SUSE

SUSE Best Practices

SAP

SAP HANA System Replication Scale-Up - Performance Optimized Scenario

SUSE Linux Enterprise Server for SAP Applications 12 SP4 and later

Fabian Herschel, Distinguished Architect SAP (SUSE)
Bernd Schubert, SAP Solution Architect (SUSE)
Lars Pinne, System Engineer (SUSE)

SUSE

Date: 2024-11-14

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 system replication in the performance optimized scenario. The document focuses on the steps to integrate an already installed and working SAP HANA with system replication.

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 system replication in the performance optimized scenario.

“SAP customers invest in SAP HANA” is the conclusion reached by a recent market study carried out by Pierre Audoin Consultants (PAC). In Germany, half of the companies expect SAP HANA to become the dominant database platform in the SAP environment. Often the “SAP Business Suite* powered by SAP HANA*” scenario is already being discussed in concrete terms.

SUSE is also accommodating this development by providing SUSE Linux Enterprise Server for SAP Applications – the recommended and supported operating system for SAP HANA. In close collaboration with SAP and hardware partners, SUSE provides two resource agents for customers to ensure the high availability of SAP HANA system replications.

1.1.1 Abstract

This guide describes planning, setup, and basic testing of SUSE Linux Enterprise Server for SAP Applications based on the high availability solution scenario "SAP HANA Scale-Up System Replication Performance Optimized".

From the application perspective the following variants are covered:

- plain system replication
- system replication with secondary site read-enabled
- multi-tier (chained) system replication
- multi-target system replication
- multi-tenant database containers for all above

From the infrastructure perspective the following variants are covered:

- 2-node cluster with disk-based SBD
- 3-node cluster with disk-less SBD

1.1.2 Scale-Up Versus Scale-Out

The first set of scenarios includes the architecture and development of *scale-up* solutions.

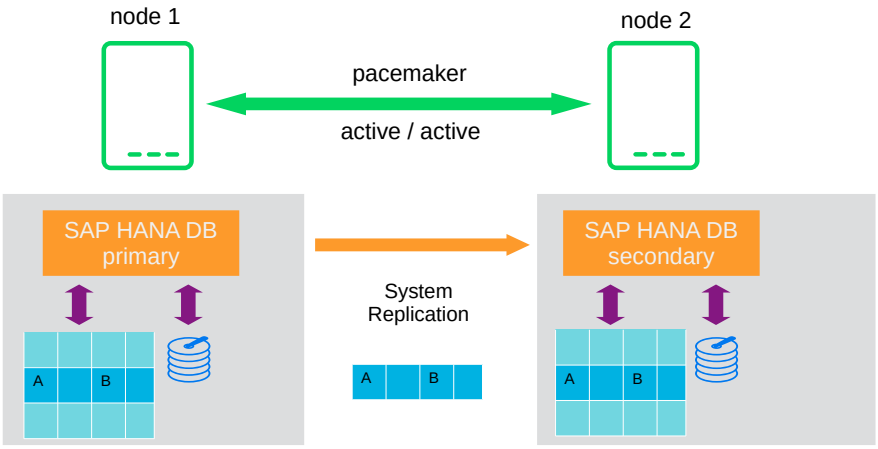


FIGURE 1: SAP HANA SYSTEM REPLICATION SCALE-UP IN THE CLUSTER

For these scenarios SUSE developed the scale-up resource agent package `SAPHanaSR`. System replication will help to replicate the database data from one computer to another computer to compensate for database failures (single-box replication).

The second set of scenarios includes the architecture and development of *scale-out* solutions (multi-box replication). For these scenarios SUSE developed the scale-out resource agent package `SAPHanaSR-ScaleOut`.

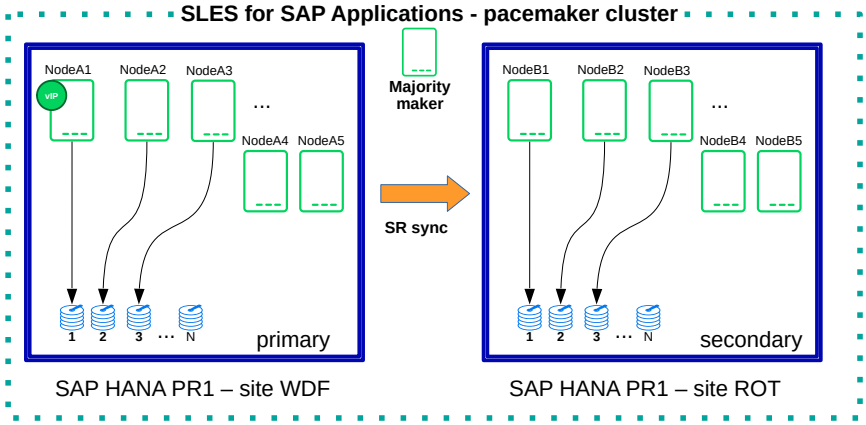


FIGURE 2: SAP HANA SYSTEM REPLICATION SCALE-OUT IN THE CLUSTER

With this mode of operation, internal SAP HANA high availability (HA) mechanisms and the resource agent must work together or be coordinated with each other. SAP HANA system replication automation for scale-out is described in an own document available on our documentation Web page at https://documentation.suse.com/sbp/sap/. The document for scale-out is named *"SAP HANA System Replication Scale-Out - Performance Optimized Scenario"*.

