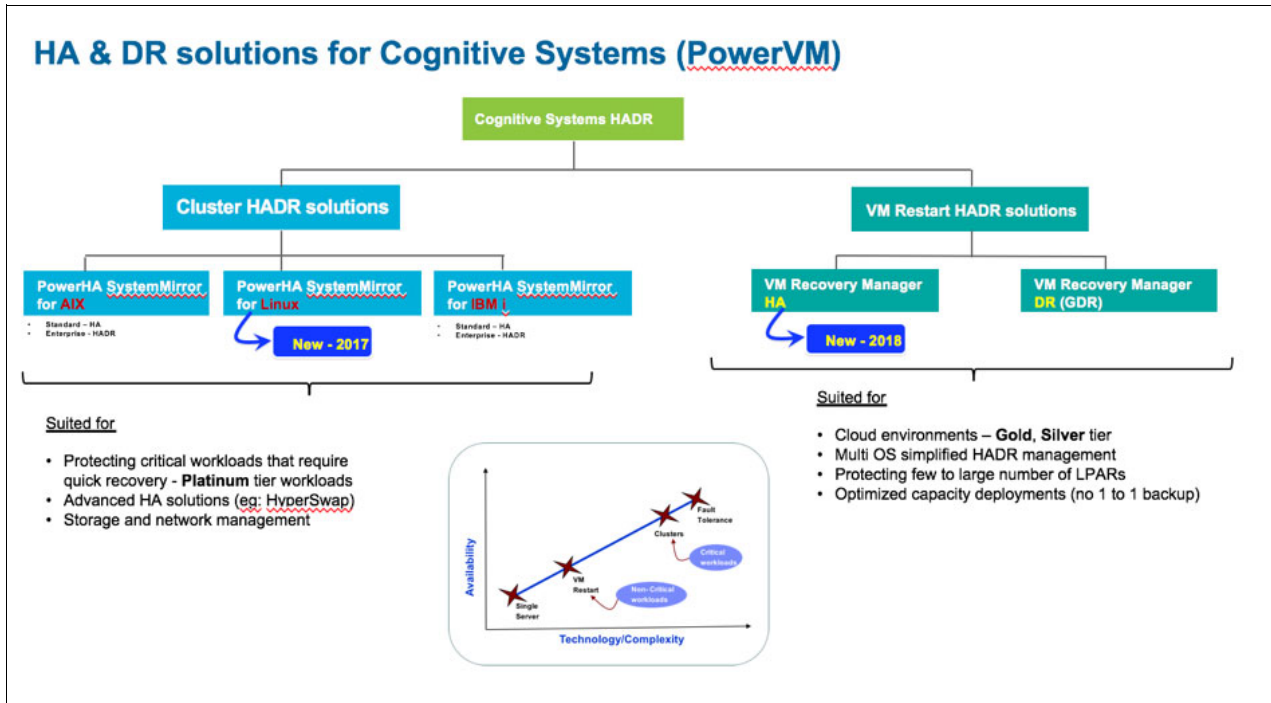

Figure 9-2   High availability and disaster recovery offerings from Power Systems servers

Here are some more solution details:

► Cluster HA and DR solutions:

  – PowerHA SystemMirror for AIX: A cluster-based HA and DR solution that has been deployed on Power Systems for a few decades.

  – PowerHA SystemMirror for Linux: Introduced in 2017, it can manage HA for SAP HANA. For more information, see 9.2, "Power Systems HA and DR solutions for SAP HANA" on page 134, and Chapter 8, "SAP HANA and IBM PowerHA SystemMirror" on page 119.

  – PowerHA SystemMirror for IBM i: Enables HA management for various workloads on IBM i.

► VM Restart Manager HA and DR solutions:

  – GDR: An easy to deploy and manage DR solution.

  – VM Recovery Manager HA: An OS-neutral and easy to deploy and use HA solution. You can use this solution to protect SAP HANA environments. For more information, see 9.2.3, "VM Recovery Manager HA: A VM Restart Manager -based HA solution for SAP HANA" on page 134.

9.2 Power Systems HA and DR solutions for SAP HANA

This section takes a closer look at the various Power Systems HA or DR solutions that support SAP HANA HA/DR management.

9.2.1 PowerHA SystemMirror for Linux: A cluster-based HA solution for SAP HANA

Beyond the built-in HA capabilities of SAP HANA, it is a best practice for customers to deploy a cluster-based HA solution to manage failovers and the proper restart of the software stack on the standby system.

You can use PowerHA SystemMirror for Linux to manage an HA for SAP HANA replication-based environment. It supports both HANA replication-based hot standby or cold restart-based HANA deployments. PowerHA provides a wizard to configure HA policies for an SAP HANA environment.

For more information, see Chapter 8, "SAP HANA and IBM PowerHA SystemMirror" on page 119 and IBM Knowledge Center.

9.2.2 IBM Geographically Dispersed Resiliency: A VM Restart Manager -based DR solution for SAP HANA

You can use IBM Geographically Dispersed Resiliency (GDR) to manage a storage replication-based DR environment. It is easy to deploy and manage the DR for the entire data center, including the SAP HANA environment.

For more information, see *Implementing High Availability and Disaster Recovery Solutions with SAP HANA on IBM Power Systems*, REDP-5443.

9.2.3 VM Recovery Manager HA: A VM Restart Manager -based HA solution for SAP HANA

A VM Recovery Manager HA solution for SAP HANA on Power Systems enables an easy to deploy and use HA solution:

► VM Restart Manager -based HA management protects against host, VM, and application failures.

► Graphically deploy and manage the HA environment.

► Practical cluster and VM Restart Manager HA management by using simplified application HA management.

► Planned HA management (Live Partition Mobility (LPM)): Vacate or restore a host with ease.

► HA agents for SAP HANA, NetWeaver, Oracle, and IBM Db2®.

► Advanced policies to control application starts, VM colocation, VM priority-based restart, and capacity adjustments.

VM Recovery Manager HA orchestrator KSYS (controller system LPAR) monitors the environment for host or VM or application failures and restarts the VM in some other host within the host group.

Figure 9-3 shows failure of a host and the resulting failover within the host group.
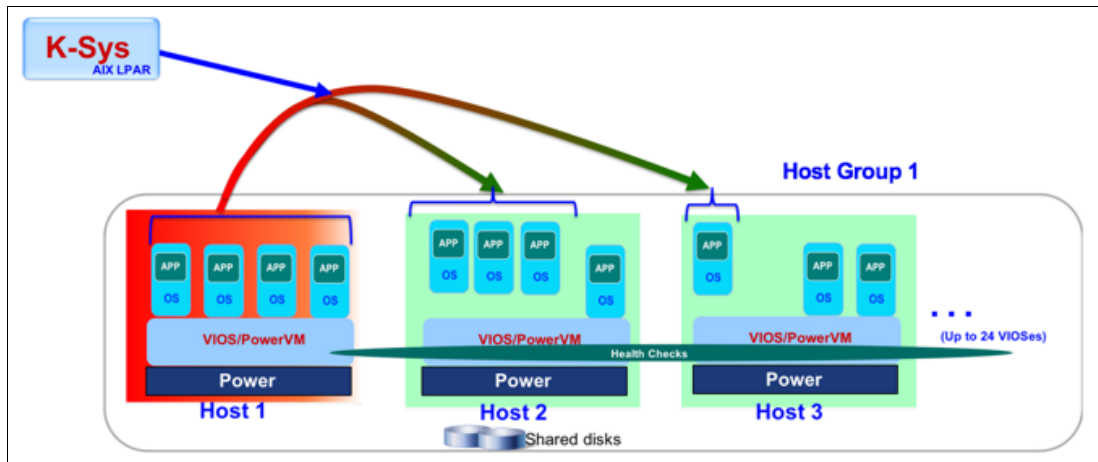

Figure 9-3   VM Recovery Manager HA management: Host failure scenario

An administrator can enable any of the monitoring options that are available:

► Host failure detections: This option detects failures of hosts and moves the VMs to the remaining hosts in the host group. This is the default detection that is done for VMs (applies to AIX, Linux, and IBM i LPARs on a PowerVM platform).

