SAP HANA System Replication Scale-Out - Performance Optimized Scenario

SUSE Linux Enterprise Server for SAP Applications 12 SP2 or newer

Fabian Herschel, Distinguished Architect SAP (SUSE)

Bernd Schubert, SAP Solution Architect (SUSE)

Lars Pinne, System Engineer (SUSE)

SUSE

SAP HANA System Replication Scale-Out - Performance Optimized Scenario

Date: 2024-12-12

SUSE® Linux Enterprise Server for SAP Applications is optimized in various ways for SAP* applications. This guide provides detailed information about how to install and customize SUSE Linux Enterprise Server for SAP Applications for SAP HANA Scale-Out system replication automation in the performance optimized scenario. It is based on SUSE Linux Enterprise Server for SAP Applications 12 SP2. The concept however can also be used with SUSE Linux Enterprise Server for SAP Applications 12 SP3 or newer.

Disclaimer: Documents published as part of the SUSE Best Practices series have been contributed voluntarily by SUSE employees and third parties. They are meant to serve as examples of how particular actions can be performed. They have been compiled with utmost attention to detail. However, this does not guarantee complete accuracy. SUSE cannot verify that actions described in these documents do what is claimed or whether actions described have unintended consequences. SUSE LLC, its affiliates, the authors, and the translators may not be held liable for possible errors or the consequences thereof.

Contents

1 About this guide

1.1 Introduction

SUSE® Linux Enterprise Server for SAP Applications is optimized in various ways for SAP* applications. This guide provides detailed information about installing and customizing *SUSE Linux Enterprise Server for SAP Applications* for SAP HANA scale-out system replication automation in the performance optimized scenario.

High availability is an important aspect of running your mission-critical SAP HANA servers.

The SAP HANA scale-out system replication is a synchronization of all data in SAP HANA to a second SAP HANA system. The SAP HANA itself replicates all of its data to a secondary SAP HANA instance. It is an out-of-the-box, standard feature.

The recovery time objective (RTO) is minimized through the data replication at regular intervals. SAP HANA supports asynchronous and synchronous modes. The document at hand describes the synchronous replication from memory into memory of the second system. This is the only method that allows the cluster to make a decision based on coded algorithms.

1.2 Additional documentation and resources

Chapters in this manual contain links to additional documentation resources that are either available on the system or on the Internet.

For the latest SUSE product documentation updates, see https://documentation.suse.com ↗.

Find white-papers, best-practices guides, and other resources at the

- SUSE Linux Enterprise Server for SAP Applications resource library: https://www.suse.com/products/sles-for-sap/resource-library/ ↗

- SUSE Best Practices Web page: https://documentation.suse.com/sbp/sap/ ↗

1.3 Feedback

Several feedback channels are available:

Bugs and Enhancement Requests

> For services and support options available for your product, refer to http://www.suse.com/support/ ↗ .

To report bugs for a product component, go to https://scc.suse.com/support/ ↗ requests, log in, and select *Submit New SR* (Service Request).

Mail

> For feedback on the documentation of this product, you can send a mail to doc-team@suse.com (mailto:doc-team@suse.com) ↗ . Make sure to include the document title, the product version and the publication date of the documentation. To report errors or suggest enhancements, provide a concise description of the problem and refer to the respective section number and page (or URL).

2 Scope of this documentation

This document describes how to set up an automation of an SAP HANA scale-out system replication cluster installed on two sites based on SUSE Linux Enterprise Server for SAP Applications 12 SP2. This concept can also be used with SUSE Linux Enterprise Server for SAP Applications 12 SP3 or newer.

To give a better overview the installation and setup is subdivided into seven steps.
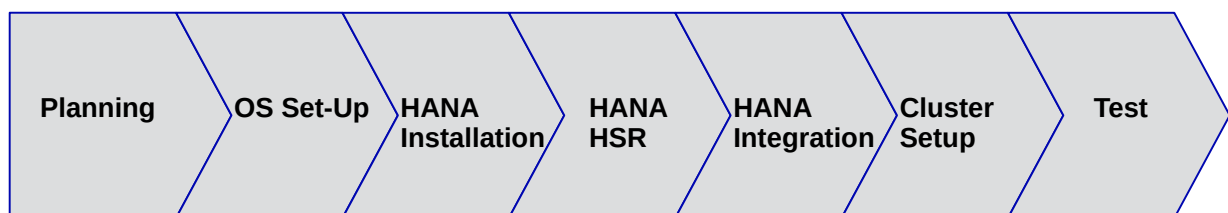

FIGURE 1: *[PLANNING] [OSSETUP] [SAPHANAINST] [SAPHANAHSR] [INTEGRATION] [CLUSTER] [TESTING]*

- Planning (section *[Planning]*)

- OS setup (section *[OsSetup]*)

- SAP HANA installation (section *[SAPHanaInst]*)

- SAP HANA system replication configuration (section *[SAPHanaHsr]*)

- SAP HANA cluster integration (section *[Integration]*)

- SLES for SAP cluster configuration (section *[Cluster]*)

- Testing (section *[Testing]*)

After the setup process you will have a SUSE Linux Enterprise Server for SAP Applications cluster controlling two 'swarms' of SAP HANA scale-out in a system replication configuration. The architecture is named the 'performance optimized scenario'.
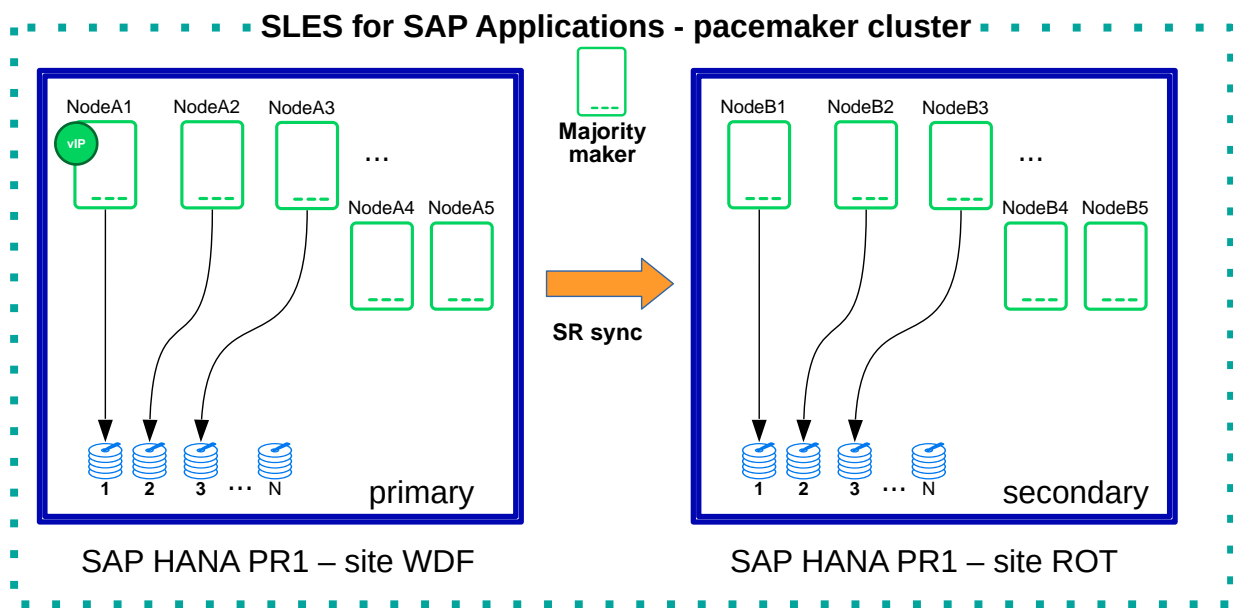

FIGURE 2: CLUSTER WITH SAP HANA SR - PERFORMANCE OPTIMIZED

3 Planning the installation

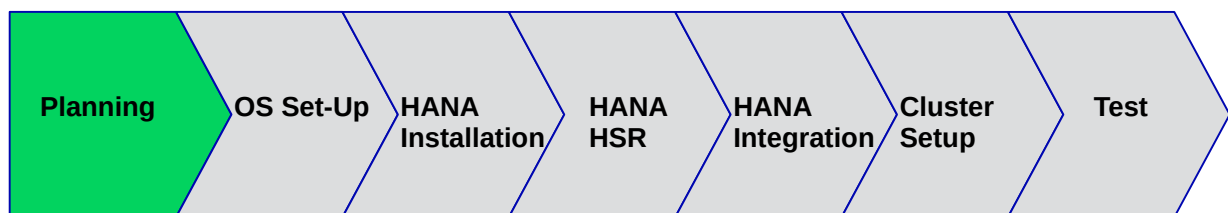Planning the installation is essential for a successful SAP HANA cluster setup.

FIGURE 3: PLANNING *[OSSETUP] [SAPHANAINST] [SAPHANAHSR] [INTEGRATION] [CLUSTER] [TESTING]*

What you need before you start:

- Software from SUSE: SUSE Linux Enterprise Server for SAP Applications installation media and a valid subscription for getting updates

- Software from SAP: SAP HANA installation media

- Physical or virtual systems including disks and NFS storage pools (see below)

- Filled parameter sheet (see below)

3.1   Minimum lab requirements and prerequisites

This section defines some minimum requirements to install SAP HANA scale-out.

Note

The minimum lab requirements mentioned here are no SAP sizing information. These data are provided only to rebuild the described cluster in a lab for test purposes. Even for such tests the requirements can increase depending on your test scenario. For productive systems ask your hardware vendor or use the official SAP sizing tools and services.

Note

Refer to SAP HANA TDI documentation for allowed storage configuration and file systems.
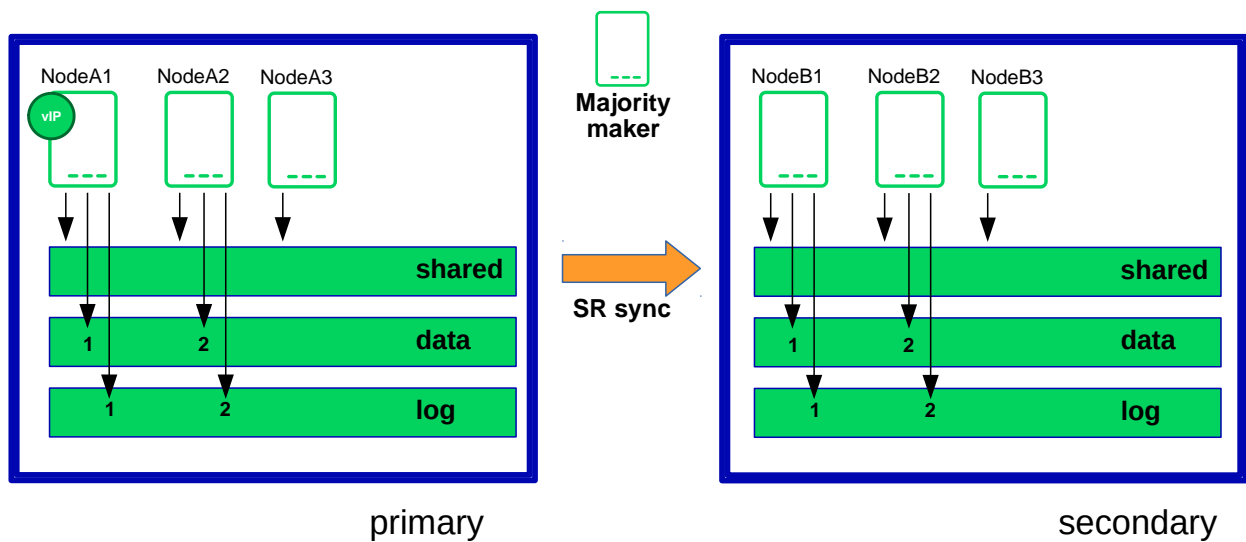
FIGURE 4: SIMPLIFIED NFS SHARE STRUCTURE OF A 2+1:2+1 SAP HANA SYSTEM REPLICATION

Requirements with 3 SAP instances per site $(2+1 : 2+1)$ - with a majority maker:

- 6 VMs with each 32 GB RAM, 50 GB disk space

- 1 VM with 2 GB RAM, 50 GB disk space

- 1 shared disk for SBD with 10 MB disk space

- 2 NFS pools (one per site) with a capacity of each 96 GB

- 1 additional IP address for takeover

FIGURE 5: SIMPLIFIED NFS SHARE STRUCTURE OF A 3+2:3+2 SAP HANA SYSTEM REPLICATION

SAP HANA System Replication Scale-Out - Performance Optimized Scenario

Requirements with 5 SAP HANA instances per site $(3+2:3+2)$ - with a majority maker:

- 10 VMs with each 32 GB RAM, 50 GB disk space

- 1 VM with 2 GB RAM, 50 GB disk space

- 1 shared disk for SBD with 10 MB disk space

- 2 NFS pools (one per site) with a capacity of each 132 GB
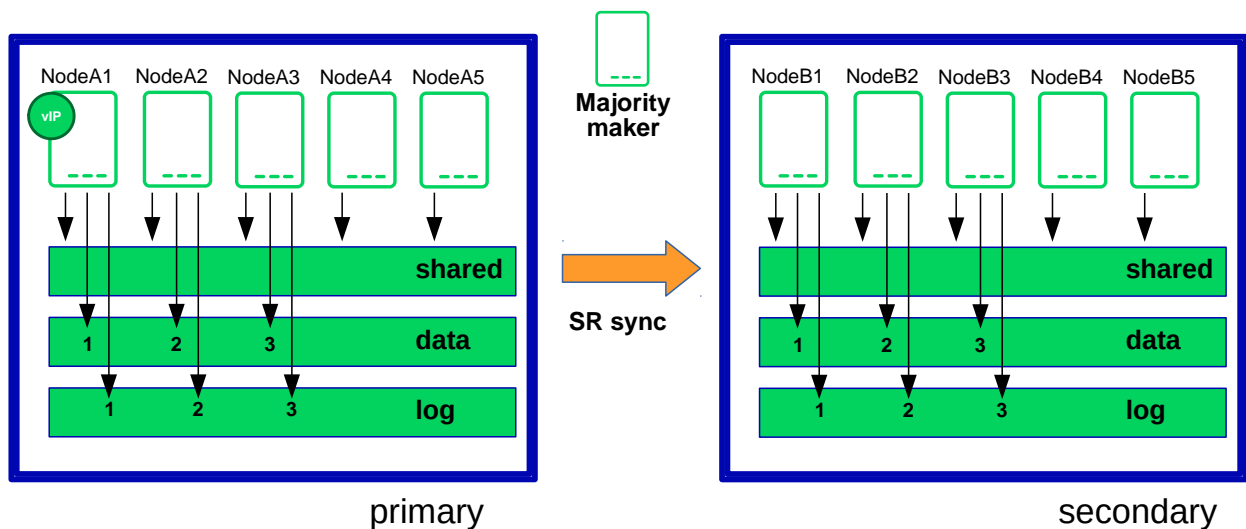
- 1 additional IP address for takeover