1.1.3 Scale-Up Scenarios and Resource Agents

SUSE has implemented the scale-up scenario with the `SAPHana` resource agent (RA), which performs the actual check of the SAP HANA database instances. This RA is configured as a master/slave resource. In the scale-up scenario, the master assumes responsibility for the SAP HANA databases running in primary mode. The slave is responsible for instances that are operated in synchronous (secondary) status.

To make configuring the cluster as simple as possible, SUSE also developed the `SAPHanaTopology` resource agent. This RA runs on all nodes of a SUSE Linux Enterprise Server for SAP Applications cluster and gathers information about the statuses and configurations of SAP HANA system replications. It is designed as a normal (stateless) clone.

SAP HANA System replication for Scale-Up is supported in the following scenarios or use cases:

- **Performance optimized** ($A \Rightarrow B$). This scenario and setup **is described in this document.**

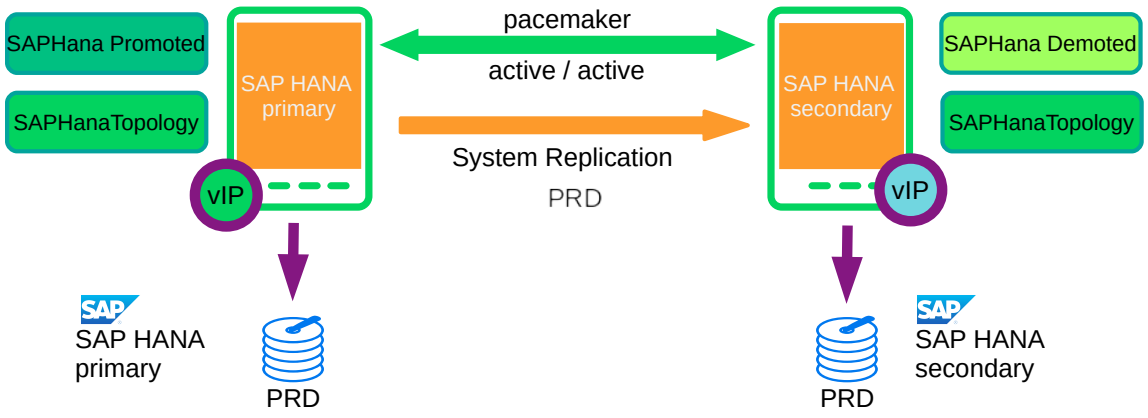


FIGURE 3: SAP HANA SYSTEM REPLICATION SCALE-UP IN THE CLUSTER - PERFORMANCE OPTIMIZED

In the performance optimized scenario an SAP HANA RDBMS site A is synchronizing with an SAP HANA RDBMS site B on a second node. As the HANA RDBMS on the second node is configured to pre-load the tables, the takeover time is typically very short.

One big advance of the performance optimized scenario of SAP HANA is the possibility to allow read access on the secondary database site. To support this **read enabled** scenario, a second virtual IP address is added to the cluster and bound to the secondary role of the system replication.

- **Cost optimized** ($A \Rightarrow B, Q$). This scenario and setup is described in another document available from the documentation Web page (https://documentation.suse.com/sbp/sap/). The document for *cost optimized* is named *"Setting up a SAP HANA SR Cost Optimized Infrastructure"*.