► VM failure detections: An administrator can optionally choose to detect failures of VMs. To enable this function, the administrator must install and initialize the VM agent component inside AIX or Linux LPARs VMs.

► Application failure detections: An administrator can optionally use the application monitoring (AppMon) framework of VM agent to register and monitor the health of the applications inside VMs.

VM Recovery Manager HA monitors for failures and takes corrective actions based on a few key components, as shown in Figure 9-4.
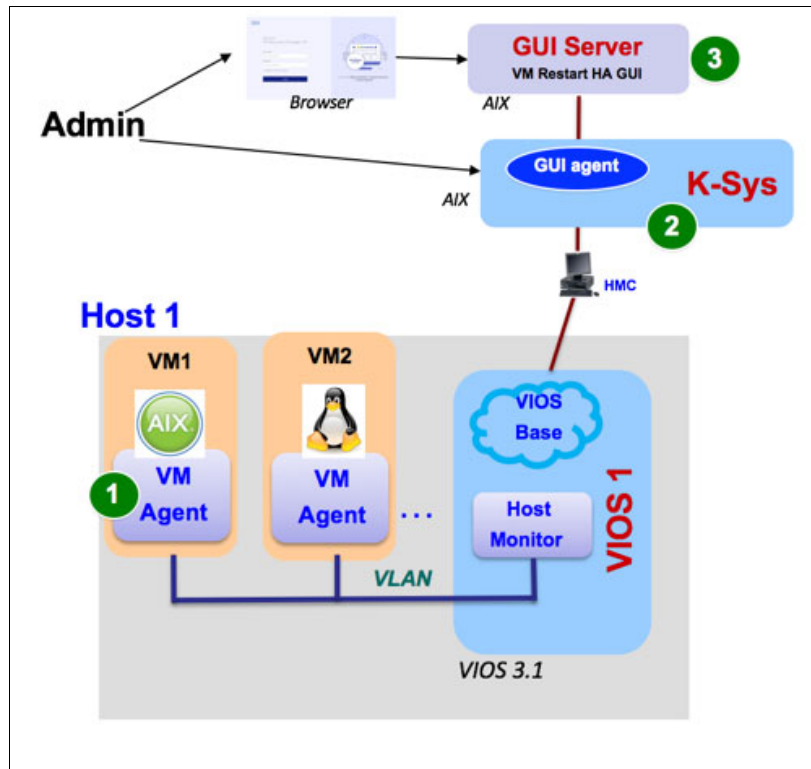

Figure 9-4   Components of VM Recovery Manager HA

The administrator must install three key components before initializing and using VM Recovery Manager HA:

► VM agent: The VM agent enables the VM health monitoring and AppMon framework. VM agent is provided for AIX and Linux. The administrator installs this component first and initializes it (if they plan to do VM or AppMon). The VM agent can be initialized and managed by using a single command inside the VM (called `ksysvmmgr`).

► KSYS: Install the KSYS software in an AIX LPAR. The KSYS software installation also deploys GUI-related agent software.

► GUI server: This software can be installed in the KSYS itself or in another AIX LPAR.

After installation, the administrator can start the browser and connect to `http://ksys_hostname:3000`. Use the KSYS LPAR login credentials to log in and then follow the instructions to deploy a host group and enable HA management.

For more information about the installation and configuration of VM Recovery Manager HA, see IBM Knowledge Center.

9.2.4 SAP HANA HA management by using VM Recovery Manager HA

VM Recovery Manager HA provides a lightweight AppMon framework as part of the VM agent. The VM agent can be installed by an administrator inside AIX or Linux LPARs. Power Systems servers support SAP HANA on Linux VMs on PowerVM.

Application monitoring framework

The AppMon framework inside VM enables simplified monitoring and management of applications running in the LPAR, as shown in Figure 9-5.
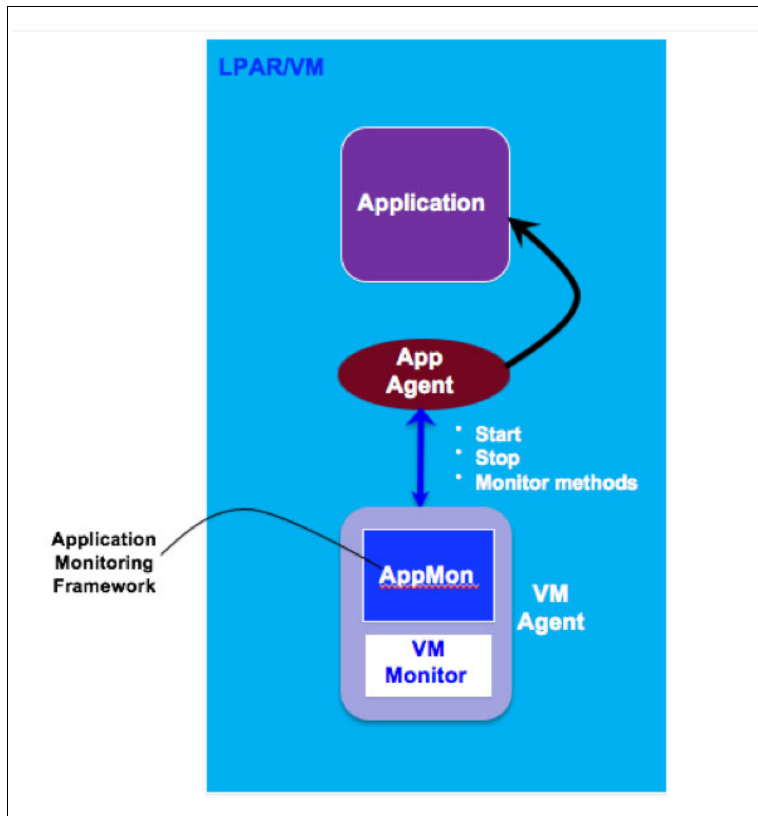

Figure 9-5   Application monitoring framework (VM Recovery Manager HA)

The VM administrator can register and monitor any applications inside the LPAR. To do this task, the administrator registers the application by running the `ksysvmmgr` command. As part of the registration, the administration provides three methods to manage the application:

► Start: This method allows AppMon to start the application.

► Stop: This method when invoked stops the application.

► Monitor: AppMon calls this method periodically (every 30 seconds) to monitor the health of the application. Based on the monitor return status, the application status is marked as green, yellow, or red, and the appropriate action is taken.

Health monitoring of an application and the resulting actions are shown in Figure 9-6.
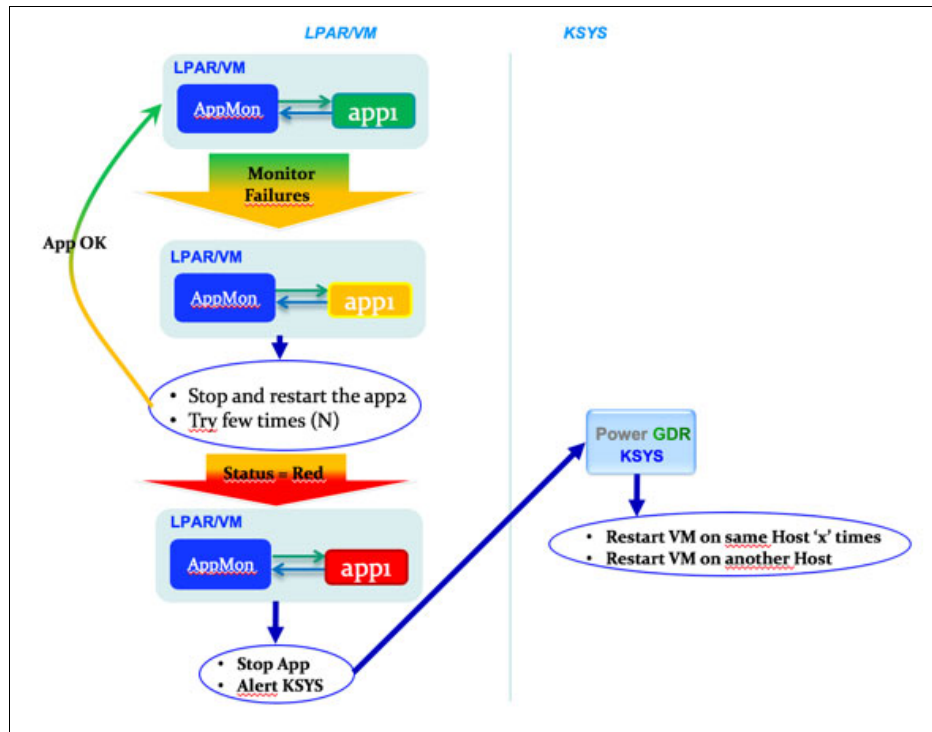

Figure 9-6   Application health monitoring (green →yellow →red status changes)

VM Recovery Manager HA provides HA agents for key middleware products. One of the agents that is supplied is to manage cold restart-based SAP HANA HA management. The next few sections describe the installation and HA management of SAP HANA by using this agent.