The SBD based fencing needs up to 3 shared block devices.
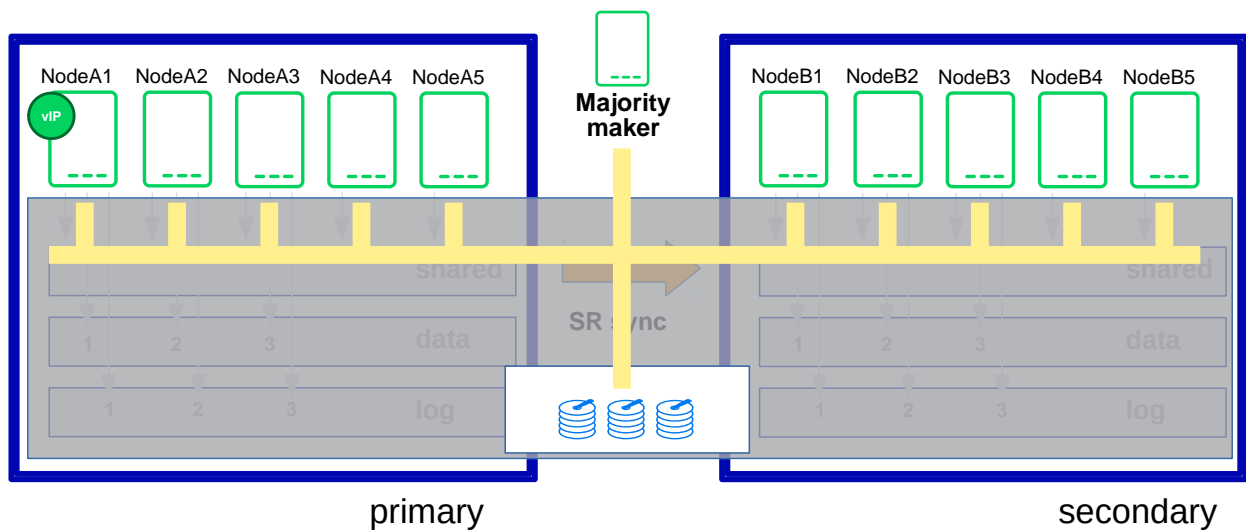
FIGURE 6: ADDITIONALLY TO THE NFS SHARES: ALL CLUSTER NODES NEED TO HAVE ACCESS TO THE SBD BLOCK DEVICES

3.2 Parameter sheet

The cluster organizing two SAP HANA 'swarms' is quite complex. The installation should be planned properly. You should have all needed parameters like SID, IP addresses and much more already in place. It is a good practice to first fill out the parameter sheet and then begin with the installation.

TABLE 1: PARAMETER SHEET TO PREPARE THE NFS BASED SETUP

Parameter	Value
Path to SLES for SAP media	

SAP HANA System Replication Scale-Out - Performance Optimized Scenario

Parameter	Value
SMT server or SCC account	
Path to SAP HANA media	
S-User for SAP marketplace	
Node names site 1	
Node names site 2	
Node name majority maker	
IP addresses of all cluster nodes	
SID	
Instance number	
Service IP address	
HANA site name site 1	
HANA site name site 2	
NFS server site 1	
NFS share "shared" site 1	
NFS share "data" site 1	
NFS share "log" site 1	
NFS server site 2	
NFS share "shared" site 2	
NFS share "data" site 2	
NFS share "log" site 2	

Parameter	Value
SBD STONITH block device(s)	
Watchdog driver	

3.3   Scale-out scenario and resource agents

To automate the failover, the High Availability Extension built into *SUSE Linux Enterprise Server for SAP Applications* is used. Two resource agents have been created to handle the scenario.

The first is the **SAPHanaController** resource agent (RA), which checks and manages the SAP HANA database instances. This RA is configured as a master/slave resource.

The master assumes responsibility for the active master name server of the SAP HANA database running in primary mode. All other instances are represented by the slave mode.
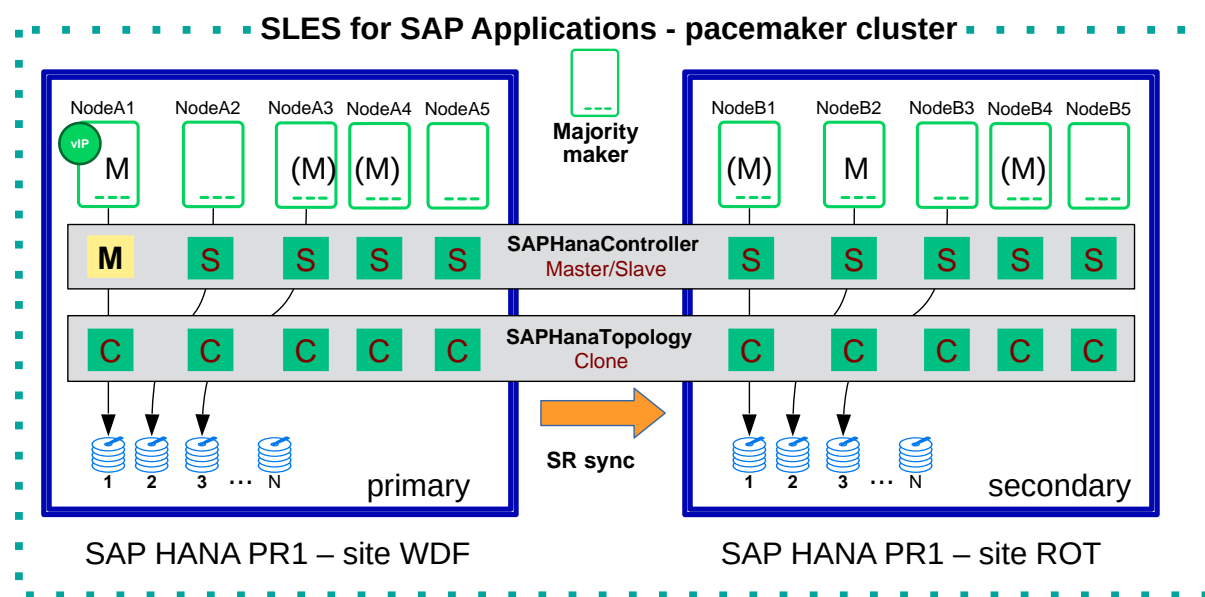
FIGURE 7: CLUSTER RESOURCE AGENTS AND MASTER/SLAVE STATUS MAPPING

The second resource agent is **SAPHanaTopology**. This RA has been created to make configuring the cluster as simple as possible. It runs on all nodes (beside the majority maker) of a SUSE Linux Enterprise High Availability Extension 12 cluster. It gathers information about the statuses and configurations of the SAP HANA system replication. It is designed as a normal (stateless) clone resource.

SAP HANA System Replication Scale-Out - Performance Optimized Scenario

SAP HANA system replication for scale-out is supported in the following scenarios or use cases:

Performance optimized, single container (A > B)

This scenario and setup is described in this document. In the performance optimized scenario an SAP HANA RDBMS on site "A" is synchronizing with an SAP HANA RDBMS on a second site "B". As the SAP HANA RDBMS on the second site is configured to preload the tables the takeover time is typically very short.

Performance optimized, multi-tenancy also named MDC (%A > %B)

Multi-tenancy is supported for all of the above scenarios and use cases. This scenario is supported since SAP HANA 1 SPS12. The setup and configuration from a cluster point of view is the same for multi-tenancy and single containers. Thus you can use the above documents for both types of scenarios.

Multi-tenancy is the default installation type for SAP HANA 2.0.

3.4   The concept of the Performance Optimized Scenario

In case of failure of the primary SAP HANA on site 1 the cluster first tries to start the takeover process. This allows to use the already loaded data at the secondary site. Typically the takeover is much faster than the local restart.

A site is noticed as "down" or "on error", if the **LandscapeHostConfiguration status** reflects this (return code 1). This happens when worker nodes are going down without any SAP HANA standby nodes left. Standby nodes are designed to perform a host auto-failover for the worker functionality.

Without any additional intervention the resource agent will wait for the SAP internal HA cluster to repair the situation locally. An additional intervention could be a custom python hook using the SAP provider **srServiceStateChanged()** available since SAP HANA 2.0 SPS01.

To achieve an automation of this resource handling process, use the SAP HANA resource agents included in the *SAPHanaSR-ScaleOut* RPM package delivered with SUSE Linux Enterprise Server for SAP Applications.

You can configure the level of automation by setting the parameter *AUTOMATED_REGISTER*. If automated registration is activated the cluster will also automatically register a former failed primary to get the new secondary.

3.5   Important prerequisites

Read the SAP Notes and papers first.

The *SAPHanaSR-ScaleOut* resource agent software package supports scale-out (multiple-box to multiple-box) system replication with the following configurations and parameters:

- The cluster must include a valid STONITH method.

- As the STONITH mechanism SBD is recommended.

- Both sites are either in the same network segment (layer 2) to allow an easy takeover of an IP Address, or you need a technique like overlay IP addresses in virtual private clouds.

- Technical users and groups, such as *<sid>adm* are defined **locally** in the Linux system.

- Name resolution of the cluster nodes and the virtual IP address should be done **locally** on **all** cluster nodes to not depend on DNS services (as it can fail, too).

- Time synchronization between the cluster nodes using reliable time services like NTP.

- Both SAP HANA sites have the same SAP Identifier (SID) and instance number.

- The SAP HANA scale-out system must have only **one** active master name server per site. It should have up to three master name server candidates (SAP HANA nodes with a configured role 'master<N>').

- The SAP HANA scale-out system must have only **one** failover group.

- The cluster described in this document does not manage any service IP address for a read-enabled secondary site.

- There is only one SAP HANA system replication like from site "A" to site "B" but no third site "C" (also called multi-tier or multi-target).

- The setup implements the performance optimized scenario but not the cost optimized scenario.

- The *saphostagent* must be running. *saphostagent* is needed to translate between the system node names and SAP host names used during the installation of SAP HANA.

- The replication mode should be either 'sync' or 'syncmem'.

- All SAP HANA instances controlled by the cluster must not be activated via *sapinit* autostart.

! Important

Automated registration of a failed primary after takeover is possible. But as a good starting configuration for projects, it is recommended to switch **off** the automated registration of a failed primary, therefore the `AUTOMATED_REGISTER="false"` is the **default**.

In this case, you need to register a failed primary after a takeover manually. Use SAP tools like SAP HANA Cockpit or **hdbnsutil**.

- Automated start of SAP HANA instances during system boot must be switched **off**.

- You need at least SAPHanaSR-ScaleOut version 0.161, SUSE Linux Enterprise Server for SAP Applications 12 SP2 and SAP HANA 1.0 SPS12 (121) or SAP HANA 2.0 SPS 2 for all mentioned setups.

! Important

You must implement a valid STONITH method. Without a valid STONITH method, the complete cluster is unsupported and will not work properly.

This setup-guide focuses on the performance optimized setup as it is the only supported scenario at the point of writing this guide.

If you need to implement a different scenario, it is strongly recommended to define a Proof-of-Concept (PoC) with SUSE. This PoC will focus on testing the existing solution in your scenario. The limitation of most of the above items is mostly due to testing limits.

4 Setting up the operating system

This section includes information you should consider during the installation of the operating system.
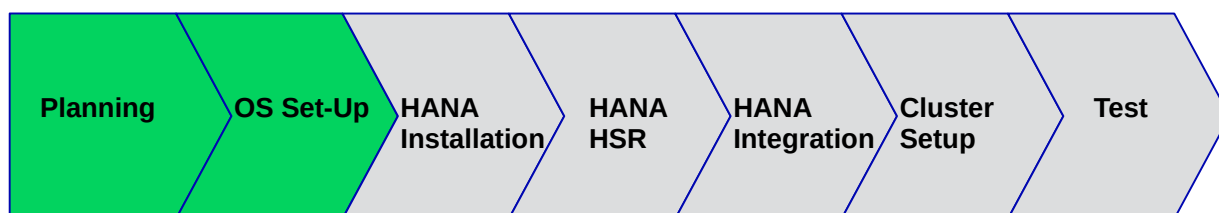

FIGURE 8: *[PLANNING]* OSSETUP *[SAPHANAINST] [SAPHANAHSR] [INTEGRATION] [CLUSTER] [TESTING]*

In this document, first SUSE Linux Enterprise Server for SAP Applications is installed and configured. Then the SAP HANA database including the system replication is set up. Finally the automation with the cluster is set up and configured.

4.1 Installing SUSE Linux Enterprise Server for SAP Applications

Multiple installation guides are already existing, with different reasons to set up the server in a certain way. Below it is outlined where this information can be found. In addition, you will find important details you should consider to get a system which is well prepared to deliver SAP HANA.

4.1.1 Installing the base operating system

Depending on your infrastructure and the hardware used, you need to adapt the installation. All supported installation methods and minimum requirement are described in the *Deployment Guide* (https://documentation.suse.com/sles/12-SP5/html/SLES-all/book-sle-deployment.html ↗). In case of automated installations you can find further information in the *AutoYaST Guide* (https://documentation.suse.com/sles/12-SP5/html/SLES-all/book-autoyast.html ↗). The major installation guide for SUSE Linux Enterprise Server for SAP Applications to fit all requirements for SAP HANA is described in the SAP notes:

- 1984787 SUSE LINUX Enterprise Server 12: Installation notes and

- 2205917 SAP HANA DB: Recommended OS settings for SLES 12 / SLES for SAP Applications 12.

4.1.2 Installing additional software

SUSE delivers with SUSE Linux Enterprise Server for SAP Applications special resource agents for SAP HANA. With the pattern *sap-hana* the resource agent for SAP HANA **ScaleUP** is installed. For the **ScaleOut** scenario you need a special resource agent. Follow the instructions below on each node if you have installed the systems based on SAP note 1984787. The pattern *High Availability* summarizes all tools recommended to be installed on **all** nodes, including the *majority maker*.