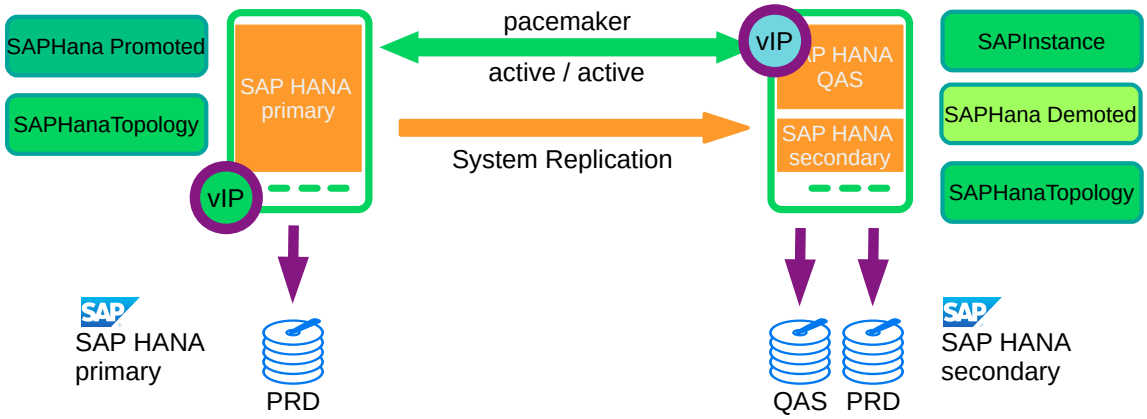


FIGURE 4: SAP HANA SYSTEM REPLICATION SCALE-UP IN THE CLUSTER - COST OPTIMIZED

In the cost optimized scenario the second node is also used for a non-productive SAP HANA RDBMS system (like QAS or TST). Whenever a takeover is needed the non-productive system must be stopped first. As the productive secondary system on this node must be limited in using system resources, the table preload must be switched off. A possible takeover needs longer than in the performance optimized use case.
In the cost optimized scenario the secondary needs to be running in a reduced memory consumption configuration. This why *read enabled* must not be used in this scenario.

- **Multi Tier** ($A \Rightarrow B \rightarrow C$) and **Multi Target** ($B \Leftarrow A \Rightarrow C$).

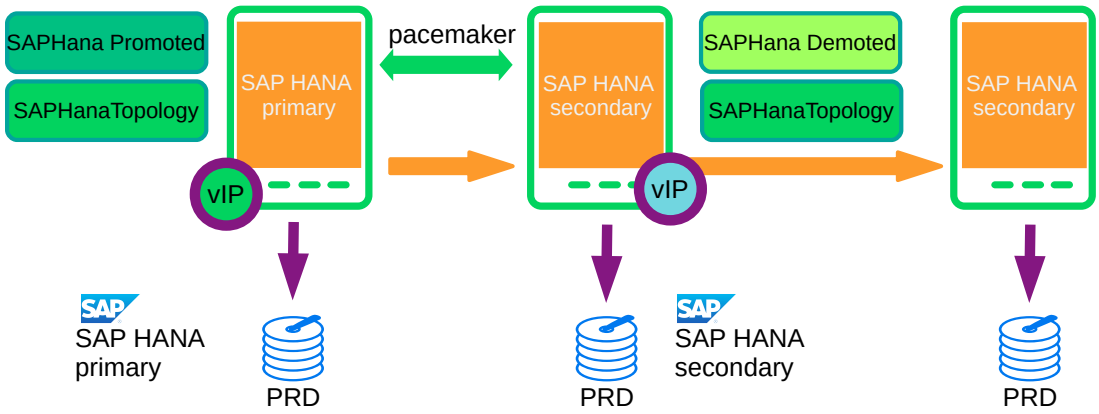


FIGURE 5: SAP HANA SYSTEM REPLICATION SCALE-UP IN THE CLUSTER - PERFORMANCE OPTIMIZED CHAIN

A *Multi Tier* system replication has an additional target. In the past this third side must have been connected to the secondary (chain topology). With current SAP HANA versions also *multiple target topology* is allowed by SAP.

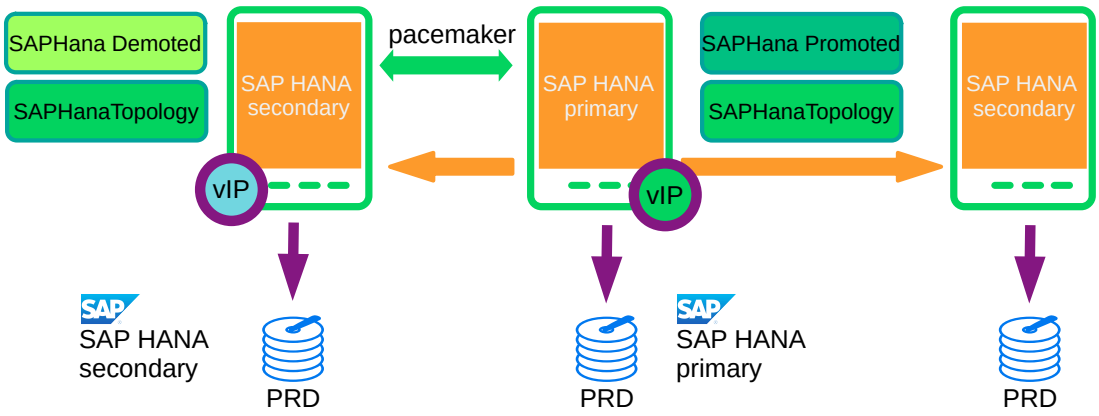


FIGURE 6: SAP HANA SYSTEM REPLICATION SCALE-UP IN THE CLUSTER - PERFORMANCE OPTIMIZED MULTI TARGET

Multi tier and multi target systems are implemented as described in this document. Only the first replication pair (A and B) is handled by the cluster itself. The main difference to the plain performance optimized scenario is that the auto registration must be switched off.