9.2.5 VM Recovery Manager HA: SAP HANA agent deployment and management

This section describes the VM Recovery Manager HA SAP HANA agent deployment and management.

Installing SAP HANA
The following steps describe the SAP HANA installation procedure:

1. Complete the following prerequisites:

   a. Install Linux Red Hat Enterprise Linux V7.4 or SUSE Linux Enterprise Server 12.2/3. For the list of supported OSes, see SAP HANA Launchpad.

   b. Extract the SAP HANA V2.0 setup files.

   c. Run `HANA2.0/DATA_UNITS/HDB_LCM_LINUX_PPC64LE/hdblcm`. For more information, see Chapter 6, "SAP HANA software stack installation for a scale-up scenario" on page 83.

d. Install all the required RPMs while running `./hdblcm`. The required RPMs for SUSE Linux Enterprise Server and Red Hat Enterprise Linux are as follows:

-   SUSE Linux Enterprise Server

    ```
    libgomp1-7.3.1+r258313-6.1.ppc64le.rpm
    libstdc++6-4.2.1-3mdv2008.0.ppc.rpm
    glibc-2.28-497.2.ppc64le.rpm
    IBM_XL_C_CPP_V13.1.5.1_LINUX_RUNTIME
    ```

-   Red Hat Enterprise Linux

    ```
    libxlc-13.1.6.1-171213.ppc64le.rpm
    compat-sap-c++-6-6.3.1-1.el7_3.ppc64le.rpm
    libtool-ltdl-2.4.6-25.fc29.ppc64le.rpm
    ```

2. Select all the default parameters (or defined values if you have any) except for SAP System ID, SAP Admin/User, and Instance number. You can use the default values for the SAP System ID, SAP Admin/User, and Instance number, but it is a best practice to provide meaningful and planned values for them because they are used when the SAP system is scaled up.

3. When prompted for passwords for system admin, SAP admin, and database (DB) user, provide the passwords, confirm them, and proceed by confirming the details by pressing *y* to proceed.

4. After waiting for approximately 15 - 20 minutes, you receive a notice about the installation's success.

5. To confirm that everything installed correctly, log in again to the host by using the newly created SAP Admin user, and then use the following commands to verify the installation:

    a. Check the SAP HANA status by running the following command:

    ```
    sapcontrol -nr 02 -function GetSystemInstanceList
    ```

    Where `02` is the instance number (change to the instance number that you provided during the installation). You see `OK` if everything is working correctly.

    > **Note:** When running these commands from a non-SAP user, use the prefix `/usr/sap/hostctrl/exe/` before the command because it cannot be identified otherwise.

    b. Stop SAP HANA by running the following command:

    ```
    sapcontrol -nr 02 -function StopSystem HDB
    ```

    Where `02` is the instance number (change to the instance number that you provided during the installation.). A confirmation report is shown.

    a. Start SAP HANA by running the following command:

    ```
    sapcontrol -nr 02 -function StartSystem HDB
    ```

    Where `02` is the instance number (change to the instance number that you provided during the installation.). A confirmation is shown. Check the status again.

6. The installation logs are stored in `/tmp` with the date of installation. For example:

    ```
    /tmp/hdb_S01_hdblcm_install_2018-04-20_02.11.21/hdblcm.log
    ```

7. To start `saphana`, run the following command:

```
CMD="${EXE_DIR}/sapcontrol -nr ${INSTANCE_NO} -function StartService ${SID}"
(without wait time)
CMD="${EXE_DIR}/sapcontrol -nr ${INSTANCE_NO} -function WaitforServiceStarted 5
0"(with wait time)
```

Where the following terms are defined as follows:

`EXE_DIR`	`/usr/sap/hostctrl/exe` (default)
`EXE_DIR`	`/home/hana/shared/${SAP_SID}/${INSTANCE_NAME}/exe` (installed path)
`EXE_DIR`	`/home/hana/shared/${SAP_SID}/exe/linuxppc64le/hdb`
`INSTANCE_NO`	`01` (configurable)
`SID/SAP_SID`	`S01` (configurable)
`INSTANCE_NAME`	`HBD01` (number same as `INSTANCE_NO`)

Configuring SAP HANA with VM Recovery Manager HA

VM Recovery Manager HA provides an agent (start, stop, and monitor methods for managing SAP HANA) for SAP HANA HA management. You can start the agent by running the following command:

```
ksysvmmgr add app sapapp <application_name> type=SAPHANA
instancename=S01<SAP_HANA_INSTANCE> database=HDB01<SAP_HANA_DATABASE_NUMBER>
```

The agent uses the default SAP HANA agent scripts that are in the `/usr/sbin/agents/sap` directory.

Alternatively, a user can use their own scripts by running the **ksysvmmgr** command:

```
ksysvmmgr add app sapapp <application_name> monitor_script=<monitor_script_path>
start_script=<start_script_path> stop_script=<stop_script_path>
```

SAP HANA HA agent internals

The following list describes the SAP HANA HA agent internals:

► Start script

The SAP HANA start script uses `type=SAPHANA username=S01 database=HDB01` to start the `saphana` application. SAP HANA takes around 2 minutes to start. The default start stabilization time is 150 seconds, which can be modified by using the **ksysvmmgr** command.

After the script starts, the servers that are listed in this command output start:

```
#  ps -aef | grep s01adm
s01adm    3111     1  0 Jun11 ?        00:02:38
/usr/sap/S01/HDB01/exe/sapstartsrv
pf=/usr/sap/S01/SYS/profile/S01_HDB01_bolts021.ausprv.stglabs.ibm.com -D -u
s01adm
root      3741 48831  0 02:54 pts/0    00:00:00 grep --color=auto s01adm
s01adm   50357     1  0 01:10 ?        00:00:00 sapstart
pf=/usr/sap/S01/SYS/profile/S01_HDB01_bolts021.ausprv.stglabs.ibm.com
s01adm   50365 50357  0 01:10 ?        00:00:00
/usr/sap/S01/HDB01/bolts021.ausprv.stglabs.ibm.com/trace/hdb.sapS01_HDB01 -d
-nw -f /usr/sap/S01/HDB01/bolts021.ausprv.stglabs.ibm.com/daemon.ini
pf=/usr/sap/S01/SYS/profile/S01_HDB01_bolts021.ausprv.stglabs.ibm.com
```

The following processes start:

```
s01adm   50381 50365  4 01:10 ?          00:04:56 hdbnameserver
s01adm   50531 50365  0 01:11 ?          00:00:42 hdbcompileserver
s01adm   50533 50365  0 01:11 ?          00:00:41 hdbpreprocessor
s01adm   50577 50365  5 01:11 ?          00:05:43 hdbindexserver -port 30103
s01adm   50579 50365  1 01:11 ?          00:01:11 hdbxsengine -port 30107
s01adm   51201 50365  0 01:11 ?          00:00:43 hdbwebdispatcher
```

► Stop script

The SAP HANA stop script uses `type=SAPHANA username=S01 database=HDB01` to stop the SAP HANA instance. It takes about 1 minute to stop and end all SAP processes. In this example, we use 100 seconds (the default) for the SAP HANA stop script run.