- remove package: patterns-sap-hana, SAPHanaSR, yast2-sap-ha

- install package: SAPHanaSR-ScaleOut, SAPHanaSR-ScaleOut-doc

- install pattern: ha_sles

To do so, for example, use Zypper:

EXAMPLE 1: UNINSTALL THE SAPHANASR AGENT FOR SCALEUP

As user root, type:

```
zypper remove SAPHanaSR
```

If the package is installed, you will get an output like this:

```
Loading repository data...
Reading installed packages...
Resolving package dependencies...

The following 3 packages are going to be REMOVED:
  patterns-sap-hana SAPHanaSR yast2-sap-ha

The following pattern is going to be REMOVED:
  sap-hana

3 packages to remove.
After the operation, 494.2 KiB will be freed.
Continue? [y/n/...? shows all options] (y): y
(1/3) Removing patterns-sap-hana-12.3-6.8.2.x86_64 ..............................
[done]
(2/3) Removing yast2-sap-ha-1.0.0-2.5.12.noarch ................................
[done]
(3/3) Removing SAPHanaSR-0.152.21-1.1.noarch ...................................
[done]
```

EXAMPLE 2: INSTALLATION OF THE SAPHANASR AGENT FOR SCALEOUT

As user root, type:

```
zypper in SAPHanaSR-ScaleOut SAPHanaSR-ScaleOut-doc
```

If the package is not installed yet, you should get an output like the below:

```
Refreshing service 'Advanced_Systems_Management_Module_12_x86_64'.
Refreshing service
  'SUSE_Linux_Enterprise_Server_for_SAP_Applications_12_SP3_x86_64'.
Loading repository data...
Reading installed packages...
```

```
Resolving package dependencies...

The following 2 NEW packages are going to be installed:
  SAPHanaSR-ScaleOut SAPHanaSR-ScaleOut-doc

2 new packages to install.
Overall download size: 539.1 KiB. Already cached: 0 B. After the operation,
 additional 763.1 KiB will be used.
Continue? [y/n/...? shows all options] (y): y
Retrieving package SAPHanaSR-ScaleOut-0.161.1-1.1.noarch
                                            (1/2),  48.7 KiB (211.8 KiB
 unpacked)
Retrieving: SAPHanaSR-
ScaleOut-0.161.1-1.1.noarch.rpm ................................[done]
Retrieving package SAPHanaSR-ScaleOut-doc-0.161.1-1.1.noarch
                                            (2/2), 490.4 KiB (551.3 KiB
 unpacked)
Retrieving: SAPHanaSR-ScaleOut-
doc-0.161.1-1.1.noarch.rpm ..............................[done (48.0 KiB/s)]
Checking for file
 conflicts: ...........................................................[done]
(1/2) Installing: SAPHanaSR-
ScaleOut-0.161.1-1.1.noarch ................................[done]
(2/2) Installing: SAPHanaSR-ScaleOut-
doc-0.161.1-1.1.noarch ............................[done]
```

Install the tools for High Availability on all nodes.

```
zypper in --type pattern ha_sles
```

4.1.3   Getting the latest updates

If you have installed the packages before, make sure to deploy the newest updates on **all** machines to have the latest versions of the resource agents and other packages. A prerequisite is a valid subscription for SUSE Linux Enterprise Server for SAP Applications. There are multiple ways to get updates via SUSE Manager, the Subscription Management Tool (SMT), or via a direct connection to the SUSE Customer Center (SCC).

Depending on your company or customer rules, use `zypper update` or `zypper patch`.

EXAMPLE 3: SOFTWARE UPDATE MUST BE TRIGGERED FROM EACH NODE

The command `zypper patch` will install all available needed patches. As user root, type:

```
# zypper patch
```

> The command `zypper update` will update all or specified installed packages with newer versions, if possible. As user root, type:

```
# zypper update
```

4.2 Configuring SUSE Linux Enterprise Server for SAP Applications to run SAP HANA

4.2.1 Tuning or modifying

Operating system tuning are described in SAP note 1275776 and 2205917. The SAP note 1275776 explains three ways to implementing the settings.

EXAMPLE 4: USING SAPTUNE (PREFERRED)

```
# saptune solution apply HANA
```

The SAP note 2205917 covers the following modifications:

- SUSE Linux Enterprise Server 12 GA (no SP) and SUSE Linux Enterprise Server 12 SP1 Linux kernel upgrade

- SUSE Linux Enterprise Server 12 SP2 Linux kernel upgrade

- SUSE Linux Enterprise Server 12 SP3 Linux kernel upgrade

- SUSE Linux Enterprise Server 12 SP4 Linux kernel upgrade

- Additional 3rd-party kernel modules

- Configure tuned to use profile "sap-hana" (applies to Intel-based systems only)

- Increase UserTasksMax

- Turn off NUMA balancing

- Disable transparent hugepages

- Configure C-States for lower latency in Linux (applies to Intel-based systems only)

- CPU Frequency/Voltage scaling (applies to Intel-based systems only)

- Energy Performance Bias (EPB, applies to Intel-based systems only)

- Turn off kernel samepage merging (KSM)

- IBM EnergyScale for POWER8 Processor-Based Systems (applies to IBM Power systems only)

- Linux page cache limit

- Network settings

- Intel Cluster-On-Die (COD) / sub-NUMA clustering technology

4.2.2 Enabling SSH access via public key (optional)

Public key authentication provides SSH users access to their servers without entering their passwords. SSH keys are also more secure than passwords, because the private key used to secure the connection is never shared. Private keys can also be encrypted. Their encrypted contents cannot easily be read. For the document at hand, a very simple but useful setup is used. This setup is based on only one ssh-key pair which enables SSH access to all cluster nodes.

Note

 Follow your company security policy to set up access to the systems.

EXAMPLE 5: SSH KEY CREATION AND EXCHANGE

As user root create an SSH key on one node.

```
# ssh-keygen -t rsa
```

The ssh-key generation asks for missing parameters.

```
Generating public/private rsa key pair.
Enter file in which to save the key (/root/.ssh/id_rsa):
Enter passphrase (empty for no passphrase):
Enter same passphrase again:
Your identification has been saved in /root/.ssh/id_rsa.
Your public key has been saved in /root/.ssh/id_rsa.pub.
The key fingerprint is:
SHA256:ip/8kdTbYZNuuEUAdsaYOAErkwnkAPBR7d2SQIpIZCU root@<host1>
The key's randomart image is:
+---[RSA 2048]----+
|XEooo+ooo+o      |
|=+.= o=.o+.      |
|..B o. + o.      |
```

```
|   o  . +... .   |
|       S.. *     |
|     . o . B o   |
|    . . o o =    |
|     o . . +     |
|      +.. .      |
+----[SHA256]-----+
```

After the `ssh-keygen` is set up, you will have two new files under `/root/.ssh/` .

```
# ls /root/.ssh/
id_rsa  id_rsa.pub
```

To allow password-free access for the user root between nodes in the cluster copy `id_rsa.pub` to `authorized_keys` and set the required permissions.

```
# cp /root/.ssh/id_rsa.pub /root/.ssh/authorized_keys
# chmod 600 /root/.ssh/authorized_keys
```

Collect the public host keys from all other node. For the document at hand, the *ssh-keyscan* command is used.

```
# ssh-keyscan
```

The SSH host key is automatically collected and stored in the file `/root/.ssh/known_host` during the first SSH connection. To avoid to confirm the first login with "yes", which accepts the host key, collect and store them beforehand.

```
# ssh-keyscan -t ecdsa-sha2-nistp256 <host1>,<host1 ip> >>.ssh/known_hosts
# ssh-keyscan -t ecdsa-sha2-nistp256 <host2>,<host2 ip> >>.ssh/known_hosts
# ssh-keyscan -t ecdsa-sha2-nistp256 <host3>,<host3 ip> >>.ssh/known_hosts
...
```

After collecting all host keys push the entire directory `/root/.ssh/` from the first node to all further cluster members.

```
# rsync -ay /root/.ssh/ <host2>:/root/.ssh/
# rsync -ay /root/.ssh/ <host3>:/root/.ssh/
# rsync -ay /root/.ssh/ <host4>:/root/.ssh/
...
```

4.2.3   Setting up disk layout for SAP HANA

An SAP certified storage system with a validated storage API is generally recommended. This is a prerequisite of a stable and reliable scale-out installation.

- /hana/shared/<SID>

- /hana/data/<SID>

- /hana/log/<SID>

Create the mount directories on all SAP HANA nodes.

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

The SAP HANA installation needs a special storage setup. The NFS setup used for this guide must be reboot-persistent. You can achieve this with entries in the `/etc/fstab`.

Note

NFS version 4 is required in the setup at hand.

EXAMPLE 6: CREATE PERMANENT MOUNT ENTRIES FOR ALL NFS POOLS

Create */etc/fstab* entries for the three NFS pools.

```
<nfs1>  /hana/data/<SID>       nfs4  defaults  0 0
<nfs2>  /hana/shared/<SID>     nfs4  defaults  0 0
<nfs3> /hana/log/<SID>         nfs4  defaults  0 0
```

In the sample environment those lines are as follows:

```
/exports/HA1_WDF1/shared  /hana/data/HA1     nfs4  defaults  0 0
/exports/HA1_WDF1/data  /hana/shared/HA1    nfs4  defaults  0 0
/exports/HA1_WDF1/log /hana/log/HA1         nfs4  defaults  0 0
```

Mount all NFS shares.

```
# mount -a
```

Create other directories (optional).

```
# mkdir -p /sapsoftware
```

File systems

SAP HANA System Replication Scale-Out - Performance Optimized Scenario

/hana/shared/<SID>

> The mount directory is used for shared files between all hosts in an SAP HANA system. This directory needs to be accessible to each of the servers in the SAP HANA cluster.

/hana/log/<SID>

> The default path to the log directory depends on the system ID of the SAP HANA host.

/hana/data/<SID>

> The default path to the data directory depends on the system ID of the SAP HANA host.

/usr/sap

> This is the path to the local SAP system instance directories. It is possible to join this location with the Linux installation.

/sapsoftware

> (optional) Space for copying the SAP install software media. This NFS pool is mounted on both sites and contains the SAP HANA installation media and installation parameter files.

Set up host name resolution for all machines.

You can either use a DNS server or modify the */etc/hosts* on **all** nodes.

With maintaining the `/etc/hosts` file, you minimize the impact of a failing DNS service. Replace the IP address and the host name in the following commands.

```
vi /etc/hosts
```

Insert the following lines to `/etc/hosts`. Change the IP address and host name to match your environment.

```
192.168.201.151 suse01
192.168.201.152 suse02
...
```

Enable NTP service on all nodes.

Simply enable an **ntp service** on all node in the cluster to have proper time synchronization.

```
yast2 ntp-client
```

5 Installing the SAP HANA databases on both sites

The infrastructure is set up. Now install the SAP HANA database on both sites. In a cluster a machine is also called a *node*.
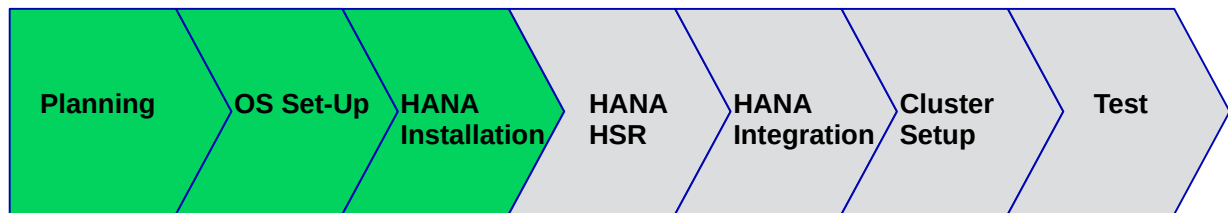
FIGURE 9: *[PLANNING] [OSSETUP]* SAPHANAINST *[SAPHANAHSR] [INTEGRATION] [CLUSTER] [TESTING]*

In the example at hand, to make it easier to follow the documentation, the machines (or nodes) are named *suse01, ... suseXX*. The nodes with odd numbers (suse01, suse03, suse05, ...) will be part of site "A" (WDF1) and the nodes with even numbers (suse02, suse04, suse06, ...) will be part of site "B"(ROT1) .

The following users are automatically created during the SAP HANA installation:

<sid>adm

> The user < sid > adm is the operating system user required for administrative tasks, such as starting and stopping the system.

sapadm

> The SAP Host Agent administrator.

SYSTEM

> The SAP HANA database superuser.

5.1 Preparation

- Read the SAP Installation and Setup Manuals available at the SAP Marketplace.

- Download the SAP HANA Software from SAP Marketplace.

- Mount the file systems to install SAP HANA database software and database content (data and log).

5.2 Installation

1. Mount /hana/shared from the nfs server.

```
# for system in suse0{1,2,3,4,5,6}; do
    ssh $system mount -a
```

SAP HANA System Replication Scale-Out - Performance Optimized Scenario

```
done
```

2. Install the SAP HANA Database as described in the SAP HANA Server Installation Guide on **all** machines (two sites) except the majority maker. Both databases need to have same SID and instance number. You can use either the graphical user interface or the command line installer *hdblcm*. The command line installer can be used in an interactive or batch mode.

EXAMPLE 7: USING HDBLCM IN INTERACTIVE MODE

```
# <path_to_sap_media>/hdblcm
```

Alternatively you can also use the batch mode of *hdblcm*. This can either be done by specifying all needed parameters via the command line or by using a parameter file.

In the example at hand the command line parameters are used. In the batch mode you need to provide an XML password file (here < path >/hana_passwords). A template of this password file can be created with the following command:

EXAMPLE 8: CREATING A PASSWORD FILE

```
# <path_to_sap_media>/hdblcm --dump_configfile_template=templateFile
```

This command creates two files:

- *templateFile* is the template for a parameter file.

- *templateFile.xml* is the XML template used to provide several hana_passwords to the hdblcm installer.

The XML password file looks as follows:

EXAMPLE 9: THE XML PASSWORD TEMPLATE

```
<?xml version="1.0" encoding="UTF-8"?>
<!-- Replace the 3 asterisks with the password -->
<Passwords>
    <root_password><![CDATA[***]]></root_password>
    <sapadm_password><![CDATA[***]]></sapadm_password>
    <master_password><![CDATA[***]]></master_password>
    <sapadm_password><![CDATA[***]]></sapadm_password>
    <password><![CDATA[***]]></password>
    <system_user_password><![CDATA[***]]></system_user_password>
    <streaming_cluster_manager_password><![CDATA[***]]></
streaming_cluster_manager_password>
    <ase_user_password><![CDATA[***]]></ase_user_password>
    <org_manager_password><![CDATA[***]]></org_manager_password>
```

SAP HANA System Replication Scale-Out - Performance Optimized Scenario

```
</Passwords>
```

After having created the XML password file, you can immediately start the SAP HANA installation in batch mode by providing all needed parameters via the command line.