- **Multi-tenancy** or MDC.

Multi-tenancy is supported for all above scenarios and use cases. This scenario is supported since SAP HANA SPS09. The setup and configuration from a cluster point of view is the same for multi-tenancy and single container. Thus you can use the above documents for both kinds of scenarios.

1.1.4 The Concept of the Performance Optimized Scenario

In case of failure of the primary SAP HANA on node 1 (node or database instance) 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.

To achieve an automation of this resource handling process, you must use the SAP HANA resource agents included in SAPHanaSR. System replication of the productive database is automated with SAPHana and SAPHanaTopology.

The cluster only allows a takeover to the secondary site if the SAP HANA system replication was in sync until the point when the service of the primary got lost. This ensures that the last commits processed on the primary site are already available at the secondary site.

SAP did improve the interfaces between SAP HANA and external software such as cluster frameworks. These improvements also include the implementation of SAP HANA call outs in case of special events such as status changes for services or system replication channels. These calls outs are also called HA/DR providers. This interface can be used by implementing SAP HANA hooks written in python. SUSE improved the SAPHanaSR package to include such SAP HANA hooks to optimize the cluster interface. Using the SAP HANA hook described in this document allows to inform the cluster immediately if the SAP HANA system replication breaks. In addition to the SAP HANA hook status the cluster continues to poll the system replication status on regular base.

You can set up 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.

Important

The solution is not designed to manually 'migrate' the primary or secondary instance using HAWK or any other cluster client commands. In the *Administration* section of this document we describe how to 'migrate' the primary to the secondary site using SAP and cluster commands.

1.1.5 Customers Receive Complete Package

Using the SAPHana and SAPHanaTopology resource agents, customers are be able to integrate SAP HANA system replications in their cluster. This has the advantage of enabling companies to use not only their business-critical SAP systems but also their SAP HANA databases without interruption while noticeably reducing needed budgets. SUSE provides the extended solution together with best practices documentation.

SAP and hardware partners who do not have their own SAP HANA high availability solution will also benefit from this development from SUSE.

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 documentation updates, see http://documentation.suse.com/.

You can also find numerous white-papers, best-practices, setup guides, and other resources at the SUSE Linux Enterprise Server for SAP Applications best practices Web page: https://documentation.suse.com/sbp/sap/.

SUSE also publishes blog articles about SAP and high availability. Join us by using the hashtag #TowardsZeroDowntime. Use the following link: https://www.suse.com/c/tag/TowardsZeroDowntime/.

1.3 Errata

To deliver urgent smaller fixes and important information in a timely manner, the Technical Information Document (TID) for this setup guide will be updated, maintained and published at a higher frequency:

- SAP HANA SR Performance Optimized Scenario - Setup Guide - Errata (https://www.suse.com/support/kb/doc/?id=7023882)
- Showing SOK Status in Cluster Monitoring Tools Workaround (https://www.suse.com/support/kb/doc/?id=7023526 - see also the blog article https://www.suse.com/c/lets-flip-the-flags-is-my-sap-hana-database-in-sync-or-not/)

In addition to this guide, check the SUSE SAP Best Practice Guide Errata for other solutions (https://www.suse.com/support/kb/doc/?id=7023713).

1.4 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 Supported Scenarios and Prerequisites

With the `SAPHanaSR` resource agent software package, we limit the support to Scale-Up (single-box to single-box) system replication with the following configurations and parameters:

- Two-node clusters are standard. Three node clusters are fine if you install the resource agents also on that third node. But define in the cluster that SAP HANA resources must never run on that third node. In this case the third node is an additional decision maker in case of cluster separation.
- The cluster must include a valid STONITH method.
  - Any STONITH mechanism supported by SUSE Linux Enterprise 12 High Availability Extension (like SDB, IPMI) is supported with SAPHanaSR.
  - This guide is focusing on the SBD fencing method as this is hardware independent.

    - If you use disk-based SBD as the fencing mechanism, you need one or more shared drives. For productive environments, we recommend more than one SBD device. For details on disk-based SBD, read the product documentation for SUSE Linux Enterprise High Availability Extension and the manual pages sbd.8 and stonith_sbd.8.
    - For disk-less SBD you need at least three cluster nodes. The disk-less SBD mechanism has the benefit that you do not need a shared drive for fencing.
- Both nodes are in the same network segment (layer 2). Similar methods provided by cloud environments such as overlay IP addresses and load balancer functionality are also fine. Follow the cloud specific guides to set up your SUSE Linux Enterprise Server for SAP Applications cluster.
- 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 must be done locally on all cluster nodes.
- Time synchronization between the cluster nodes like NTP.
- Both SAP HANA instances (primary and secondary) have the same SAP Identifier (SID) and instance number.
- If the cluster nodes are installed in different data centers or data center areas, the environment must match the requirements of the SUSE Linux Enterprise High Availability Extension cluster product. Of particular concern are the network latency and recommended maximum distance between the nodes. Review our product documentation for SUSE Linux Enterprise High Availability Extension about those recommendations.
- Automated registration of a failed primary after takeover.
    - As a good starting configuration for projects, we recommend to switch off the automated registration of a failed primary. The setup `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*.
    - For optimal automation, we recommend `AUTOMATED_REGISTER="true"`.
- Automated start of SAP HANA instances during system boot must be switched off.
- Multi-tenancy (MDC) databases are supported.