After the stop script run finishes, the stop script shuts down SAP processes by making them unavailable, as shown in the following output:

```
#  ps -aef | grep s01adm
s01adm 3111 1 0 Jun11 ?           00:02:38 /usr/sap/S01/HDB01/exe/sapstartsrv
pf=/usr/sap/S01/SYS/profile/S01_HDB01_bolts021.ausprv.stglabs.ibm.com -D -u
s01adm
s01adm 4216 1  0 02:55 ?          00:00:00 hdbrsutil -f -D -p 30101 -i 1537167311
s01adm 4249 1  0 02:55 ?          00:00:00 hdbrsutil -f -D -p 30103 -i 1537167313
root 4455 48831  0 02:56 pts/0    00:00:00 grep --color=auto s01adm
```

► Monitor script

The monitor script checks the status of the SAP HANA instance. If the process returns a green online status, then the check is a success; any other returned status is a failure.

The start and stop stabilization time must be more than the time that is required for the script to run, that is, the start stabilization time must be more than 150 seconds, and the stop stabilization time must be more than 100 seconds.

The SAP HANA monitor script checks the SAP HANA instance and returns the running process list and the processes' status. The processes are:

– **hdbdaemon**
– **hdbcompileserver**
– **hdbindexserver**
– **hdbnameserver**
– **hdbpreprocessor**
– **hdbxsengine**

The application status can be verified by checking the process information by running the **ps -aef** command as the `sapadmin` user. The following output shows this information:

```
#  ps -aef | grep s01adm
s01adm    3111     1 0 Jun11 ?          00:02:38
/usr/sap/S01/HDB01/exe/sapstartsrv
pf=/usr/sap/S01/SYS/profile/S01_HDB01_bolts021.ausprv.stglabs.ibm.com -D -u
s01adm
root      3741 48831 0 02:54 pts/0    00:00:00 grep --color=auto s01adm
s01adm   50357     1 0 01:10 ?          00:00:00 sapstart
pf=/usr/sap/S01/SYS/profile/S01_HDB01_bolts021.ausprv.stglabs.ibm.com
s01adm   50365 50357 0 01:10 ?          00:00:00
/usr/sap/S01/HDB01/bolts021.ausprv.stglabs.ibm.com/trace/hdb.sapS01_HDB01 -d
-nw -f /usr/sap/S01/HDB01/bolts021.ausprv.stglabs.ibm.com/daemon.ini
pf=/usr/sap/S01/SYS/profile/S01_HDB01_bolts021.ausprv.stglabs.ibm.com
s01adm   50381 50365  4 01:10 ?          00:04:56 hdbnameserver
s01adm   50531 50365  0 01:11 ?          00:00:42 hdbcompileserver
```

```
s01adm     50533 50365  0 01:11 ?          00:00:41 hdbpreprocessor
s01adm     50577 50365  5 01:11 ?          00:05:43 hdbindexserver -port 30103
s01adm     50579 50365  1 01:11 ?          00:01:11 hdbxsengine -port 30107
s01adm     51201 50365  0 01:11 ?          00:00:43 hdbwebdispatcher
```

VM agent behavior with the SAP HANA application

This section describes the VM agent behavior with the SAP HANA application. It follows the following chain of events:

1. The monitor script runs after the **ksysvmmgr add app** and **sync** functions check whether the application is already started, and then the script checks the state and updates. The state is `FAILING`.

2. The stop script runs while the state is failing so that all the residual configurations are removed and the application can be started normally. The state is `TO STOP`.

3. If the stop script returns a failure, the VM agent reruns the stop script until the maximum failure times threshold is reached (3 is the default) before moving to permanent failure. the state is `ABNORMAL/PERMANENT FAILURE`.

4. If the stop script returns a success, the start script is triggered after the stop stabilization time is reached. The application waits until the start stabilization time is reached (30 seconds is the default) to start the application. The state is `TO START`.

5. The monitor script is triggered after the start script runs successfully. The state is `NORMAL`.

SAP HANA logs

Here are the SAP HANA logs that are provided by the VM agent:

► `LOG :: /usr/sbin/agents/saphana/monitorsaphana :: Tue Apr 24 01:45:37 EDT 2018 :: check status of sap hana instance.` The monitor script is called to check the SAP HANA status.

► `LOG :: /usr/sbin/agents/sap/sapstartctrl :: Tue Apr 24 01:45:38 EDT 2018 :: Enter function Control_instance() Check_cmd.\n`

► `LOG :: /usr/sbin/agents/sap/sapstartctrl :: Tue Apr 24 01:45:38 EDT 2018 :: Control_instance sapcontrol GetProcessList rc=4:`

```
24.04.2018 01:45:38
GetProcessList
OK
0 name: hdbdaemon
0 description: HDB Daemon
0 dispstatus: GRAY
0 textstatus: Stopped    → The SAP HANA status is not normal.
0 starttime:
0 elapsedtime:
0 pid: 60964
```

► `LOG :: /usr/sbin/agents/sap/sapstartctrl :: Tue Apr 24 01:45:38 EDT 2018 :: The HDB process is stopped.`

► `LOG :: /usr/sbin/agents/saphana/stopsaphana :: Tue Apr 24 01:47:41 EDT 2018 :: SAP_HANA is not running.` The stop script is called to stop or clean up SAP HANA.

► `LOG :: /usr/sbin/agents/saphana/stopsaphana :: Tue Apr 24 01:47:41 EDT 2018 :: sap hana instance is already stopped.`

- ► LOG :: /usr/sbin/agents/saphana/startsaphana :: Tue Apr 24 01:47:43 EDT 2018 :: SAP_HANA is not running. Starting SAP hana The start script is called to start SAP HANA.

- ► LOG :: /usr/sbin/agents/saphana/startsaphana :: Tue Apr 24 01:47:43 EDT 2018 :: Calling doStart()...

- ► LOG :: /usr/sbin/agents/sap/saphdbctrl :: Tue Apr 24 01:47:43 EDT 2018 :: saphdb_ci_start start issued.

- ► LOG :: /usr/sbin/agents/sap/sapsrvctrl :: Tue Apr 24 01:47:43 EDT 2018 :: Enter function Control_sapstartsrv() Start.\n

- ► LOG :: /usr/sbin/agents/sap/sapsrvctrl :: Tue Apr 24 01:47:43 EDT 2018 :: Control_sapstartsrv sapcontrol StartService rc=0:

  StartService
  OK .\n

- ► LOG :: /usr/sbin/agents/sap/saphdbctrl :: Tue Apr 24 01:47:43 EDT 2018 :: saphdb_ci_start: sapsrvctrl -a start -p HDB S01_HDB01 OpState=0.

- ► LOG :: /usr/sbin/agents/sap/sapstartctrl :: Tue Apr 24 01:47:43 EDT 2018 :: Enter function Control_instance() Start_cmd.\n

- ► LOG :: /usr/sbin/agents/sap/sapstartctrl :: Tue Apr 24 01:47:43 EDT 2018 :: Control_instance sapcontrol Start rc=0:

  24.04.2018 01:47:43
  Start
  OK

- ► LOG :: /usr/sbin/agents/sap/sapstartctrl :: Tue Apr 24 01:48:55 EDT 2018 :: Executed '/bin/su - s01adm -c /home/hana/shared/S01/exe/linuxppc64le/hdb/sapcontrol -host localhost -nr 01 -function WaitforStarted 120 1' returncode: 0

- ► LOG :: /usr/sbin/agents/sap/sapstartctrl :: Tue Apr 24 01:48:55 EDT 2018 :: Start instance returned with a returncode of 0.

- ► LOG :: /usr/sbin/agents/sap/sapstartctrl :: Tue Apr 24 01:48:55 EDT 2018 :: Start completed successfully.