EXAMPLE 10: USING HDBLCM IN BATCH MODE

In the example below the password file is used to provide the password during the installation dialog. All installation parameters are named directly as one command.

```
# cat <path>/hana_passwords | \
<path_to_sap_media>/hdblcm \
  --batch \
  --sid=<SID>\
  --number=<Inst> \
  --action=install \
  --hostname=<node1> \
  --addhosts=<node2>:role=worker,<node3>:role=standby  \
  --certificates_hostmap=<node1>=<node1> \
  --certificates_hostmap=<node2>=<node2> \
  --certificates_hostmap=<node3>=<node3> \
  --install_hostagent \
  --system_usage=test \
  --sapmnt=/hana/shared \
  --datapath=<datapath> \
  --logpath=<logpath> \
  --root_user=root  \
  --workergroup=default \
  --home=/usr/sap/<SID>/home \
  --userid=<uid> \
  --shell=/bin/bash \
  --groupid=<gid> \
  --read_password_from_stdin=xml
```

The second example use the modified template file as answering file.

```
# cat <path>/hana_passwords | \
<path_to_sap_media>/hdblcm \
 -b \
 --configfile=<path_to_templateFile>/<mod_templateFile> \
 --read_password_from_stdin=xml
```

5.3 Checks

Verify that **both** database sites are up and all processes of these databases are running correctly.

1. As Linux user *<sid>adm* use the SAP command line tool *HDB* to get an overview of all running SAP HANA processes. The output of *HDB info* should look like the example below for **both** sites:

EXAMPLE 11: CALLING HDB INFO (AS USER <SID>ADM)

```
~> HDB info
```

The *HDB info* command lists the processes currently running for that SID.

```
USER          PID  ...  COMMAND
ha1adm        6561 ...  -csh
ha1adm        6635 ...    \_ /bin/sh /usr/sap/HA1/HDB00/HDB info
ha1adm        6658 ...      \_ ps fx -U HA1 -o
 user,pid,ppid,pcpu,vsz,rss,args
ha1adm        5442 ...  sapstart pf=/hana/shared/HA1/profile/HA1_HDB00_suse01
ha1adm        5456 ...    \_ /usr/sap/HA1/HDB00/suse01/trace/hdb.sapha1_HDB00
 -d -nw -f /usr/sap/ha1/HDB00/suse
ha1adm        5482 ...        \_ hdbnameserver
ha1adm        5551 ...        \_ hdbpreprocessor
ha1adm        5554 ...        \_ hdbcompileserver
ha1adm        5583 ...        \_ hdbindexserver
ha1adm        5586 ...        \_ hdbstatisticsserver
ha1adm        5589 ...        \_ hdbxsengine
ha1adm        5944 ...        \_ sapwebdisp_hdb pf=/usr/sap/HA1/HDB00}/suse01/
wdisp/sapwebdisp.pfl -f /usr/sap/SL
ha1adm        5363 ...  /usr/sap/HA1/HDB00/exe/sapstartsrv pf=/hana/shared/
HA1/profile/HA1_HDB00_suse02 -D -u s
```

2. Use the python script *landscapeHostConfiguration.py* to show the status of an entire SAP HANA site.

EXAMPLE 12: QUERY THE HOST ROLES (AS USER <SID>ADM)

```
~> HDBSettings.sh landscapeHostConfiguration.py
```

The landscape host configuration is shown with a line per SAP HANA host.

```
| Host   | Host   |... NameServer  | NameServer  | IndexServer | IndexServer
|        | Active |... Config Role | Actual Role | Config Role | Actual Role
| ------ | ------ |... ----------- | ----------- | ----------- | -----------
| suse01 | yes    |... master 1    | master      | worker      | master
| suse03 | yes    |... master 2    | slave       | worker      | slave
| suse05 | yes    |... master 3    | slave       | standby     | standby
```

SAP HANA System Replication Scale-Out - Performance Optimized Scenario

```
        overall host status: ok
```

3. Get an overview of instances of that site (as user <sid>adm)

```
~> sapcontrol -nr <Inst> -function GetSystemInstanceList
```

You should get a list of SAP HANA instances belonging to that site.

```
12.06.2018 17:25:16
GetSystemInstanceList
OK
hostname, instanceNr, httpPort, httpsPort, startPriority, features, dispstatus
suse01, 00, 50013, 50014, 0.3, HDB|HDB_WORKER, GREEN
suse05, 00, 50013, 50014, 0.3, HDB|HDB_WORKER, GREEN
suse03, 00, 50013, 50014, 0.3, HDB|HDB_WORKER, GREEN
```

6 Setting up the SAP HANA system replication

This section describes the setup of the system replication (HSR) after SAP HANA has been installed properly.

Procedure

1. Back up the primary database

2. Enable the primary database

3. Register the secondary database
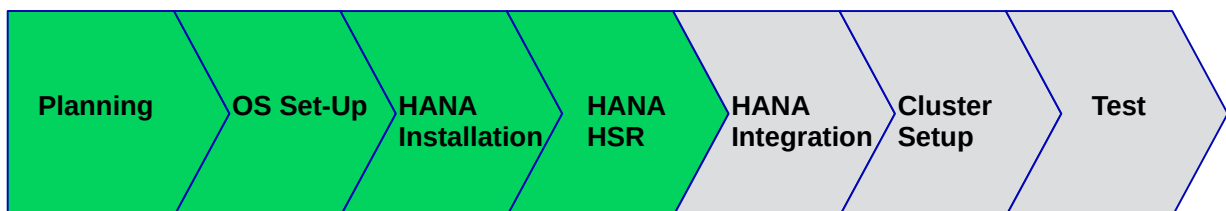
4. Verify the system replication

FIGURE 10: *[PLANNING] [OSSETUP] [SAPHANAINST]* SAPHANAHSR *[INTEGRATION] [CLUSTER] [TESTING]*

For more information read the Section *Setting Up System Replication* of the SAP HANA Administration Guide.

6.1　Backing up the primary database

First back up the primary database as described in the *SAP HANA Administration Guide, Section SAP HANA Database Backup and Recovery*.

Below find examples to back up SAP HANA with SQL Commands:

EXAMPLE 14: SIMPLE BACKUP FOR THE SYSTEM DATABASE AND ALL TENANTS WITH ONE SINGLE BACKUP CALL

As user <sid>adm enter the following command:

```
~> hdbsql -i {refInst} -u SYSTEM -d SYSTEMDB \
    "BACKUP DATA FOR FULL SYSTEM USING FILE ('backup')"
```

You get the following command output (or similar):

```
0 rows affected (overall time 15.352069 sec; server time 15.347745 sec)
```

EXAMPLE 15: SIMPLE BACKUP FOR A SINGLE CONTAINER (NON-MDC) DATABASE

Enter the following command as user <sid>adm:

```
~> hdbsql -i <Inst> -u <dbuser> \
    "BACKUP DATA USING FILE ('backup')"
```

!　Important

Without a valid backup, you **cannot** bring SAP HANA into a system replication configuration.

6.2　Enabling the primary database

As Linux user *<sid>adm* enable the system replication at the primary node. You need to define a site name (like *WDF1*) which must be unique for all SAP HANA databases which are connected via system replication. This means the secondary must have a different site name.

EXAMPLE 16: ENABLE THE SYSTEM REPLICATION ON THE PRIMARY SITE

As user <sid>adm enable the primary:

```
~> hdbnsutil -sr_enable --name=WDF1
```

Check if the command output is similar to:

```
nameserver is active, proceeding ...
successfully enabled system as system replication source site
done.
```

The command line tool *hdbnsutil* can be used to check the system replication mode and site name.

EXAMPLE 17: CHECK THE SYSTEM REPLICATION CONFIGURATION STATUS AS USER <SID>ADM ON THE PRIMARY

```
~> hdbnsutil -sr_stateConfiguration
```

If the system replication enablement was successful at the primary, the output should be as follows:

```
checking for active or inactive nameserver ...
System Replication State
~~~~~~~~~~~~~~~~~~~~~~~

mode: primary
site id: 1
site name: WDF1
done.
```

The mode has changed from "none" to "primary" and the site now has a site name and a site ID.

6.3   Registering the secondary database

The SAP HANA database instance on the secondary side must be stopped before the system can be registered for the system replication. You can use your preferred method to stop the instance (like *HDB* or *sapcontrol*). After the database instance has been stopped successfully, you can register the instance using *hdbnsutil*.

EXAMPLE 18: STOP THE SECONDARY AS LINUX USER *<SID>ADM*:

```
~> sapcontrol -nr <Inst> -function StopSystem
```

EXAMPLE 19: COPY THE KEY AND KEY-DATA FILE FROM THE PRIMARY TO THE SECONDARY SITE

The copy of key and key-data should only be done on the master name server. As the files are in the global file space, you do not need to run the command on all cluster nodes.

```
cd /usr/sap/<SID>/SYS/global/security/rsecssfs
rsync -va {,<node1-siteB>:}$PWD/data/SSFS_<SID>.DAT
rsync -va {,<node1-siteB>:}$PWD/key/SSFS_<SID>.KEY
```

EXAMPLE 20: REGISTER THE SECONDARY AS LINUX USER *<SID>ADM*:

```
~> hdbnsutil -sr_register --name=<site2> \
    --remoteHost=<node1-siteA> --remoteInstance=<Inst> \
    --replicationMode=sync --operationMode=logreplay
```

```
adding site ...
checking for inactive nameserver ...
nameserver suse02:30001 not responding.
collecting information ...
updating local ini files ...
done.
```

The *remoteHost* is the primary node in our case, the *remoteInstance* is the database instance number (here 00).

Now start the database instance again and verify the system replication status. On the secondary site, the mode should be one of „SYNC", „SYNCMEM" or „ASYNC". The mode depends on the **sync** option defined during the registration of the secondary.

EXAMPLE 21: START THE SYSTEM ON THE SECONDARY SITE AS USER <SID>ADM

```
~> sapcontrol -nr <Inst> -function StartSystem
```

Wait until the SAP HANA database is started completely.

EXAMPLE 22: CHECK THE SYSTEM REPLICATION CONFIGURATION AS LINUX USER <SID>ADM

```
~> hdbnsutil -sr_stateConfiguration
```

The output should look like the following:

```
System Replication State
~~~~~~~~~~~~~~~~~~~~~~~~~
mode: sync
site id: 2
site name: ROT1
active primary site: 1

primary masters: suse01 suse03 suse05
```

```
done.
```

6.4 Verifying the system replication

To view the replication state of the whole SAP HANA cluster, use the following command as $<sid>adm$ user on the primary site.

EXAMPLE 23: CHECK THE SYSTEM REPLICATION STATUS AT THE PRIMARY SITE (AS <SID>ADM)

```
~> HDBSettings.sh systemReplicationStatus.py
```

This script prints a human-readable table of the system replication channels and their status. The most interesting column is the **Replication Status**, which should be **ACTIVE**.

```
| Database | Host   | .. Site Name | Secondary | .. Secondary | .. Replication
|          |        | ..           | Host      | .. Site Name | .. Status
| -------- | ------ | .. --------- | --------- | .. --------- | .. ------
| SYSTEMDB | suse01 | .. WDF1      | suse02    | .. ROT1      | .. ACTIVE
| HA1      | suse01 | .. WDF1      | suse02    | .. ROT1      | .. ACTIVE
| HA1      | suse01 | .. WDF1      | suse02    | .. ROT1      | .. ACTIVE
| HA1      | suse03 | .. WDF1      | suse04    | .. ROT1      | .. ACTIVE

status system replication site "2": ACTIVE
overall system replication status: ACTIVE

Local System Replication State

~~~~~~~~~~

mode: PRIMARY
site id: 1
site name: WDF1
```

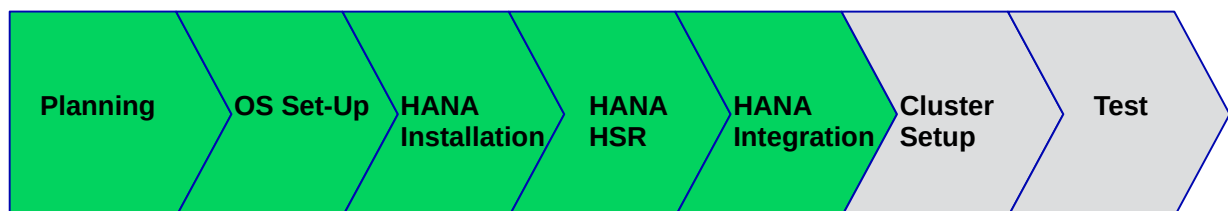# 7 Integrating SAP HANA with the cluster

FIGURE 11: *[PLANNING] [OSSETUP] [SAPHANAINST] [SAPHANAHSR]* INTEGRATION *[CLUSTER] [TESTING]*

Proceed with the following steps:

Procedure

1. Implement the python hook SAPHanaSR

2. Configure system replication operation mode

3. Allow <sid>adm to access the cluster

4. Start SAP HANA

5. Test the hook integration

7.1 Implementing the Python hook SAPHanaSR

This step must be done on both sites. The hook must be available on all SAP HANA nodes. SAP HANA must be stopped to change the *global.ini* and allow SAP HANA to integrate the HA/ DR hook script during start. Use the hook script SAPHanaSR.py from the SAPHanaSR-ScaleOut package. See manual pages SAPHanaSR.py(7) and SAPHanaSR-manageProvider(8) for details.

- Install the HA/DR hook script into a read/writable directory

- Integrate the hook into *global.ini* (SAP HANA needs to be stopped for doing that offline)

- Check integration of the hook during start-up

Note

All hook scripts should be used directly from the SAPHanaSR-ScaleOut package. If the scripts are moved or copied, regular SUSE package updates will not work.

Stop SAP HANA.