- Multi-tenancy databases could be used in combination with any other setup (performance based, cost optimized and multi-tier).
- In MDC configurations the SAP HANA RDBMS is treated as a single system including all database containers. Therefore cluster takeover decisions are based on the complete RDBMS status independent of the status of individual database containers.
- For SAP HANA 1.0 you need version SPS10 rev3, SPS11 or newer if you want to stop tenants during production and you want the cluster to be able to take over. Older SAP HANA versions are marking the system replication as failed if you stop a tenant.
- Tests on Multi-tenancy databases could force a different test procedure if you are using strong separation of the tenants. As an example, killing the complete SAP HANA instance using *HDB kill* does not work, because the tenants are running with different Linux user UIDs. <sidadm> is not allowed to terminate the processes of the other tenant users.

You need at least SAPHanaSR version 0.152 and in best SUSE Linux Enterprise Server for SAP Applications 12 SP4 or newer. SAP HANA 1.0 is supported since SPS09 (095) for all mentioned setups. SAP HANA 2.0 is supported with all known SPS versions.

Important

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

If you need to implement a different scenario, we strongly recommend to define a PoC with SUSE. This PoC will focus on testing the existing solution in your scenario. Most of the above mentioned limitations are because careful testing is needed.

Besides SAP HANA, you need SAP Host Agent to be installed on your system.

3 Scope of This Documentation

This document describes how to set up the cluster to control SAP HANA in System Replication Scenarios. The document focuses on the steps to integrate an already installed and working SAP HANA with System Replication.

The described example setup builds an SAP HANA HA cluster in two data centers in Walldorf (WDF) and in Rot (ROT), installed on two SLES for SAP 12 SP4 systems.

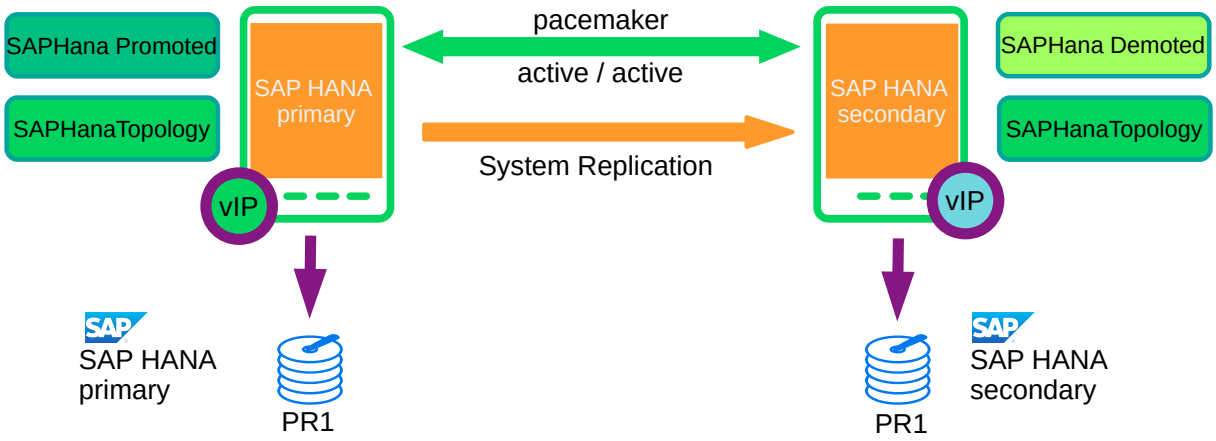


FIGURE 7: CLUSTER WITH SAP HANA SR - PERFORMANCE OPTIMIZED

You can either set up the cluster using the YaST wizard, doing it manually or using your own automation.

If you like to use the YaST wizard, you can use the shortcut *yast sap_ha* to start the module. The procedure to set up SAPHanaSR using YaST is described in the product documentation of SUSE Linux Enterprise Server for SAP Applications in section *Setting Up an SAP HANA Cluster* available at: https://documentation.suse.com/sles-sap/12-SP4/single-html/SLES4SAP-guide/#cha-s4s-cluster