- ► LOG :: /usr/sbin/agents/sap/saphdbctrl :: Tue Apr 24 01:48:55 EDT 2018 :: SAP HDB Start done. rc:0

- ► LOG :: /usr/sbin/agents/saphana/startsaphana :: Tue Apr 24 01:48:55 EDT 2018 :: sap hana instance started !!

- ► LOG :: /usr/sbin/agents/saphana/startsaphana :: Tue Apr 24 01:48:55 EDT 2018 :: sap hana instance started successfully. SAP HANA started successfully.

- ► LOG :: /usr/sbin/agents/saphana/monitorsaphana :: Tue Apr 24 01:48:57 EDT 2018 :: check status of sap hana instance.

- ► LOG :: /usr/sbin/agents/sap/sapstartctrl :: Tue Apr 24 01:48:57 EDT 2018 :: Enter function Control_instance() Check_cmd.\n

- ► LOG :: /usr/sbin/agents/sap/sapstartctrl :: Tue Apr 24 01:48:58 EDT 2018 :: Control_instance sapcontrol GetProcessList rc=3:

  24.04.2018 01:48:58
  GetProcessList
  OK
  0 name: hdbdaemon
  0 description: HDB Daemon

```
0 dispstatus: GREEN→ The monitor script is called again to check the status.
The status is GREEN.
0 textstatus: Running
0 starttime: 2018 04 24 01:47:44
0 elapsedtime: 0:01:14
0 pid: 54774
1 name: hdbcompileserver
1 description: HDB Compileserver
1 dispstatus: GREEN
1 textstatus: Running
1 starttime: 2018 04 24 01:47:51
1 elapsedtime: 0:01:07
1 pid: 54973
2 name: hdbindexserver
2 description: HDB Indexserver-S01
2 dispstatus: GREEN
2 textstatus: Running
2 starttime: 2018 04 24 01:47:52
2 elapsedtime: 0:01:06
2 pid: 55013
3 name: hdbnameserver
3 description: HDB Nameserver
3 dispstatus: GREEN
3 textstatus: Running
3 starttime: 2018 04 24 01:47:45
3 elapsedtime: 0:01:13
3 pid: 54839
4 name: hdbpreprocessor
4 description: HDB Preprocessor
4 dispstatus: GREEN
4 textstatus: Running
4 starttime: 2018 04 24 01:47:51
4 elapsedtime: 0:01:07
4 pid: 54975
5 name: hdbwebdispatcher
5 description: HDB Web Dispatcher
5 dispstatus: GREEN
5 textstatus: Running
5 starttime: 2018 04 24 01:48:40
5 elapsedtime: 0:00:18
5 pid: 55588
6 name: hdbxsengine
6 description: HDB XSEngine-S01
6 dispstatus: GREEN
6 textstatus: Running
6 starttime: 2018 04 24 01:47:52
6 elapsedtime: 0:01:06
6 pid: 55015
```

► LOG :: /usr/sbin/agents/sap/sapstartctrl :: Tue Apr 24 01:48:58 EDT 2018 :: The instance is running.

► LOG :: /usr/sbin/agents/sap/saphdbctrl :: Tue Apr 24 01:48:58 EDT 2018 :: saphdb_ci_status sapstartctrl -a status -p HDB S01_HDB01 OpState=1.

- LOG :: /usr/sbin/agents/saphana/monitorsaphana :: Tue Apr 24 01:48:58 EDT 2018
  :: SAP_HANA is already running.
- LOG :: /usr/sbin/agents/saphana/monitorsaphana :: Tue Apr 24 01:48:58 EDT 2018
  :: sap hana is monitorable.

HANA OS Healthchecker

This appendix introduces the HANA OS Healthchecker (HOH) tool. This tool comes *as is* with no support of any kind by IBM or anyone else.

This appendix covers the following topics:

- ► Introduction
- ► What it checks
- ► How to run the tool

Introduction

When installing or reviewing an SAP HANA on IBM Power Systems installation, there are a fair number of settings that are not automatically checked by any of the officially supported tools.

These settings can be forgotten or ignored. To fill this gap in a manner that is *not supported by IBM or anyone else*, a tool that is called HOH was developed, which checks multiple configuration settings.

This tool was developed with maintenance in mind, so the settings that it checks are not part of the HOH core but of the JSON files that come with it. Hence, the tool is easy to maintain when changes by any of the vendors occur.

> **Note:** This tool is not an official one, and is *not supported by anybody*. If you choose to run it, you run it *at your own risk and accept all responsibility for it*.

What it checks

At the time of writing, the current HOH version is Version 1.17. The tool can be found at GitHub.

The tool checks the following settings:

► Network Time Protocol (NTP) configuration by running the **datetimectl systemd** command. This function covers both ntpd and chrony.
► The `sysctl` settings.
► SELINUX (only Red Hat).
► `saptune` (only SUSE).
► `tuned` (only Red Hat).
► Installed packages.
► IBM service and productivity tools installation statuses.
► Multipath basic checks for Extents File System (XFS) and 2145 storage combination only.

How to run the tool

The tool is hosted in a public repository of GitHub, so it can be cloned or downloaded directly from there. To clone it from GitHub, install the Git client if not yet installed on your system, and then clone the repository as shown in Example A-1.

Example A-1   Running git clone to clone HANA OS Healthchecker

```
# git clone https://github.com/bolinches/HANA-TDI-healthcheck
Cloning into 'HANA-TDI-healthcheck'...
remote: Enumerating objects: 127, done.
remote: Counting objects: 100% (127/127), done.
remote: Compressing objects: 100% (58/58), done.
remote: Total 294 (delta 78), reused 114 (delta 69), pack-reused 167
Receiving objects: 100% (294/294), 52.72 KiB | 0 bytes/s, done.
Resolving deltas: 100% (188/188), done.
```

If you already downloaded the HOH in the past and want to update to the latest version, run **git pull** inside of the directory to where it was cloned, as shown in Example A-2.

Example A-2   Running git pull to update HANA OS Healthchecker

```
# cd HANA-TDI-healthcheck
# git pull
```

To run the tool, go to the cloned repository and call it directly by passing one of the storages (XFS, IBM Enterprise Storage Server®, or Network File System (NFS)), as shown in Example A-3.

Example A-3   Running HANA OS Healthchecker with XFS

```
# cd HANA-TDI-healthcheck
# ./hoh.py XFS
Welcome to HANA OS Healthchecker (HOH) version 1.17

Please use https://github.com/bolinches/HANA-TDI-healthcheck to get latest
versions and report issues about HOH.

The purpose of HOH is to supplement the official tools like HWCCT not to
substitute them, always refer to official documentation from IBM, SuSE/RedHat, and
SAP

You should always check your system with latest version of HWCCT as explained on
SAP note:1943937 - Hardware Configuration Check Tool - Central Note

JSON files versions:
        Supported OS:                       0.6
        sysctl:                             1.3
        Packages:                           0.2
        IBM Power packages:                 0.4
        IBM Spectrum Virtualize multipath:  1.0

This software comes with absolutely no warranty of any kind. Use it at your own
risk

Do you want to continue? (y/n):
```

When you choose to continue at your own risk, the tool generates an output for your system. A ready for use output with SUSE 12 SP2 is shown in Example A-4.