```
~> sapcontrol -nr <Inst> -function StopSystem
```

EXAMPLE 24: ADDING SAPHANASR VIA GLOBAL.INI

```
[ha_dr_provider_SAPHanaSR]
provider = SAPHanaSR
path = /usr/share/SAPHanaSR-ScaleOut
```

```
execution_order = 1

[trace]
ha_dr_saphanasr = info
```

7.2 Configuring the system replication operation mode

When your system is connected as an SAPHanaSR target, you can find an entry in the *global.ini* file which defines the operation mode. Up to now there are two modes available:

- *delta_datashipping*

- *logreplay*

Until performing a takeover and re-registration in the opposite direction, the entry for the operation mode is missing on your primary site. The "classic" operation mode is *delta_datashipping*. The preferred mode for HA is *logreplay*. Using the operation mode *logreplay* makes your secondary site in the SAP HANA system replication a hot standby system. For more details regarding both modes check the SAP documentation such as "How To Perform System Replication for SAP HANA" (see https://www.sap.com/documents/2013/10/26c02b58-5a7c-0010-82c7-eda71af511fa.html ↗).

Check both *global.ini* files and add the operation mode, if needed.

section

    [system_replication]

key

    operation_mode = logreplay

Path for the *global.ini*: /hana/shared/<SID>/global/hdb/custom/config/

```
[system_replication]
operation_mode = logreplay
```

7.3 Allowing <sid>adm to access the cluster

The current version of the SAPHanaSR python hook uses the command *sudo* to allow the <sid>adm user to access the cluster attributes. In Linux you can use *visudo* to start the vi editor for the **/etc/sudoers** configuration file.

The user <sid>adm must be able to set the cluster attribute hana_<sid>_glob_srHook. The SAP HANA system replication hook needs password free access. The following example limits the sudo access to exactly setting the needed attribute. See manual page sudoers(5) for details.

Replace the <sid> by the lowercase SAP system ID.

EXAMPLE 25: ENTRY IN SUDO PERMISSIONS /ETC/SUDOERS FILE

Basic parameter option to allow <sidadm> to use the srHook.

```
# SAPHanaSR-ScaleOut needs for srHook
<sid>adm ALL=(ALL) NOPASSWD: /usr/sbin/crm_attribute -n hana_<sid>_glob_srHook -v *
```

More specific parameters option to meet a high security level. All Cmnd_Alias entries must be each defined as a single line entry. In the following example the lines might include a line-break forced by document formatting. The alias identifier needs to be in capitals.

```
# SAPHanaSR-ScaleOut needs for srHook
Cmnd_Alias SOK   = /usr/sbin/crm_attribute -n hana_<sid>_glob_srHook -v SOK   -t
 crm_config -s SAPHanaSR
Cmnd_Alias SFAIL = /usr/sbin/crm_attribute -n hana_<sid>_glob_srHook -v SFAIL -t
 crm_config -s SAPHanaSR
<sid>adm ALL=(ALL) NOPASSWD: SOK, SFAIL
```

EXAMPLE 26: RESULT OF REPLACE <SID> WITH HA1

```
# SAPHanaSR-ScaleOut needs for srHook
ha1adm ALL=(ALL) NOPASSWD: /usr/sbin/crm_attribute -n hana_ha1_glob_srHook -v *
```

7.4 Starting SAP HANA

After having completed the SAP HANA integration and having configured the communication between SAP HANA and the cluster, you can start the SAP HANA databases on both sites.

EXAMPLE 27: STARTING A COMPLETE SAP HANA SITE AS USE <SID>ADM

```
~> sapcontrol -nr <Inst> -function StartSystem
```

The *sapcontrol* service commits the request with OK.

```
11.06.2018 18:30:16
StartSystem
OK
```

Check if SAP HANA has finished starting.

```
~> sapcontrol -nr <Inst> -function WaitforStarted 300 20
```

7.5 Testing the hook integration

When the SAP HANA database has been restarted after the changes, check if the hook script is called correctly.

First check if SAP HANA did create a compiled version of the python script. The file list in `/hana/shared/myHooks` should now also contain a file with the extension **pyc**.

```
cd /hana/shared/myHooks; ll
```

```
-rw-r--r-- 1 <sid>adm sapsys 4890 May  4 14:40 SAPHanaSR.py
-rw-r--r-- 1 <sid>adm sapsys 4932 Jun 11 15:00 SAPHanaSR.pyc
```

A second verification is to check the SAP HANA trace files as <sid>adm:

```
suse01:ha1adm> cdtrace
suse01:ha1adm> awk '/ha_dr_SAPHanaS.*crm_attribute/ \
    { printf "%s %s %s %s\n",$2,$3,$5,$16 }' nameserver_suse01.*
2018-05-04 12:34:04.476445 ha_dr_SAPHanaS...SFAIL
2018-05-04 12:53:06.316973 ha_dr_SAPHanaS...SOK
```

8 Configuring the cluster and SAP HANA resources

This chapter describes the configuration of the SUSE Linux Enterprise High Availability cluster. The SUSE Linux Enterprise High Availability Extension is part of SUSE Linux Enterprise Server for SAP Applications. Further, the integration of SAP HANA System Replication with the SUSE Linux Enterprise High Availability cluster is explained. The integration is done by using the SAPHanaSR-ScaleOut package which is also part of SUSE Linux Enterprise Server for SAP Applications.

FIGURE 12: *[PLANNING] [OSSETUP] [SAPHANAINST] [SAPHANAHSR] [INTEGRATION]* CLUSTER *[TESTING]*

Procedure

1. Installation of cluster packages

2. Basic Cluster Configuration

3. Configure Cluster Properties and Resources

4. Final steps

8.1 Installing the cluster packages

If not already done, install the pattern *High Availability* on **all** nodes.

To do so, use Zypper.

```
# zypper in -t pattern ha_sles
```

Now the Resource Agents for controlling the SAP HANA system replication need to be installed at **all** cluster nodes, including the majority maker.

```
# zypper in SAPHanaSR-ScaleOut
```

If you have the packages installed before, make sure to get the newest updates on **all** nodes

```
zypper patch
```

8.2 Configuring the basic cluster

After having installed the cluster packages, the next step is to set up the basic cluster framework. For convenience, use YaST or the *ha-cluster-init* script.

! Important

It is strongly recommended to add a second corosync ring, implement unicast (UCAST) communication and adjust the timeout values to your environment.

Prerequisites

- Name resolution

- Time synchronization

- Redundant network for cluster intercommunication

- STONITH method

8.2.1 Setting up watchdog for "Storage-based Fencing"

It is recommended to use Storage-based Fencing (SBD) as central STONITH device, as done in the example at hand. Each node constantly monitors connectivity to the storage device, and terminates itself in case the partition becomes unreachable. Whenever SBD is used, a correctly working watchdog is crucial. Modern systems support a hardware watchdog that needs to be "tickled" or "fed" by a software component. The software component (usually a daemon) regularly writes a service pulse to the watchdog. If the daemon stops feeding the watchdog, the hardware will enforce a system restart. This protects against failures of the SBD process itself, such as dying, or getting stuck on an I/O error.

EXAMPLE 28: SET UP FOR WATCHDOG

> **❗ Important**
>
> Access to the watchdog timer: No other software must access the watchdog timer. Some hardware vendors ship systems management software that uses the watchdog for system resets (for example, HP ASR daemon). Disable such software, if watchdog is used by SBD.

> Determine the right watchdog module. Alternatively, you can find a list of installed drivers with your kernel version.
>
> ```
> # ls -l /lib/modules/$(uname -r)/kernel/drivers/watchdog
> ```
>
> Check if any watchdog module is already loaded.
>
> ```
> # lsmod | egrep "(wd|dog|i6|iT|ibm)"
> ```
>
> If you get a result, the system has already a loaded watchdog. If the watchdog does not match your watchdog device, you need to unload the module.

SAP HANA System Replication Scale-Out - Performance Optimized Scenario

To safely unload the module, check first if an application is using the watchdog device.

```
# lsof /dev/watchdog
# rmmod <wrong_module>
```

Enable your watchdog module and make it persistent. For the example below, *softdog* has been used which has some restrictions and should not be used as first option.

```
# echo softdog > /etc/modules-load.d/watchdog.conf
# systemctl restart systemd-modules-load
```

Check if the watchdog module is loaded correctly.

```
# lsmod | grep dog
```

Testing the watchdog can be done with a simple action. Ensure to switch of your SAP HANA first because watchdog will force an unclean reset / shutdown of your system.

In case of a hardware watchdog a desired action is predefined after the timeout of the watchdog has reached. If your watchdog module is loaded and not controlled by any other application, do the following:

> **!** Important
>
> Triggering the watchdog without continuously updating the watchdog resets/switches off the system. This is the intended mechanism. The following commands will force your system to be reset/switched off.

```
# touch /dev/watchdog
```

In case the softdog module is used the following action can be performed:

```
# echo 1> /dev/watchdog
```

After your test was successful you can implement the watchdog on all cluster members. The example below applies to the softdog module. Replace <**wrong_module**> by the module name queried before.

```
# for i in suse{02,03,04,05,06,-mm}; do
    ssh -T $i <<EOSSH
        hostname
        rmmod <wrong_module>
```

```
            echo softdog > /etc/modules-load.d/watchdog.conf
            systemctl restart systemd-modules-load
            lsmod |grep -e dog
EOSSH
done
```

8.2.2   Setting up the initial cluster using *ha-cluster-init*

For more detailed information about ha-cluster-* tools, see section *Overview of the Bootstrap Scripts* of the Installation and Setup Quick Start Guide for SUSE Linux Enterprise High Availability Extension at https://documentation.suse.com/sle-ha/12-SP4/html/SLE-HA-install-quick/art-ha-install-quick.html#sec-ha-inst-quick-bootstrap ↗

Create an initial setup by using *ha-cluster-init* command. Follow the dialog steps.

> ### Note
>
> This is **only** to be done on the **first** cluster node. If you are using SBD as STONITH mechanism, you need to first load the watchdog kernel module matching your setup. In the example at hand the *softdog* kernel module is used.

The command *ha_cluster-init* configures the basic cluster framework including:

- SSH keys

- csync2 to transfer configuration files

- SBD (at least one device)

- corosync (at least one ring)

- HAWK Web interface

```
# ha-cluster-init -u -s <sbd-device>
```

As requested by *ha-cluster-init*, change the password of the user *hacluster* on all cluster nodes.

> ### Note
>
> Do not forget to change the password of the user *hacluster*.

8.2.3　Configuring the cluster for all other cluster nodes

The other nodes of the cluster could be integrated by starting the command *ha-cluster-join*. This command asks for the IP address or name of the **first** cluster node. Then all needed configuration files are copied over. As a result the cluster is started on **all** nodes. Do not forget the majority maker.

If you are using SBD as STONITH method, you need to activate the *softdog* kernel module matching your systems. In the example at hand the *softdog* kernel module is used.

```
# ha-cluster-join -c <host1>
```

8.2.4　Checking the cluster for the first time

Now it is time to check and optionally start the cluster for the first time on all nodes.

> **Note**
>
> All nodes should be started in parallel. Otherwise unseen nodes might get fenced.

Check the cluster status with `crm_mon`. Use the option `-r` to also see resources which are configured but stopped.

```
# crm_mon -r
```

The command will show the "empty" cluster and will print something like the screen output below. The most interesting information in this output is that there are seven nodes in the status "online" and the message "partition with quorum".

```
Stack: corosync
Current DC: suse05 (version 1.1.16-4.8-77ea74d) - partition with quorum
Last updated: Mon Jun 11 16:55:04 2018
Last change: Mon Jun 11 16:53:58 2018 by root via crm_attribute on suse02

7 nodes configured
1 resource configured

Online: [ suse-mm suse01 suse02 suse03 suse04 suse05 suse06 ]

Full list of resources:
```

```
stonith-sbd      (stonith:external/sbd): Started suse-mm
```

8.3 Configuring cluster properties and resources

This section describes how to configure bootstrap, STONITH, resources, and constraints using the *crm* configure shell command as described in section *Configuring and Managing Cluster Resources (Command Line)* of the SUSE Linux Enterprise High Availability Administration Guide (see https://documentation.suse.com/sle-ha/12-SP4/html/SLE-HA-all/cha-ha-manual-config.html ↗).

Use the command *crm* to add the objects to the Cluster Resource Management (CRM). Copy the following examples to a local file and then load the configuration to the Cluster Information Base (CIB). The benefit is that you have a scripted setup and a backup of your configuration.

Perform all *crm* commands only on **one** node, for example on machine suse01.

First write a text file with the configuration, which you load into your cluster in a second step. This step is as follows:

```
# vi crm-file<XX>
# crm configure load update crm-file<XX>
```

8.3.1 Cluster bootstrap and more

The first example defines the cluster bootstrap options including the resource and operation defaults.

The `stonith-timeout` should be greater than 1.2 times the SBD `msgwait` timeout. Find more details and examples in manual page SAPHanaSR-ScaleOut_basic_cluster(7).

```
# vi crm-bs.txt
```

Enter the following to *crm-bs.txt*:

```
property $id="cib-bootstrap-options" \
            no-quorum-policy="freeze" \
            stonith-enabled="true" \
            concurrent-fencing="true" \
            stonith-action="reboot" \
            stonith-timeout="150s"
rsc_defaults $id="rsc-options" \
            resource-stickiness="1000" \
            migration-threshold="50"
op_defaults $id="op-options" \
```

SAP HANA System Replication Scale-Out - Performance Optimized Scenario

```
        timeout="600"
```

Now add the configuration to the cluster.

```
crm configure load update crm-bs.txt
```

8.3.2 STONITH

As already explained in the requirements, STONITH is crucial for a supported cluster setup. Without a valid fencing mechanism your cluster is unsupported.

A standard STONITH mechanism implements SBD based fencing. The SBD STONITH method is very stable and reliable and has proved very good road capability.

You can use other fencing methods available for example from your public cloud provider. However, it is crucial to intensively test the server fencing.

For SBD based fencing you can use one up to three SBD devices. The cluster will react differently when an SBD device is lost. The differences and SBD fencing are explained very well in the SUSE product documentation of the SUSE Linux Enterprise High Availability Extension available at https://documentation.suse.com/ ↗.

You need to adapt the SBD resource for the SAP HANA scale-out cluster.

As user <sid>adm create a file named for crm-fencing.txt.

EXAMPLE 29: CONFIGURE FENCING

```
# vi crm-fencing.txt
```

Enter the following to *crm-fencing.txt*:

```
primitive stonith-sbd stonith:external/sbd \
 params pcmk_action_limit=-1 pcmk_delay_max=1
```

Now load the configuration from the file to the cluster.

```
# crm configure load update crm-fencing.txt
```

8.3.3 Cluster in maintenance mode

Load the configuration for the resources and the constraints step-by-step to the cluster to explain the different parts. The best way to avoid unexpected cluster reactions is to

- first set the complete cluster to maintenance mode,

- then do all needed changes and,

- as last step, end the cluster maintenance mode.