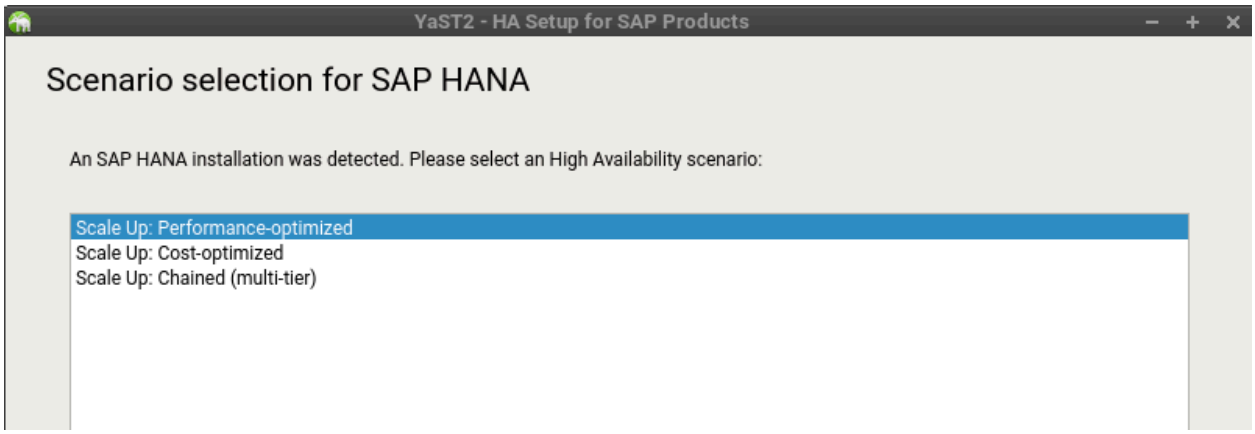
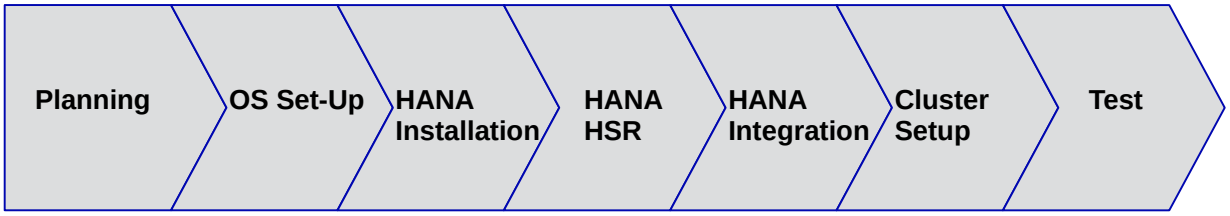


FIGURE 8: SCENARIO SELECTION FOR SAP HANA IN THE YAST MODULE SAP_HA

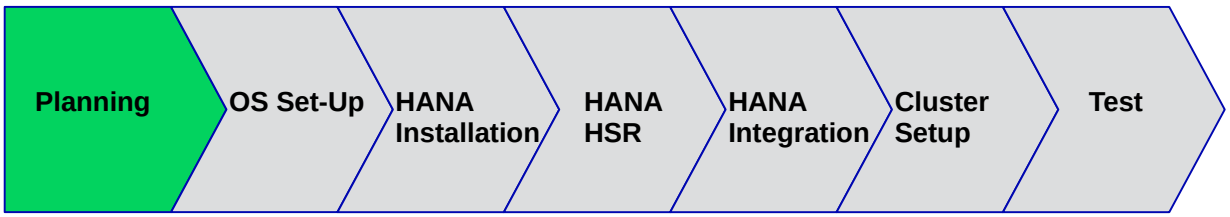
This guide focuses on the manual setup of the cluster to explain the details and to give you the possibility to create your own automation.

The seven main setup steps are:



- Planning (see *Section 4, "Planning the Installation"*)
- OS installation (see *Section 5, "Operating System Setup"*)
- Database installation (see *Section 6, "Installing the SAP HANA Databases on both cluster nodes"*)
- SAP HANA system replication setup (see *Section 7, "Set Up SAP HANA System Replication"*
- SAP HANA HA/DR provider hooks (see *Section 8, "Set Up SAP HANA HA/DR providers"*)
- Cluster configuration (see *Section 9, "Configuration of the Cluster"*)
- Testing (see *Section 10, "Testing the Cluster"*)

4 Planning the Installation



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

What you need before you start:

- Software from SUSE: SUSE Linux Enterprise Server for SAP Applications installation media, a valid subscription, and access to update channels
- Software from SAP: SAP HANA installation media

- Physical or virtual systems including disks
- Filled parameter sheet (see below *Section 4.2, "Parameter Sheet"*)

4.1 Minimum Lab Requirements and Prerequisites

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.