Example A-4   SUSE 12 SP2 HANA OS Healthchecker dirty run

```
Checking OS version

OK:  SUSE Linux Enterprise Server 12 SP2 is a supported OS for this tool

Checking NTP status with timedatectl

OK: NTP is configured in this system
ERROR: NTP sync is not activated in this system. Please check timedatectl command


Checking if saptune solution is set to HANA
```

```
2205917 - SAP HANA DB: Recommended OS settings for SLES 12 / SLES for SAP
Applications 12 -
    KernelMMTransparentHugepage Expected: never
    KernelMMTransparentHugepage Actual  :
The parameters listed above have deviated from the specified SAP solution
recommendations.

ERROR: saptune is *NOT* fully using the solution HANA

The following individual SAP Notes recommendations are available via sapnote
Consider enabling ALL of them, including 2161991 as only sets NOOP as I/O
scheduler

All notes (+ denotes manually enabled notes, * denotes notes enabled by
solutions):
    1275776Linux: Preparing SLES for SAP environments
    1557506Linux paging improvements
    1984787SUSE LINUX Enterprise Server 12: Installation notes
    2161991VMware vSphere (guest) configuration guidelines
    2205917SAP HANA DB: Recommended OS settings for SLES 12 / SLES for SAP
Applications 12
    SAP_ASESAP_Adaptive_Server_Enterprise
    SAP_BOBJSAP_Business_OBJects
    SUSE-GUIDE-01SLES 12 OS Tuning & Optimization Guide – Part 1
    SUSE-GUIDE-02SLES 12: Network, CPU Tuning and Optimization – Part 2

Remember: if you wish to automatically activate the solution's tuning options
after a reboot,you must instruct saptune to configure "tuned" daemon by running:
    saptune daemon start

Checking sysctl settings:

ERROR: net.core.rmem_max is 229376 and should be 56623104
ERROR: net.core.somaxconn is 128 and should be 4096
ERROR: net.ipv4.tcp_mem is 97923 130564 195846 and should be 56623104 56623104
56623104
ERROR: net.ipv4.tcp_tw_reuse is 0 and should be 1
OK: net.ipv4.tcp_timestamps it is set to the recommended value of 1
ERROR: net.ipv4.tcp_max_syn_backlog is 2048 and should be 8192
OK: net.ipv4.tcp_slow_start_after_idle it is set to the recommended value of 0
ERROR: net.ipv4.tcp_rmem is 65536 87380 6291456 and should be 65536 262088
56623104
ERROR: net.ipv4.tcp_wmem is 65536 16384 4194304 and should be 65536 262088
56623104
ERROR: net.core.wmem_max is 229376 and should be 56623104
ERROR: net.ipv4.tcp_syn_retries is 6 and should be 8
OK: kernel.numa_balancing it is set to the recommended value of 0
ERROR: net.ipv4.tcp_tw_recycle is 0 and should be 1

Checking packages install status:

OK: ipmitool installation status is as expected
OK: powerpc-utils installation status is as expected
OK: pseries-energy installation status is as expected
ERROR: ibmPMLinux installation status is *NOT* as expected
```

```
OK: ppc64-diag installation status is as expected

Checking IBM service and productivity tools packages install status:

WARNING: ibm-power-nonmanaged-rhel7 installation status is *NOT* as expected.
Check that at least one package is installed
WARNING: ibm-power-nonmanaged-sles12 installation status is *NOT* as expected.
Check that at least one package is installed
OK: ibm-power-kvmguest-sles12 installation status is not installed
WARNING: ibm-power-managed-rhel7 installation status is *NOT* as expected. Check
that at least one package is installed
OK: ibm-power-kvmguest-rhel7 installation status is not installed
WARNING: ibm-power-nonmanaged-sles15 installation status is *NOT* as expected.
Check that at least one package is installed
WARNING: ibm-power-baremetal-rhel7 installation status is *NOT* as expected. Check
that at least one package is installed
WARNING: ibm-power-baremetal-sles15 installation status is *NOT* as expected.
Check that at least one package is installed
WARNING: ibm-power-managed-sles12 installation status is *NOT* as expected. Check
that at least one package is installed
WARNING: ibm-power-managed-sles15 installation status is *NOT* as expected. Check
that at least one package is installed
OK: ibm-power-kvmguest-sles15 installation status is not installed
WARNING: ibm-power-baremetal-sles12 installation status is *NOT* as expected.
Check that at least one package is installed

Checking simple multipath.conf test

OK:  2145 disk type detected
ERROR:  multipath.conf does not exists

The summary of this run:

   SELinux not tested
   time configuration reported 1 deviation[s]
   saptune/tuned reported deviations
   sysctl reported 10 deviation[s] and 0 warning[s]
   packages reported 1 deviation[s]
   IBM service and productivity tools packages reported deviations
   XFS with IBM Spectrum Virtualize in use and no multipath.conf file detected
   2145 disk detected. Be sure to follow IBM Storage sizing guidelines:
https://www-01.ibm.com/support/docview.wss?uid=tss1flash10859&aid=1
```

There are multiple issues to fix in Example A-4 on page 149. Address them and run the tool
again until it reports no errors. Then, you are ready to proceed to the next step.

> **Note:** As a reminder about the time, be sure is fixed on the **timedatectl** level, and not ntpd
> or chrony only. Hint: Run **timedatectl set-ntp 1**.

As a final comment on HOH, this tool is *not supported by IBM or anybody* because it is a
collaborative effort. If you have questions, bug reports, requests, and so on, add them to the
tool GitHub page.

Example of a multipath.conf file for SAP HANA systems

This appendix provides a tested `multipath.conf` file.

This appendix covers the following topics:

- ► Introduction
- ► The multipath.conf file

Introduction

Example B-1 presented here is, at the time of writing, the recommended file for IBM Power Systems running Linux, SAP HANA, and IBM Spectrum Virtualize storages (disk type 2145).

The multipath.conf file

The `multipaths` section uses alias names for easier identification at the operating system (OS) level. The `wwid` in Example B-1 is going to be different in your environment.

Example B-1   The multipath.conf example for disk type 2145

```
defaults {
  fast_io_fail_tmo     5
  user_friendly_names  no
}

multipaths {

    #ROOTVG
    multipath {
        wwid 3600507640081811fe800000000003e4a
        alias ROOTVG
    }

    #HANA DATA
    multipath {
        wwid 3600507640081811fe800000000003e7a
        alias HANA_DATA_1_1
    }
    multipath {
        wwid 3600507640081811fe800000000003e79
        alias HANA_DATA_1_2
    }
     multipath {
        wwid 3600507640081811fe800000000003e78
        alias HANA_DATA_1_3
    }
      multipath {
        wwid 3600507640081811fe800000000003e77
        alias HANA_DATA_1_4
    }

    #HANA LOG
      multipath {
        wwid 3600507640081811fe800000000003e7e
        alias HANA_LOG_1_1
    }
      multipath {
        wwid 3600507640081811fe800000000003e7d
        alias HANA_LOG_1_2
    }
        multipath {
```

```
            wwid 3600507640081811fe800000000003e7c
            alias HANA_LOG_1_3
    }
        multipath {
            wwid 3600507640081811fe800000000003e7b
            alias HANA_LOG_1_4
    }

    #HANA SHARED
        multipath {
            wwid 3600507640081811fe800000000003e7f
            alias HANA_SHARED01
      }
}


devices {
    device {
        vendor   "IBM"
        product "2145"
        path_grouping_policy  group_by_prio
        prio  "alua"
        path_checker   "tur"
        path_selector "service-time 0"
        failback   "immediate"
        rr_weight "priorities"
        no_path_retry "fail"
        rr_min_io_rq  32
        dev_loss_tmo   600
        fast_io_fail_tmo  5
    }
}
```

This is a base example that was tested with both Red Hat Enterprise Linux 7.x and SUSE Linux Enterprise Server 12/15 series. Always refer to the official documentation of the specific storage and versions that you are using before using this example in production.

> **Note:** For all `multipath.conf` files, perform the following tests to check whether the settings match for a better solution about how to get to the correct configuration:
>
> 1. Perform a rolling takeover of Virtual I/O Server (VIOS) and reintegration expectation: path recovery. (Simulate a rolling VIOS upgrade. The start and stop sequence of VIOS must match the time that is typically needed for VIOS maintenance).
>
> 2. Pull the cable and reattach.
>
> 3. Simulate a rolling maintenance of storage head nodes and check whether the paths are recovered in SUSE Linux Enterprise Server.

The critical tests

Here are more details about the tests that are mentioned in the Note box above:

1. Rolling takeover of VIOS and storage head nodes.

   The objective is to ensure that the timeout settings allow the paths to reintegrate themselves.

2. Pull the cable and reattach it.

   The objective is to test the instant path recovery.

C

SAP HANA software stack installation for a scale-out scenario

This appendix describes the differences between scale-up and scale-out HANA installations.