```
# crm configure property maintenance-mode=true
```

8.3.4 SAPHanaTopology

Next, define the group of resources needed, before the SAP HANA instances can be started. Prepare the changes in a text file, for example *crm-saphanatop.txt,* and load these with the *crm* command.

If necessary, change the **SID** and **instance number** (bold) to appropriate values for your setup.

EXAMPLE 30: **CONFIGURE SAPHANATOPOLOGY**

```
suse01:~ # vi crm-saphanatop.txt
```

Enter the following to *crm-saphanatop.txt*:

```
primitive rsc_SAPHanaTop_<SID>_HDB<Inst> ocf:suse:SAPHanaTopology \
    op monitor interval="10" timeout="600" \
    op start interval="0" timeout="600" \
    op stop interval="0" timeout="300" \
    params SID="<SID>" InstanceNumber="<Inst>"

clone cln_SAPHanaTop_<SID>_HDB<Inst> rsc_SAPHanaTop_<SID>_HDB<Inst> \
    meta clone-node-max="1" interleave="true"
```

```
primitive rsc_SAPHanaTop_HA1_HDB00 ocf:suse:SAPHanaTopology \
    op monitor interval="10" timeout="600" \
    op start interval="0" timeout="600" \
    op stop interval="0" timeout="300" \
    params SID="HA1" InstanceNumber="00"

clone cln_SAPHanaTop_HA1_HDB00 rsc_SAPHanaTop_HA1_HDB00 \
    meta clone-node-max="1" interleave="true"
```

For additional information about all parameters, use the command `man ocf_suse_SAPHanaTopology`.

Again, add the configuration to the cluster.

```
# crm configure load update crm-saphanatop.txt
```

The most important parameters here are *SID* (HA1) and *InstanceNumber* (00), which are self explaining in an SAP context.

Beside these parameters, the timeout values or the operations (start, monitor, stop) are typical values to be adjusted for your environment.

8.3.5   SAPHanaController

Next, define the group of resources needed, before the SAP HANA instances can be started. Edit the changes in a text file, for example `crm-saphanacon.txt` and load these with the command `crm.`

```
# vi crm-saphanacon.txt
```

EXAMPLE 31: CONFIGURE SAPHANACONTROLLER

Enter the following to crm-saphanacon.txt

```
primitive rsc_SAPHanaCon_<SID>_HDB<Inst> ocf:suse:SAPHanaController \
    op start interval="0" timeout="3600" \
    op stop interval="0" timeout="3600" \
    op promote interval="0" timeout="3600" \
    op monitor interval="60" role="Master" timeout="700" \
    op monitor interval="61" role="Slave" timeout="700" \
    params SID="<SID>" InstanceNumber="<Inst>" \
        PREFER_SITE_TAKEOVER="true" \
        DUPLICATE_PRIMARY_TIMEOUT="7200" AUTOMATED_REGISTER="false"

ms msl_SAPHanaCon_<SID>_HDB<Inst> rsc_SAPHanaCon_<SID>_HDB<Inst> \
        meta clone-node-max="1" master-max="1" interleave="true"
```

The most important parameters here are <SID> (HA1) and <Inst> (00), which are in the SAP context quite self explaining. Beside these parameters, the timeout values or the operations (start, monitor, stop) are typical tuneables.

```
primitive rsc_SAPHanaCon_HA1_HDB00 ocf:suse:SAPHanaController \
    op start interval="0" timeout="3600" \
    op stop interval="0" timeout="3600" \
    op promote interval="0" timeout="3600" \
    op monitor interval="60" role="Master" timeout="700" \
```

```
      op monitor interval="61" role="Slave" timeout="700" \
      params SID="HA1" InstanceNumber="00" PREFER_SITE_TAKEOVER="true" \
          DUPLICATE_PRIMARY_TIMEOUT="7200" AUTOMATED_REGISTER="false"

 ms msl_SAPHanaCon_HA1_HDB00 rsc_SAPHanaCon_HA1_HDB00 \
      meta clone-node-max="1" master-max="1" interleave="true"
```

Add the configuration to the cluster.

```
# crm configure load update crm-saphanacon.txt
```

TABLE 2: TABLE DESCRIPTION OF IMPORTANT RESOURCE AGENT PARAMETER

Name	Description
PREFER_SITE_TAKEOVER	Defines whether RA should prefer to takeover to the secondary instance instead of restarting the failed primary locally. Set to **true** for SAPHanaSR-ScaleOut.
AUTOMATED_REGISTER	Defines whether a former primary should be automatically registered to be secondary of the new primary. Defaults to **false**.
	If set to **false**, the former primary must be manually registered. The cluster will not start this SAP HANA RDBMS until it is registered to avoid double primary up situations.
DUPLICATE_PRIMARY_TIME-OUT	Time difference needed between two primary time stamps if a dual-primary situation occurs. If the time difference is less than the time gap, the cluster holds one or both sites in a "WAITING" status. This is to give an administrator the chance to react on a failover. If the complete node of the former primary crashed, the former primary will be registered after the time difference is passed. If "only" the SAP HANA RDBMS has crashed, then the former primary will be registered immediately. After this registration to the new primary, all data will be overwritten by the system replication.

Additional information about all parameters can be found with the command `man ocf_suse_SAPHanaController`.

8.3.6 The virtual IP address

The last resource to be added to the cluster is covering the virtual IP address. Replace the bold string with your instance number, SAP HANA system ID and the virtual IP address.

EXAMPLE 32: CONFIGURE THE IP ADDRESS

```
# vi crm-vip.txt
```

Enter the following to *crm-vip.txt*:

```
primitive rsc_ip_<SID>_HDB<Inst> ocf:heartbeat:IPaddr2 \
    op monitor interval="10s" timeout="20s" \
    params ip="<IP>"
```

```
primitive rsc_ip_HA1_HDB00 ocf:heartbeat:IPaddr2 \
    op monitor interval="10s" timeout="20s" \
    params ip="192.168.201.109"
```

Load the file to the cluster.

```
# crm configure load update crm-vip.txt
```

In most installations, only the parameter **ip** needs to be set to the virtual IP address to be presented to the client systems. Use the command `man ocf_heartbeat_IPaddr2` for details on additional parameters.

8.3.7 Constraints

The constraints are organizing the correct placement of the virtual IP address for the client database access and the start order between the two resource agents SAPHana and SAPHanaTopology. The rules help to remove false positive messages from the *crm_mon* command.

EXAMPLE 33: CONFIGURE NEEDED CONSTRAINTS

```
# vi crm-cs.txt
```

Enter the following to *crm-cs.txt*:

```
colocation col_saphana_ip_<SID>_HDB<Inst> 2000: rsc_ip_<SID>_HDB<Inst>:Started \
```

```
     msl_SAPHanaCon_<SID>_HDB<Inst>:Master

order ord_SAPHana_<SID>_HDB<Inst> Optional: cln_SAPHanaTop_<SID>_HDB<Inst> \
     msl_SAPHanaCon_<SID>_HDB<Inst>

location SAPHanaCon_not_on_majority_maker msl_SAPHanaCon_<SID>_HDB<Inst> -inf:
 <majority maker>
location SAPHanaTop_not_on_majority_maker cln_SAPHanaTop_<SID>_HDB<Inst> -inf:
 <majority maker>
```

```
colocation col_saphana_ip_HA1_HDB00 2000: rsc_ip_HA1_HDB00:Started \
     msl_SAPHanaCon_HA1_HDB00:Master

order ord_SAPHana_HA1_HDB00 Optional: cln_SAPHanaTop_HA1_HDB00 \
     msl_SAPHanaCon_HA1_HDB00

location SAPHanaCon_not_on_majority_maker msl_SAPHanaCon_HA1_HDB00 -inf: suse-mm
location SAPHanaTop_not_on_majority_maker cln_SAPHanaTop_HA1_HDB00 -inf: suse-mm
```

Load the file to the cluster.

```
# configure load update crm-cs.txt
```

8.4 Final Steps

8.4.1 Verifying the communication between the hook and the cluster

Now check if the HA/DR provider could set the appropriate cluster attribute hana_<sid>_glob_srHook:

EXAMPLE 34: QUERY THE SRHOOK CLUSTER ATTRIBUTE

```
# crm_attribute -G  -n hana_<sid>_glob_srHook
```

You should get an output similar to the following:

```
scope=crm_config   name=hana_<sid>_glob_srHook value=SFAIL
```

In this case the HA/DR provider sets the attribute to SFAIL to inform the cluster about a broken system replication.

8.4.2 Using special virtual host names or FQHN during the installation of SAP HANA

If you have used special virtual host names or the fully qualified host name (FQHN) instead of the short node name, the resource agents need to map these names. To be able to match the short node name with the used SAP 'virtual host name', the saphostagent needs to report the list of installed instances correctly:

EXAMPLE 35: **IN THE SETUP AT HAND THE VIRTUAL HOST NAME MATCHES THE NODE NAME**

```
suse01:ha1adm> /usr/sap/hostctrl/exe/saphostctrl -function ListInstances
 Inst Info : HA1 - 00 - suse01 - 749, patch 418, changelist 1816226
```

8.4.3 Ending the cluster maintenance mode

After all changes, as last step end the cluster maintenance mode.

```
# crm configure property maintenance-mode=false
```

9 Testing the cluster

Testing is one of the most important project tasks for implementing clusters. Proper testing is crucial. Make sure that all test cases derived from project or customer expectations are defined and passed completely. **Without testing the project is likely to fail in production use**.

FIGURE 13: *[PLANNING] [OSSETUP] [SAPHANAINST] [SAPHANAHSR] [INTEGRATION] [CLUSTER]* TESTING

The test prerequisite, if not described differently, is always that all cluster nodes are booted, are already normal members of the cluster and the SAP HANA RDBMS is running. The system replication is in state **SYNC** represented by 'SOK'. The cluster is idle, no actions are pending, no migration constraints left over, no failcounts left over.

In the current version of this setup guide a plain list of test cases is provided. A more detailed description of the test cases is planned for future versions. These details will be provided either in an updated version of this guide or the test cases will be extracted to a separate test plan document.

9.1   Generic cluster tests

The cluster tests described in this section cover the cluster reaction during operations. This includes starting and stopping the complete cluster or simulating SBD failures and much more.

- Parallel start of all cluster nodes (`systemctl start pacemaker` should be done in a short time frame).

- Stop of the complete cluster.

- Isolate ONE of the two SAP HANA sites.

- Power-off the majority maker.

- Isolate the SBD.

- Simulate a maintenance procedure with cluster continuously running.

- Simulate a maintenance procedure with cluster restart.

- Kill the corosync process of one of the cluster nodes.

9.2   Tests on the primary site

The tests described in this section are checking the reaction on several failures of the primary site.

9.2.1   Tests regarding cluster nodes of the primary site

The tests listed here check the SAP HANA and cluster reaction if one or more nodes of the primary site are failing or re-joining the cluster.

- Power-off master name server of the primary. The test assumes that there is still an SAP HANA standby instance.

- Power-off master name server of the primary after all standby instances are already in use.

- Power-off any worker node but not the master name server of the primary. The test assumes that there is still an SAP HANA standby instance.

- Power-off any worker node but not the master name server of the primary after all standby instances are already in use.

- Power-off any standby node of the primary.

- Re-join of a previously power-off cluster node.

9.2.2 Tests regarding the complete primary site

This test category is simulating a complete site failure.

- Power-off all nodes of the primary site in parallel.

9.2.3 Tests regarding the SAP HANA instances of the primary site

The tests listed here are checks about the SAP HANA and cluster reactions triggered by application failures such as a crashed SAP HANA instance.

- Kill the SAP HANA instance of the master name server of the primary. The test assumes that there is still an SAP HANA standby instance.

- Kill the SAP HANA instance of the master name server of the primary after all standby instances are already in use.

- Kill the SAP HANA instance of any worker node but not the master name server of the primary. The test assumes that there is still an SAP HANA standby instance.

- Kill the SAP HANA instance of any worker node but not the master name server of the primary after all standby instances are already in use.

- Kill the SAP HANA instance of any standby node.

- Kill sapstartsrv of any SAP HANA instance of the primary.

9.3 Tests on the secondary site

The tests described in this section are checking the reaction on several failures of the secondary site.

9.3.1   Tests regarding cluster nodes of the secondary site

The tests listed here check the SAP HANA and cluster reaction if one or more nodes of the secondary site are failing or re-joining the cluster.

- Power-off master name server of the secondary. The test assumes that there is still an SAP HANA standby instance.

- Power-off master name server of the secondary after all standby instances are already in use.

- Power-off any worker node but not the master name server of the secondary. The test assumes that there is still an SAP HANA standby instance.

- Power-off any worker node but not the master name server of the secondary after all standby instances are already in use.

- Power-off any standby node of the secondary.

- Re-join of a previously power-off cluster node.

9.3.2   Tests regarding the complete secondary site

This test category is simulating a complete site failure.

- Power-off all nodes of the secondary site in parallel.

9.3.3   Tests regarding the SAP HANA instances of the secondary site

The tests listed here are checks about the SAP HANA and cluster reactions triggered by application failures such as a crashed SAP HANA instance.

- Kill the SAP HANA instance of the master name server of the secondary. The test assumes that there is still an SAP HANA standby instance.

- Kill the SAP HANA instance of the master name server of the secondary after all standby instances are already in use.

- Kill the SAP HANA instance of any worker node but not the master name server of the secondary. The test assumes that there is still an SAP HANA standby instance.

- Kill the SAP HANA instance of any worker node but not the master name server of the secondary after all standby instances are already in use.

- Kill the SAP HANA instance of any standby node.

- Kill *sapstartsrv* of any SAP HANA instance of the secondary.

10 Administration

10.1 Dos and don'ts

In your project, you should **do** the following:

- Define (and test) STONITH **before** adding other resources to the cluster.

- Do **intensive** testing.

- **Tune** the timeouts of operations of SAPHanaController and SAPHanaTopology.

- Start with **PREFER_SITE_TAKEOVER = true**, **AUTOMATED_REGISTER = false** and **DUPLICATE_PRIMARY_TIMEOUT = "7200"**.

- Always make sure that the cluster configuration does not contain any left-over client-prefer location constraints or failcounts.

- Before testing or beginning maintenance procedures, check if the cluster is in idle state.

In your project, **avoid** the following:

- Rapidly changing/changing back cluster configuration, such as: Setting nodes to standby and online again or stopping/starting the master/slave resource.

- Creating a cluster without proper time synchronization or unstable name resolutions for hosts, users, and groups.

- Adding location rules for the clone, master/slave or IP resource. Only location rules mentioned in this setup guide are allowed.

- As "migrating" or "moving" resources in *crm-shell*, HAWK or other tools would add client-prefer location rules, these activities are completely **forbidden!**.

10.2 Monitoring and tools

You can use the High Availability Web Konsole (HAWK), SAP HANA Studio and different command line tools for cluster status requests.

10.2.1 HAWK – cluster status and more

You can use an Internet browser to check the cluster status. Use the following URL: https://<node>:7630 ↗

The login credentials are provided during the installation dialog of *ha-cluster-init*. Keep in mind to change the default password of the Linux user *hacluster* .
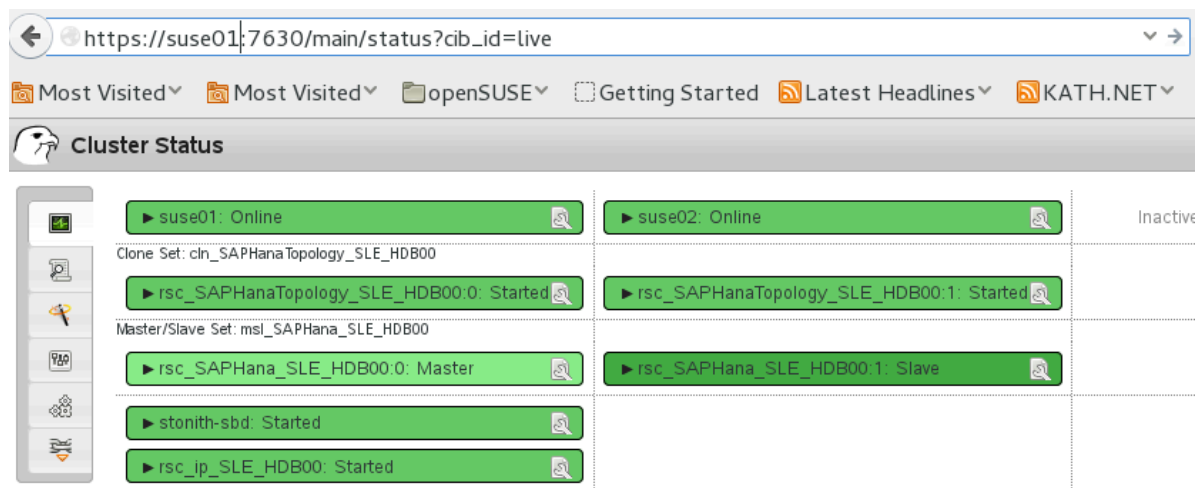

FIGURE 14: CLUSTER STATUS IN HAWK

If you set up the cluster using *ha-cluster-init* and you have installed all packages as described above, your system will provide a very useful Web interface. You can use this graphical Web interface to get an overview of the complete cluster status, perform administrative tasks or even configure resources and cluster bootstrap parameters.

Read the product manuals for a complete documentation of this powerful user interface.

10.2.2 SAP HANA Studio

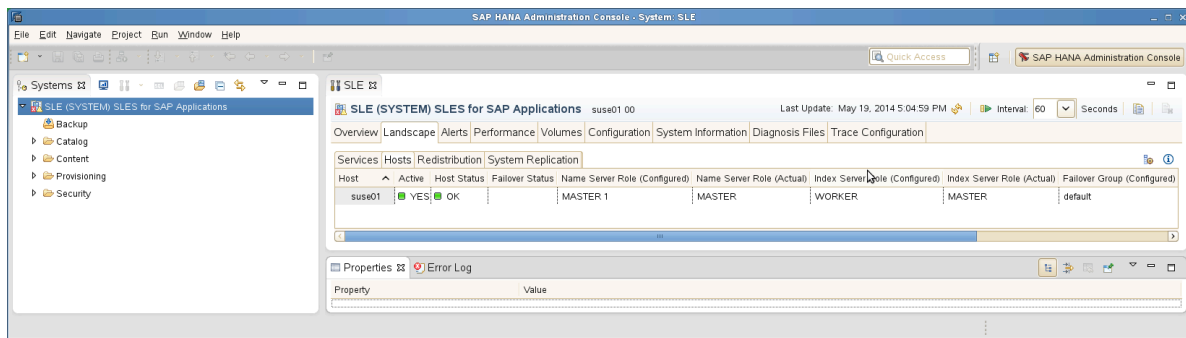Database-specific administration and checks can be done with SAP HANA Cockpit.

FIGURE 15: SAP HANA STUDIO – LANDSCAPE OF A SCALE-OUT SYSTEM

> **Note**
>
> Be extremely careful with changing any parameter or the topology of the system replication. This might cause an interference with the cluster resource management.

A positive example is to register a former primary as new secondary and you have set *AUTOMATED_REGISTER = false*.

A negative example is to un-register a secondary, disable the system replication on the primary, and similar actions.

For all actions that change the system replication it is recommended to first check for the maintenance procedure.

10.2.3   Cluster command line tools

crm_mon

> A simple overview can be obtained by calling *crm_mon*. Using the option *-r* shows also stopped but already configured resources. Option *-1* tells *crm_mon* to output the status once instead of periodically.