Requirements with 1 SAP instance per site (1 : 1) - without a majority maker (2 node cluster):

- 2 VMs with each 32GB RAM, 50GB disk space for the system
- 1 shared disk for SBD with 10 MB disk space
- 2 data disks (one per site) with a capacity of each 96GB for SAP HANA
- 1 additional IP address for takeover
- 1 optional IP address for the read-enabled setup
- 1 optional IP address for HAWK Administration GUI

Requirements with 1 SAP instance per site (1 : 1) - with a majority maker (3 node cluster):

- 2 VMs with each 32GB RAM, 50GB disk space for the system
- 1 VM with 2GB RAM, 50GB disk space for the system
- 2 data disks (one per site) with a capacity of each 96GB for SAP HANA
- 1 additional IP address for takeover
- 1 optional IP address for the read-enabled setup
- 1 optional IP address for HAWK Administration GUI

4.2 Parameter Sheet

Even if the setup of the cluster organizing two SAP HANA sites is quite simple, the installation should be planned properly. You should have all needed parameters like SID, IP addresses and much more in place. It is good practice to first fill out the parameter sheet and then begin with the installation.

TABLE 1: PARAMETER SHEET FOR PLANNING

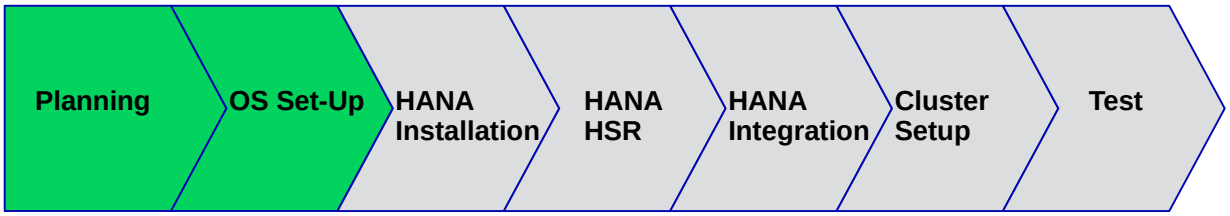
Parameter	Value	Role
node 1		Cluster node name and IP address.
node 2		Cluster node name and IP address.
SID		SAP System Identifier
Instance Number		Number of the SAP HANA database. For system replication also Instance Number + 1 is blocked.
Network mask		
vIP primary		Virtual IP address to be assigned to the primary SAP HANA site
vIP secondary		Virtual IP address to be assigned to the read-enabled secondary SAP HANA site (optional)
Storage		Storage for HDB data and log files is connected “locally” (per node; not shared)
SBD		STONITH device (two for production)
HAWK Port	`7630`	
NTP Server		Address or name of your time server

TABLE 2: PARAMETER SHEET WITH VALUES USED IN THIS DOCUMENT

Parameter	Value	Role
node 1	`suse01`, `192.168.1.11`	Cluster node name and IP address.

Parameter	Value	Role
node 2	`suse02`, `192.168.1.12`	Cluster node name and IP address.
SID	`HA1`	SAP System Identifier
Instance Number	`10`	Number of the SAP HANA database. For system replication also Instance Number + 1 is blocked.
Network mask	`255.255.255.0`	
vIP primary	`192.168.1.20`	
vIP secondary	`192.168.1.21`	(optional)
Storage		Storage for HDB data and log files is connected “locally” (per node; not shared)
SBD	`/dev/disk/by-id/SBDA`	STONITH device (two for production)
HAWK Port	`7630`	
NTP Server	pool pool.ntp.org	Address or name of your time server

5 Operating System Setup



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

For the scope of 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.

5.1 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 well-working system.

5.1.1 Installing 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-SP4/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-SP4/html/SLES-all/book-autoyast.html). The main installation guides for SUSE Linux Enterprise Server for SAP Applications that fit all requirements for SAP HANA are available from 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.

5.1.2 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 **scale-up** is installed. For the **scale-out** 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*.

EXAMPLE 1: INSTALLING ADDITIONAL SOFTWARE FOR THE HA CLUSTER

1. Install the `High Availability` pattern on all nodes

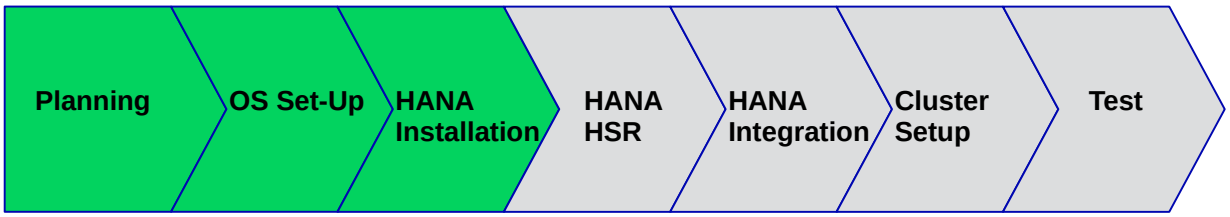
```
suse01:~> zypper in --type pattern ha_sles
```

2. Install the `SAPHanaSR` resource agents on all nodes

```
suse01:~> zypper in SAPHanaSR SAPHanaSR-doc
```

For more information, see *Installation and Basic Setup*, SUSE Linux Enterprise High Availability Extension.