This chapter also provides the scale-out prerequisites that you must meet when you plan to use the Storage Connector API for sharing the data and log areas among the cluster nodes.

This appendix covers the following topics:

- ► Differences between scale-out and scale-up installations
- ► Installing HANA scale-out clusters
- ► Postinstallation notes

Differences between scale-out and scale-up installations

When installing HANA on a number of scale-out nodes, run the installer on the first node, and then request to add more nodes to the HANA instance. The installer then prompts you for the additional host names and adds the nodes to the existing HANA instance.

The HANA binary files are installed in the `/hana/shared` directory, which is shared among all the cluster nodes. As such, there is no duplication of the binary files on each node. After installation, each worker node has an entry inside the `/hana/data/<SID>` and `/hana/log/<SID>` directories, in the form of `mntNNNN`, characterizing a cluster-based layout of data and logs.

If you are using a shared storage approach, Elastic Storage Server, or Network File System (NFS), you do not need any special configuration for installing HANA. If you are using the storage connector API, then you must start the installer with a setup file, as described in "Storage Connector API setup" on page 164.

> **Note:** When using shared storage for the HANA data and log areas, Elastic Storage Server, or NFS, validate that these file systems are mounted on all nodes before installing HANA. If you use the storage connector for the data and log areas, check that they are unmounted on all nodes before installing HANA.
>
> The HANA shared file system is mounted for both cases.

Prerequisites

Your nodes must comply with the following prerequisites before you start a HANA scale-out installation:

- ► The date and time of all the nodes must be synchronized. Use a suitable Network Time Protocol (NTP) server to comply with this requirement. If you do not have any NTP servers available, one of the nodes can act as one.
- ► Ensure that all nodes can ping one another by name by using both their short and fully qualified host names.
- ► A scale-out environment is characterized by a true cluster that is built at the application layer, that is, the HANA database (DB). To ease the management of the cluster, set up password-less communication among the cluster nodes.

Installing HANA scale-out clusters

Both the graphical and text-mode HANA installations are available for a scale-out installation. The next sections show only the differences in the steps when compared to a scale-up cluster installation.

All our installations use a four-node cluster with three worker nodes (`saphana005`, `hana006`, and `hana007`), and one standby node (`hana008`).

Scale-out graphical installation

To start a graphical installation of HANA, follow the instructions in 6.2.1, "GUI installation" on page 86 until you get to the window that is shown in Figure 6-5 on page 90. Perform a **Multiple-Host System** installation instead, as shown in Figure C-1 and shown as **1**. Then, check that the root user and password are entered correctly, as shown by **2**. Keep the installation path as `/hana/shared`, and then click **Add Host**, as shown by **3**, to add the other nodes into the cluster. The node where you are running the wizard becomes the master node.

Figure C-1   Multiple host (scale-out) installation

Every time that you click **Add Host**, a window similar to Figure C-2 opens. Add one node at a time by using its host name, which is shown by **1**, and select the appropriate node role, which is shown by **2**. There is no need to change the other parameters.

Figure C-2   Adding a node to the installation

In our scenario, we have two more worker nodes and one standby node. So, we perform the add node step three times until we have a layout with all of our nodes, as shown in Figure C-1 on page 159 and shown by **3**.

The remaining of the installation process looks the same as a scale-up installation from this point. You can resume the installation by following the steps from Figure 6-6 on page 91 onward.

Scale-out text-mode installation

To start a text-mode installation of HANA, see 6.2.2, "Text-mode installation" on page 97. The installation flow looks similar to Example 6-3 on page 98, except that you enter yes to the question Do you want to add hosts to the system?. If you do so, you are prompted to add the information that is required for each one of the nodes. The complete installation flow is shown in Example C-1 with the user inputs in **bold**.

Example C-1   Scale-out installation: Text mode

```
saphana005:/tmp/51052031/DATA_UNITS/HDB_LCM_LINUX_PPC64LE # ./hdblcm

SAP HANA Lifecycle Management - SAP HANA Database 2.00.010.00.1491294693
*************************************************************************

Scanning Software Locations...
Detected components:
    SAP HANA Database (2.00.010.00.1491294693) in /mnt/SW/HANA/HANA
2.0/SPS01/51052031/DATA_UNITS/HDB_SERVER_LINUX_PPC64LE/server

[... snip ...]

    Xsa Cockpit 1 (1.000.0) in /mnt/SW/HANA/HANA
2.0/SPS01/51052031/DATA_UNITS/XSA_CONTENT_10/XSACXSACOCKPIT00_0.zip

Choose an action

  Index | Action             | Description
```

```
              ------------------------------------------------
       1      │ install            │ Install new system
       2      │ extract_components │ Extract components
       3      │ Exit (do nothing)  │
```

Enter selected action index [3]: 1

SAP HANA Database version '2.00.010.00.1491294693' will be installed.

Select additional components for installation:

```
  Index │ Components │ Description

  ---------------------------------------------------------------------------------
    1    │ server     │ No additional components
    2    │ all        │ All components
    3    │ afl        │ Install SAP HANA AFL (incl.PAL,BFL,OFL,HIE) version
2.00.010.0000.1491308763
    4    │ client     │ Install SAP HANA Database Client version 2.1.37.1490890836
    5    │ smartda    │ Install SAP HANA Smart Data Access version 2.00.0.000.0
    6    │ xs         │ Install SAP HANA XS Advanced Runtime version 1.0.55.288028
    7    │ epmmds     │ Install SAP HANA EPM-MDS version 2.00.010.0000.1491308763
```

Enter comma-separated list of the selected indices [4]: 1,4
Enter Installation Path [/hana/shared]:
Enter Local Host Name [saphana005]:
Do you want to add hosts to the system? (y/n) [n]: y
Enter comma-separated host names to add: hana006,hana007,hana008
Enter Root User Name [root]:
Collecting information from host 'hana006'...
Collecting information from host 'hana007'...
Collecting information from host 'hana008'...
Information collected from host 'hana008'.
Information collected from host 'hana007'.
Information collected from host 'hana006'.

Select roles for host 'hana006':

```
  Index │ Host Role                │ Description
  ------------------------------------------------------------------
    1    │ worker                   │ Database Worker
    2    │ standby                  │ Database Standby
    3    │ extended_storage_worker  │ Dynamic Tiering Worker
    4    │ extended_storage_standby │ Dynamic Tiering Standby
    5    │ streaming                │ Smart Data Streaming
    6    │ rdsync                   │ Remote Data Sync
    7    │ ets_worker               │ Accelerator for SAP ASE Worker
    8    │ ets_standby              │ Accelerator for SAP ASE Standby
    9    │ xs_worker                │ XS Advanced Runtime Worker
   10    │ xs_standby               │ XS Advanced Runtime Standby
```

Enter comma-separated list of selected indices [1]: 1
Enter Host Failover Group for host 'hana006' [default]:
Enter Storage Partition Number for host 'hana006' [<<assign automatically>>]:
Enter Worker Group for host 'hana006' [default]:

```
Select roles for host 'hana007':

  Index │ Host Role               │ Description
  -------------------------------------------------------------------
  1     │ worker                  │ Database Worker
  2     │ standby                 │ Database Standby
  3     │ extended_storage_worker │ Dynamic Tiering Worker
  4     │ extended_storage_standby │ Dynamic Tiering Standby
  5     │ streaming               │ Smart Data Streaming
  6     │ rdsync                  │ Remote Data Sync
  7     │ ets_worker              │ Accelerator for SAP ASE Worker
  8     │ ets_standby             │ Accelerator for SAP ASE Standby
  9     │ xs_worker               │ XS Advanced Runtime Worker
  10    │ xs_standby              │ XS Advanced Runtime Standby
```