```
Stack: corosync
Current DC: suse05 (version 1.1.16-4.8-77ea74d) - partition with quorum
Last updated: Mon Jun 11 16:55:04 2018
Last change: Mon Jun 11 16:53:58 2018 by root via crm_attribute on suse02

7 nodes configured
16 resources configured

Online: [ suse-mm suse01 suse02 suse03 suse04 suse05 suse06 ]
```

```
Full list of resources:

stonith-sbd      (stonith:external/sbd): Started suse-mm
rsc_ip_HA1_HDB00         (ocf::heartbeat:IPaddr2):        Started suse02
 Master/Slave Set: msl_SAPHanaCon_HA1_HDB00 [rsc_SAPHanaCon_HA1_HDB00]
     Masters: [ suse02 ]
     Slaves: [ suse01 suse03 suse04 suse05 suse06 ]
     Stopped: [ suse-mm ]
 Clone Set: cln_SAPHanaTop_HA1_HDB00 [rsc_SAPHanaTop_HA1_HDB00]
     Started: [ suse01 suse02 suse03 suse04 suse05 suse06 ]
     Stopped: [ suse-mm ]
```

See the manual page *crm_mon(8)* for details.

SAPHanaSR-showAttr

To show some SAPHanaController and SAPHanaTopology resource agent internal values, you can call the program *SAPHanaSR-showAttr*. The internal values, storage location and their parameter names may change in the next versions. The command *SAPHanaSR-showAttr* will always fetch the values from the correct storage location.

> ⓘ Important
>
> Do **not** use cluster commands like *crm_attribute* to fetch the values directly from the cluster. Your methods will be broken, when you need to move an attribute to a different storage location or even out of the cluster. *SAPHanaSR-showAttr* is firstly a test program only and should not be used for automated system monitoring.

EXAMPLE 36: CHECK SAPHANASR-SHOWATTR AS USER ROOT

```
suse-mm:~ # SAPHanaSR-showAttr --sid=<SID>
```

The tool displays all interesting cluster attributes in three areas.

- The **global** section includes the information about the cib time stamp and the attributes covering the status of the system replication

- The **site** section includes the attributes per site and shows which site is the primary and the return code of the *landscapeHostConfiguration.py* script. In addition the active master name server is shown per site.

- The **hosts** section includes the node status, the roles of the host inside the SAP HANA database, the calculated score to get the primary master name server and the site name the host belongs to.

```
Global cib-time                  prim sec  srHook sync_state
-------------------------------------------------------------
global Tue Jun 12 15:02:58 2018 WDF1 ROT1 SOK    SOK


Site lpt          lss mns     srr
--------------------------------
WDF1 1528808568 4    suse02 P
ROT1 30         4    suse01 S


Hosts   clone_state node_state roles                       score site
----------------------------------------------------------------------
suse-mm online
suse01  DEMOTED      online      master1:master:worker:master 100 ROT1
suse02  PROMOTED     online      master1:master:worker:master 150 WDF1
suse03  DEMOTED      online      master3:slave:worker:slave   80  ROT1
suse04  DEMOTED      online      master2:slave:worker:slave   110 WDF1
suse05  DEMOTED      online      master2:slave:worker:slave   80  ROT1
suse06  DEMOTED      online      master3:slave:worker:slave   110 WDF1
```

The majority maker *suse-mm* does not run an SAP HANA instance and therefore neither has a role attribute nor a score or site value.

10.2.4   SAP HANA LandscapeHostConfiguration

To check the status of an SAP HANA database and to figure out if the cluster should react, you can use the script `landscapeHostConfiguration.py`.

EXAMPLE 37: CHECK THE LANDSCAPE STATUS AS USER <SID>ADM

```
~> HDBSettings.sh landscapeHostConfiguration.py
```

The landscape host configuration is shown with a line per SAP HANA host.

```
| Host  | Host   | ... NameServer  | NameServer  | IndexServer | IndexServer |
|       | Active | ... Config Role | Actual Role | Config Role | Actual Role |
| ------ | ------ | ... ----------- | ----------- | ----------- | ----------- |
| suse01 | yes    | ... master 1    | master      | worker      | master      |
| suse03 | yes    | ... master 2    | slave       | worker      | slave       |
| suse05 | yes    | ... master 3    | slave       | standby     | standby     |