Enter comma-separated list of selected indices [1]: 1
Enter Host Failover Group for host 'hana007' [default]:
Enter Storage Partition Number for host 'hana007' [<<assign automatically>>]:
Enter Worker Group for host 'hana007' [default]:

```
Select roles for host 'hana008':

  Index │ Host Role               │ Description
  -------------------------------------------------------------------
  1     │ worker                  │ Database Worker
  2     │ standby                 │ Database Standby
  3     │ extended_storage_worker │ Dynamic Tiering Worker
  4     │ extended_storage_standby │ Dynamic Tiering Standby
  5     │ streaming               │ Smart Data Streaming
  6     │ rdsync                  │ Remote Data Sync
  7     │ ets_worker              │ Accelerator for SAP ASE Worker
  8     │ ets_standby             │ Accelerator for SAP ASE Standby
  9     │ xs_worker               │ XS Advanced Runtime Worker
  10    │ xs_standby              │ XS Advanced Runtime Standby
```

Enter comma-separated list of selected indices [1]: 2
Enter Host Failover Group for host 'hana008' [default]:
Enter Worker Group for host 'hana008' [default]:

Enter SAP HANA System ID: RB1
Enter Instance Number [00]: 13
Enter Local Host Worker Group [default]:

```
  Index │ System Usage │ Description
  --------------------------------------------------------------------------------
  1     │ production   │ System is used in a production environment
  2     │ test         │ System is used for testing, not production
  3     │ development  │ System is used for development, not production
  4     │ custom       │ System usage is neither production, test nor development
```

Select System Usage / Enter Index [4]: 2
Enter Location of Data Volumes [/hana/data/RB1]:
Enter Location of Log Volumes [/hana/log/RB1]:
Restrict maximum memory allocation? [n]:

```
Enter Certificate Host Name For Host 'saphana005' [saphana005]:
Enter Certificate Host Name For Host 'hana006' [hana006]:
Enter Certificate Host Name For Host 'hana007' [hana007]:
Enter Certificate Host Name For Host 'hana008' [hana008]:
Enter System Administrator (rb1adm) Password: ********
Confirm System Administrator (rb1adm) Password: ********
Enter System Administrator Home Directory [/usr/sap/RB1/home]:
Enter System Administrator Login Shell [/bin/sh]:
Enter System Administrator User ID [1001]:
Enter Database User (SYSTEM) Password: ********
Confirm Database User (SYSTEM) Password: ********
Restart system after machine reboot? [n]:

Summary before execution:
=========================

SAP HANA Database System Installation
   Installation Parameters
       Remote Execution: ssh
       Database Isolation: low

[ ... snip ...]

Additional Hosts
      hana008
         Role: Database Standby (standby)
         High-Availability Group: default
         Worker Group: default
         Storage Partition: N/A
      hana007
         Role: Database Worker (worker)
         High-Availability Group: default
         Worker Group: default
         Storage Partition: <<assign automatically>>
      hana006
         Role: Database Worker (worker)
         High-Availability Group: default
         Worker Group: default
         Storage Partition: <<assign automatically>>

Do you want to continue? (y/n): y

Installing components...
Installing SAP HANA Database...
  Preparing package 'Saphostagent Setup'...

[ ... snip ...]

  Updating SAP HANA Database instance integration on host 'hana008'...
Creating Component List...
SAP HANA Database System installed
You can send feedback to SAP with this form:
https://saphana005:1129/lmsl/HDBLCM/RB1/feedback/feedback.html
```

```
Log file written to
'/var/tmp/hdb_RB1_hdblcm_install_2017-07-11_21.57.20/hdblcm.log' on host
'saphana005'.
```

Storage Connector API setup

There is more than one method to use the storage connector API. In this publication, we use a logical volume manager (LVM) -based one that is called *hdb_ha.fcClientLVM*. For more information and to check how to use a different storage connector, see *SAP HANA Fibre Channel Storage Connector Admin Guide*.

Complete the following steps:

1. To use the storage connector during a scale-out installation, you must create a text file that is used as the initial input to what will be part of the HANA instance `global.ini` configuration file. This text file provides instructions about the storage connector about how to map the volume groups (VGs) and logical volumes (LVs) that you created in "Storage Connector API for the data and log areas" on page 78.

   Example C-2 shows a `global.ini` file that works for our four-node cluster, with one master node, two worker nodes, and one standby node. The file system layout that we use is the one that is described in "Storage Connector API for the data and log areas" on page 78. Create this file in the `/hana/shared` directory so that all nodes have access to it.

   Example C-2   A global.ini file to be used at installation time for using the logical volume manager storage connector

```
# cat /hana/shared/global.ini
[storage]
ha_provider = hdb_ha.fcClientLVM
partition_1_data__lvmname = hanadata01-datalv01
partition_1_log__lvmname = hanalog01-loglv01
partition_2_data__lvmname = hanadata02-datalv02
partition_2_log__lvmname = hanalog02-loglv02
partition_3_data__lvmname = hanadata03-datalv03
partition_3_log__lvmname = hanalog03-loglv03

partition_*_*__prtype = 5
partition_*_*__mountoptions = -t xfs
```

2. Call the installer by using the parameter **--storage_cfg=<global.ini directory>**, where the input is the path to the directory that contains the `global.ini` file (not the path to the file itself), which is `/hana/shared`, as shown in Example C-2.

   To check how to start the graphical installer, see Figure 6-2 on page 87. Start it by running the following command:

   `./hdblcmgui --storage_cfg=/hana/shared`

3. Follow the guidelines in "Scale-out graphical installation" on page 159 to proceed with the installation.

   Similarly, if you want to install HANA by using the text-mode installer, go to the `HDB_LCM_LINUX_PPC64LE` folder, as explained in Example 6-3 on page 98, and run the following command:

   `./hdblcm --storage_cfg=/hana/shared`

4. Follow the guidelines in "Scale-out text-mode installation" on page 160 to proceed with the installation.

Postinstallation notes

After installing HANA on the scale-out cluster, you can connect to it by using the HANA Studio interface. The process for adding the HANA instance is the same as outlined in 6.3, "Postinstallation notes" on page 101. When you add the instance, use the master node of your cluster (the one you from which you ran the installation) as the node to which to connect.

> **Note:** When you add `System` to HANA, you must select **Multiple containers** because HANA V2.0 sps01 uses multiple containers DB mode. Otherwise, you see the following error message.

After adding the instance in HANA Studio, go to the **Landscape** → **Services** tab to confirm that the services are distributed among all the nodes, as shown in Figure C-3.

Figure C-3   Scale-out instance in HANA Studio

Also, review the **Landscape** → **Hosts** tab, as shown in Figure C-4. Node `hana008` is displayed as `STANDBY` for the services for our installation.

Figure C-4   *Scale-out system: Current node roles*

As a best practice, perform failover tests by shutting down the HANA service in one of the worker nodes, or shut down the node, and observe that the standby node takes over its role to open a HANA Studio connection to another node that is running to check the cluster status.

> **Note:** A scale-out cluster can handle only as many simultaneous node outages as the number of standby nodes in the cluster. For example, if you have only one standby node, you can sustain an outage of a single node. If two nodes fail at the same time, your HANA DB is brought offline. If you must protect your business against the failure of multiple nodes at the same time, add as many standby nodes as you need.

Related publications

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

IBM Redbooks

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

► *Implementing High Availability and Disaster Recovery Solutions with SAP HANA on IBM Power Systems*, REDP-5443

► *Implementing the IBM System Storage SAN Volume Controller with IBM Spectrum Virtualize V8.2.1*, SG24-7933

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

ibm.com/redbooks

Online resources

These websites are also relevant as further information sources:

► IBM Infrastructure for SAP HANA

https://www.ibm.com/it-infrastructure/power/sap-hana

► IBM PowerVM

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

► IBM Service and productivity tools

https://www14.software.ibm.com/webapp/set2/sas/f/lopdiags/home.html

► Red Hat Enterprise Linux evaluation

https://access.redhat.com/products/red-hat-enterprise-linux/evaluation

► SAP HANA

https://www.sap.com/products/hana/implementation/sizing.html

► SAP HANA System Replication in pacemaker cluster

https://access.redhat.com/articles/3004101

Help from IBM

IBM Support and downloads

ibm.com/support

IBM Global Services

ibm.com/services

IBM®

Printed in U.S.A.

Get connected

Redbooks®
ibm.com/redbooks