overall host status: ok
```

Following the SAP HA guideline, the *SAPHana* resource agent interprets the return codes in the following way:

TABLE 3: TABLE INTERPRETATION OF RETURN CODES

Return Code	Description
4	SAP HANA database is up and OK. The cluster does interpret this as correctly running database.
3	SAP HANA database is up and in status INFO. The cluster does interpret this as a correctly running database.
2	SAP HANA database is up and in status warning. The cluster does interpret this as a correctly running database.
1	SAP HANA database is down. If the database should be up and is not own by intention, this could trigger a takeover.
0	Internal Script Error – to be ignored.

11 Useful links, manuals, and SAP Notes

11.1 SUSE Best Practices and more

Best Practices for SUSE Linux Enterprise Server for SAP Applications 12

https://documentation.suse.com/sbp/sap/ ↗

SAP blog article

Fail-Safe Operation of SAP HANA*: SUSE Extends Its High-Availability Solution (http://scn.sap.com/community/hana-in-memory/blog/2014/04/04/fail-safe-operation-of-sap-hana-suse-extends-its-high-availability-solution) ↗

SAP Community Wiki article

HOW TO SET UP SAPHanaSR IN THE COST OPTIMIZED SAP HANA SR SCENARIO (http://scn.sap.com/docs/DOC-65899) ↗

11.2 SUSE product documentation

The SUSE product manuals and documentation can be downloaded at https://documentation.suse.com/ ↗ .

Current online documentation of SUSE Linux Enterprise Server for SAP Applications

   https://documentation.suse.com/sles-sap/12-SP5/ ↗

Current online documentation of SUSE Linux Enterprise High Availability

   https://documentation.suse.com/sle-ha/12-SP5/ ↗

Tuning guide for SUSE Linux Enterprise Server

   https://documentation.suse.com/sles/12-SP5/html/SLES-all/book-sle-tuning.html ↗

Storage admin guide for SUSE Linux Enterprise Server

   https://documentation.suse.com/sles/12-SP5/html/SLES-all/stor-admin.html ↗

Release notes

   https://www.suse.com/releasenotes/ ↗

TID multipath system unable to boot after installing dracut-037-98.2.x86_64

   https://www.suse.com/support/kb/doc/?id=7020912 ↗

TID Systemd-udev-settle timing out

   https://www.suse.com/support/kb/doc/?id=7022681 ↗

TID How to load the correct watchdog kernel module

   http://www.suse.com/support/kb/doc.php?id=7016880 ↗

TID rpcbind won't start after upgrade from SLES 11 to SLES 12

   https://www.suse.com/support/kb/doc/?id=7017144 ↗

TID Memory, I/O and DefaultTasksMax related considerations for SLES for SAP servers with huge memory

   https://www.suse.com/support/kb/doc/?id=7021211 ↗

TID XFS metadata corruption and invalid checksum on SAP Hana servers

   https://www.suse.com/support/kb/doc/?id=7022921 ↗

SUSE Linux Enterprise Server technical information

   https://www.suse.com/products/server/technical-information/ ↗

XFS file system

   https://www.suse.com/communities/conversations/xfs-the-file-system-of-choice/ ↗

11.3 SAP product documentation

SAP HANA Installation and Update Guide

http://help.sap.com/hana/SAP_HANA_Server_Installation_Guide_en.pdf ↗

SAP HANA Administration Guide

http://help.sap.com/hana/SAP_HANA_Administration_Guide_en.pdf ↗

11.4 SAP Notes

As SAP Notes are changing over time, this list is only a starting point

- 611361 Hostnames of SAP ABAP Platform servers

- 1275776 Preparing SLES for SAP Environments

- 1514967 SAP HANA: Central Note

- 1523337 SAP In-Memory Database 1.0: Central Note

- 1501701 Single Computing Unit Performance and Sizing

- 1824819 SAP HANA DB: Recommended OS settings for SLES 11 / SLES for SAP Applications 11 SP4

- 1846872 "No space left on device" error reported from HANA

- 1855805 Recommended SLES 11 packages for HANA support on OS level

- 1867783 XFS Data Inconsistency Bug with SLES 11 SP2

- 1876398 Network configuration for System Replication in HANA SP6

- 1888072 SAP HANA DB: Indexserver crash in strcmp sse42

- 1890444 Slow HANA system due to CPU power save mode

- 1944799 SAP HANA Guidelines for SLES Operating System Installation

- 1954788 SAP HANA DB: Recommended OS settings for SLES 11 / SLES for SAP Applications 11 SP3

- 1984787 SUSE LINUX Enterprise Server 12: Installation notes and

- 1999993 How-To: Interpreting SAP HANA Mini Check Results

- 2000000 FAQ: SAP HANA Performance Optimization

- 2100040 FAQ: SAP HANA CPU

- 2205917 SAP HANA DB: Recommended OS settings for SLES 12 / SLES for SAP Applications 12.

- 2380229 SAP HANA Platform 2.0 – Central Note

- 2470289 FAQ: SAP HANA Non-Uniform Memory Access (NUMA)

- 2647673 HANA Installation Failure

12 Legal notice

13 GNU Free Documentation License

Copyright © 2000, 2001, 2002 Free Software Foundation, Inc. 51 Franklin St, Fifth Floor, Boston, MA 02110-1301 USA. Everyone is permitted to copy and distribute verbatim copies of this license document, but changing it is not allowed.

0. PREAMBLE

The purpose of this License is to make a manual, textbook, or other functional and useful document "free" in the sense of freedom: to assure everyone the effective freedom to copy and redistribute it, with or without modifying it, either commercially or noncommercially. Secondarily, this License preserves for the author and publisher a way to get credit for their work, while not being considered responsible for modifications made by others.

This License is a kind of "copyleft", which means that derivative works of the document must themselves be free in the same sense. It complements the GNU General Public License, which is a copyleft license designed for free software.

We have designed this License in order to use it for manuals for free software, because free software needs free documentation: a free program should come with manuals providing the same freedoms that the software does. But this License is not limited to software manuals; it can be used for any textual work, regardless of subject matter or whether it is published as a printed book. We recommend this License principally for works whose purpose is instruction or reference.

1. APPLICABILITY AND DEFINITIONS

This License applies to any manual or other work, in any medium, that contains a notice placed by the copyright holder saying it can be distributed under the terms of this License. Such a notice grants a world-wide, royalty-free license, unlimited in duration, to use that work under the conditions stated herein. The "Document", below, refers to any such manual or work. Any member of the public is a licensee, and is addressed as "you". You accept the license if you copy, modify or distribute the work in a way requiring permission under copyright law.

A "Modified Version" of the Document means any work containing the Document or a portion of it, either copied verbatim, or with modifications and/or translated into another language.

SAP HANA System Replication Scale-Out - Performance Optimized Scenario

A "Secondary Section" is a named appendix or a front-matter section of the Document that deals exclusively with the relationship of the publishers or authors of the Document to the Document's overall subject (or to related matters) and contains nothing that could fall directly within that overall subject. (Thus, if the Document is in part a textbook of mathematics, a Secondary Section may not explain any mathematics.) The relationship could be a matter of historical connection with the subject or with related matters, or of legal, commercial, philosophical, ethical or political position regarding them.

The "Invariant Sections" are certain Secondary Sections whose titles are designated, as being those of Invariant Sections, in the notice that says that the Document is released under this License. If a section does not fit the above definition of Secondary then it is not allowed to be designated as Invariant. The Document may contain zero Invariant Sections. If the Document does not identify any Invariant Sections then there are none.

The "Cover Texts" are certain short passages of text that are listed, as Front-Cover Texts or Back-Cover Texts, in the notice that says that the Document is released under this License. A Front-Cover Text may be at most 5 words, and a Back-Cover Text may be at most 25 words.

A "Transparent" copy of the Document means a machine-readable copy, represented in a format whose specification is available to the general public, that is suitable for revising the document straightforwardly with generic text editors or (for images composed of pixels) generic paint programs or (for drawings) some widely available drawing editor, and that is suitable for input to text formatters or for automatic translation to a variety of formats suitable for input to text formatters. A copy made in an otherwise Transparent file format whose markup, or absence of markup, has been arranged to thwart or discourage subsequent modification by readers is not Transparent. An image format is not Transparent if used for any substantial amount of text. A copy that is not "Transparent" is called "Opaque".

Examples of suitable formats for Transparent copies include plain ASCII without markup, Texinfo input format, LaTeX input format, SGML or XML using a publicly available DTD, and standard-conforming simple HTML, PostScript or PDF designed for human modification. Examples of transparent image formats include PNG, XCF and JPG. Opaque formats include proprietary formats that can be read and edited only by proprietary word processors, SGML or XML for which the DTD and/or processing tools are not generally available, and the machine-generated HTML, PostScript or PDF produced by some word processors for output purposes only.

The "Title Page" means, for a printed book, the title page itself, plus such following pages as are needed to hold, legibly, the material this License requires to appear in the title page. For works in formats which do not have any title page as such, "Title Page" means the text near the most prominent appearance of the work's title, preceding the beginning of the body of the text.

A section "Entitled XYZ" means a named subunit of the Document whose title either is precisely XYZ or contains XYZ in parentheses following text that translates XYZ in another language. (Here XYZ stands for a specific section name mentioned below, such as "Acknowledgements", "Dedications", "Endorsements", or "History".) To "Preserve the Title" of such a section when you modify the Document means that it remains a section "Entitled XYZ" according to this definition.

The Document may include Warranty Disclaimers next to the notice which states that this License applies to the Document. These Warranty Disclaimers are considered to be included by reference in this License, but only as regards disclaiming warranties: any other implication that these Warranty Disclaimers may have is void and has no effect on the meaning of this License.

2. VERBATIM COPYING

You may copy and distribute the Document in any medium, either commercially or noncommercially, provided that this License, the copyright notices, and the license notice saying this License applies to the Document are reproduced in all copies, and that you add no other conditions whatsoever to those of this License. You may not use technical measures to obstruct or control the reading or further copying of the copies you make or distribute. However, you may accept compensation in exchange for copies. If you distribute a large enough number of copies you must also follow the conditions in section 3.

You may also lend copies, under the same conditions stated above, and you may publicly display copies.

3. COPYING IN QUANTITY

If you publish printed copies (or copies in media that commonly have printed covers) of the Document, numbering more than 100, and the Document's license notice requires Cover Texts, you must enclose the copies in covers that carry, clearly and legibly, all these Cover Texts: Front-Cover Texts on the front cover, and Back-Cover Texts on the back cover. Both covers must also clearly and legibly identify you as the publisher of these copies. The front cover must present the full title with all words of the title equally prominent and visible. You may add other material on the covers in addition. Copying with changes limited to the covers, as long as they preserve the title of the Document and satisfy these conditions, can be treated as verbatim copying in other respects.

If the required texts for either cover are too voluminous to fit legibly, you should put the first ones listed (as many as fit reasonably) on the actual cover, and continue the rest onto adjacent pages.

If you publish or distribute Opaque copies of the Document numbering more than 100, you must either include a machine-readable Transparent copy along with each Opaque copy, or state in or with each Opaque copy a computer-network location from which the general network-using public has access to download using public-standard network protocols a complete Transparent copy of the Document, free of added material. If you use the latter option, you must take reasonably prudent steps, when you begin distribution of Opaque copies in quantity, to ensure that this Transparent copy will remain thus accessible at the stated location until at least one year after the last time you distribute an Opaque copy (directly or through your agents or retailers) of that edition to the public.

It is requested, but not required, that you contact the authors of the Document well before redistributing any large number of copies, to give them a chance to provide you with an updated version of the Document.

4. MODIFICATIONS

You may copy and distribute a Modified Version of the Document under the conditions of sections 2 and 3 above, provided that you release the Modified Version under precisely this License, with the Modified Version filling the role of the Document, thus licensing distribution and modification of the Modified Version to whoever possesses a copy of it. In addition, you must do these things in the Modified Version:

A. Use in the Title Page (and on the covers, if any) a title distinct from that of the Document, and from those of previous versions (which should, if there were any, be listed in the History section of the Document). You may use the same title as a previous version if the original publisher of that version gives permission.

B. List on the Title Page, as authors, one or more persons or entities responsible for authorship of the modifications in the Modified Version, together with at least five of the principal authors of the Document (all of its principal authors, if it has fewer than five), unless they release you from this requirement.

C. State on the Title page the name of the publisher of the Modified Version, as the publisher.

D. Preserve all the copyright notices of the Document.

E. Add an appropriate copyright notice for your modifications adjacent to the other copyright notices.

F. Include, immediately after the copyright notices, a license notice giving the public permission to use the Modified Version under the terms of this License, in the form shown in the Addendum below.

G. Preserve in that license notice the full lists of Invariant Sections and required Cover Texts given in the Document's license notice.

H. Include an unaltered copy of this License.

I. Preserve the section Entitled "History", Preserve its Title, and add to it an item stating at least the title, year, new authors, and publisher of the Modified Version as given on the Title Page. If there is no section Entitled "History" in the Document, create one stating the title, year, authors, and publisher of the Document as given on its Title Page, then add an item describing the Modified Version as stated in the previous sentence.

J. Preserve the network location, if any, given in the Document for public access to a Transparent copy of the Document, and likewise the network locations given in the Document for previous versions it was based on. These may be placed in the "History" section. You may omit a network location for a work that was published at least four years before the Document itself, or if the original publisher of the version it refers to gives permission.

K. For any section Entitled "Acknowledgements" or "Dedications", Preserve the Title of the section, and preserve in the section all the substance and tone of each of the contributor acknowledgements and/or dedications given therein.

L. Preserve all the Invariant Sections of the Document, unaltered in their text and in their titles. Section numbers or the equivalent are not considered part of the section titles.

M. Delete any section Entitled "Endorsements". Such a section may not be included in the Modified Version.

N. Do not retitle any existing section to be Entitled "Endorsements" or to conflict in title with any Invariant Section.

O. Preserve any Warranty Disclaimers.

If the Modified Version includes new front-matter sections or appendices that qualify as Secondary Sections and contain no material copied from the Document, you may at your option designate some or all of these sections as invariant. To do this, add their titles to the list of Invariant Sections in the Modified Version's license notice. These titles must be distinct from any other section titles.

You may add a section Entitled "Endorsements", provided it contains nothing but endorsements of your Modified Version by various parties—for example, statements of peer review or that the text has been approved by an organization as the authoritative definition of a standard.

You may add a passage of up to five words as a Front-Cover Text, and a passage of up to 25 words as a Back-Cover Text, to the end of the list of Cover Texts in the Modified Version. Only one passage of Front-Cover Text and one of Back-Cover Text may be added by (or through arrangements made by) any one entity. If the Document already includes a cover text for the same cover, previously added by you or by arrangement made by the same entity you are acting on behalf of, you may not add another; but you may replace the old one, on explicit permission from the previous publisher that added the old one.

The author(s) and publisher(s) of the Document do not by this License give permission to use their names for publicity for or to assert or imply endorsement of any Modified Version.

5. COMBINING DOCUMENTS

You may combine the Document with other documents released under this License, under the terms defined in section 4 above for modified versions, provided that you include in the combination all of the Invariant Sections of all of the original documents, unmodified, and list them all as Invariant Sections of your combined work in its license notice, and that you preserve all their Warranty Disclaimers.

The combined work need only contain one copy of this License, and multiple identical Invariant Sections may be replaced with a single copy. If there are multiple Invariant Sections with the same name but different contents, make the title of each such section unique by adding at the end of it, in parentheses, the name of the original author or publisher of that section if known, or else a unique number. Make the same adjustment to the section titles in the list of Invariant Sections in the license notice of the combined work.

In the combination, you must combine any sections Entitled "History" in the various original documents, forming one section Entitled "History"; likewise combine any sections Entitled "Acknowledgements", and any sections Entitled "Dedications". You must delete all sections Entitled "Endorsements".

6. COLLECTIONS OF DOCUMENTS

You may make a collection consisting of the Document and other documents released under this License, and replace the individual copies of this License in the various documents with a single copy that is included in the collection, provided that you follow the rules of this License for verbatim copying of each of the documents in all other respects.

You may extract a single document from such a collection, and distribute it individually under this License, provided you insert a copy of this License into the extracted document, and follow this License in all other respects regarding verbatim copying of that document.

7. AGGREGATION WITH INDEPENDENT WORKS

A compilation of the Document or its derivatives with other separate and independent documents or works, in or on a volume of a storage or distribution medium, is called an "aggregate" if the copyright resulting from the compilation is not used to limit the legal rights of the compilation's users beyond what the individual works permit. When the Document is included in an aggregate, this License does not apply to the other works in the aggregate which are not themselves derivative works of the Document.

If the Cover Text requirement of section 3 is applicable to these copies of the Document, then if the Document is less than one half of the entire aggregate, the Document's Cover Texts may be placed on covers that bracket the Document within the aggregate, or the electronic equivalent of covers if the Document is in electronic form. Otherwise they must appear on printed covers that bracket the whole aggregate.

8. TRANSLATION

Translation is considered a kind of modification, so you may distribute translations of the Document under the terms of section 4. Replacing Invariant Sections with translations requires special permission from their copyright holders, but you may include translations of some or all

Invariant Sections in addition to the original versions of these Invariant Sections. You may include a translation of this License, and all the license notices in the Document, and any Warranty Disclaimers, provided that you also include the original English version of this License and the original versions of those notices and disclaimers. In case of a disagreement between the translation and the original version of this License or a notice or disclaimer, the original version will prevail.

If a section in the Document is Entitled "Acknowledgements", "Dedications", or "History", the requirement (section 4) to Preserve its Title (section 1) will typically require changing the actual title.

9. TERMINATION

You may not copy, modify, sublicense, or distribute the Document except as expressly provided for under this License. Any other attempt to copy, modify, sublicense or distribute the Document is void, and will automatically terminate your rights under this License. However, parties who have received copies, or rights, from you under this License will not have their licenses terminated so long as such parties remain in full compliance.

10. FUTURE REVISIONS OF THIS LICENSE

The Free Software Foundation may publish new, revised versions of the GNU Free Documentation License from time to time. Such new versions will be similar in spirit to the present version, but may differ in detail to address new problems or concerns. See http://www.gnu.org/copyleft/ ↗.

Each version of the License is given a distinguishing version number. If the Document specifies that a particular numbered version of this License "or any later version" applies to it, you have the option of following the terms and conditions either of that specified version or of any later version that has been published (not as a draft) by the Free Software Foundation. If the Document does not specify a version number of this License, you may choose any version ever published (not as a draft) by the Free Software Foundation.

ADDENDUM: How to use this License for your documents

```
Copyright (c) YEAR YOUR NAME.
    Permission is granted to copy, distribute and/or modify this document
    under the terms of the GNU Free Documentation License, Version 1.2
```

SAP HANA System Replication Scale-Out - Performance Optimized Scenario

```
or any later version published by the Free Software Foundation;
with no Invariant Sections, no Front-Cover Texts, and no Back-Cover Texts.
A copy of the license is included in the section entitled "GNU
Free Documentation License".
```

If you have Invariant Sections, Front-Cover Texts and Back-Cover Texts, replace the " with… Texts." line with this:

```
with the Invariant Sections being LIST THEIR TITLES, with the
   Front-Cover Texts being LIST, and with the Back-Cover Texts being LIST.
```

If you have Invariant Sections without Cover Texts, or some other combination of the three, merge those two alternatives to suit the situation.

If your document contains nontrivial examples of program code, we recommend releasing these examples in parallel under your choice of free software license, such as the GNU General Public License, to permit their use in free software.