6 Installing the SAP HANA Databases on both cluster nodes



Even though this document focuses on the integration of an installed SAP HANA with system replication already set up into the pacemaker cluster, this chapter summarizes the test environment. Always use the official documentation from SAP to install SAP HANA and to set up the system replication.

6.1 Installing the SAP HANA Databases on both cluster nodes

PREPARATION

- Read the SAP Installation and Setup Manuals available at the SAP Marketplace.
- Download the SAP HANA Software from SAP Marketplace.

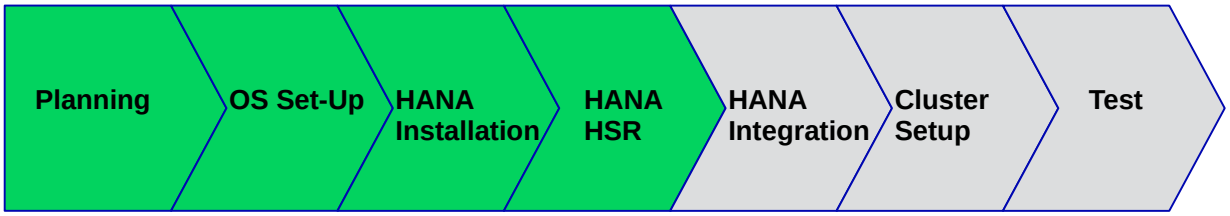
ACTIONS

1. Install the SAP HANA Database as described in the SAP HANA Server Installation Guide.
2. Check if the SAP Host Agent is installed on all cluster nodes. If this SAP service is not installed, install it now.
3. Verify that both databases are up and all processes of these databases are running correctly.

As Linux user *<sid>adm* use the command line tool *HDB* to get an overview of running HANA processes. The output of `HDB info` should be similar to the output shown below:

```
suse02:~> HDB info
USER          PID  ...  COMMAND
ha1adm        6561 ...  -csh
ha1adm        6635 ...    \_ /bin/sh /usr/sap/HA1/HDB10/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_HDB10_suse02
ha1adm        5456 ...   \_ /usr/sap/HA1/HDB10/suse02/trace/hdb.sapHA1_HDB10 -d
-nw -f /usr/sap/HA1/HDB10/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/HDB10/suse02/wdisp/sapwebdisp.pfl -f /usr/sap/SL
ha1adm        5363 ...  /usr/sap/HA1/HDB10/exe/sapstartsrv
pf=/hana/shared/HA1/profile/HA1_HDB10_suse02 -D -u s
```

7 Set Up SAP HANA System Replication



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

Procedure

1. Back up the primary database
2. Enable primary database
3. Register the secondary database
4. Verify the system replication

7.1 Back Up the Primary Database

Back up the primary database as described in the SAP HANA Administration Guide, section *SAP HANA Database Backup and Recovery*. We provide an example with SQL commands. You need to adapt these backup commands to match your backup infrastructure.

EXAMPLE 2: SIMPLE BACKUP FOR THE SYSTEM DATABASE AND ALL TENANTS WITH ONE SINGLE BACKUP CALL

As user <sidadm> enter the following command:

```
hdbsql -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 3: SIMPLE BACKUP FOR A SINGLE CONTAINER (NON MDC) DATABASE

Enter the following command as user <sidadm>:

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

Important

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

7.2 Enable Primary Node

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

Note

Do not use strings like "primary" and "secondary" as site names.

EXAMPLE 4: ENABLE THE PRIMARY

Enable the primary using the -sr_enable option.

```
suse01:~> hdbnsutil -sr_enable --name=WDF
checking local nameserver:
checking for active nameserver ...
nameserver is running, proceeding ...
configuring ini files ...
successfully enabled system as primary site ...
done.
```

EXAMPLE 5: CHECK SR CONFIGURATION ON THE PRIMARY

Check the primary using the command `hdbnsutil -sr_stateConfiguration`.

```
suse01:~> hdbnsutil -sr_stateConfiguration --sapcontrol=1
SAPCONTROL-OK: <begin>
mode=primary
site id=1
site name=WDF
SAPCONTROL-OK: <end>
done.
```

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

7.3 Register the Secondary Node

The SAP HANA database instance on the secondary side must be stopped before the instance 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`. Again, use Linux user *<sid>adm*:

EXAMPLE 6: STOP THE SECONDARY

To stop the secondary you can use the command line tool *HDB*.

```
suse02:~> HDB stop
```

EXAMPLE 7: COPY THE KEY AND KEY-DATA FILE FROM THE PRIMARY TO THE SECONDARY SITE

Beginning with SAP HANA 2.0 the system replication is running encrypted. This is why the key files needs to copied-over from the primary to the secondary site.

```
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 8: REGISTER THE SECONDARY

The registration of the secondary is triggered by calling *hdbnsutil -sr_register* ….

```
...
suse02:~> hdbnsutil -sr_register --name=ROT \
     --remoteHost=suse01 --remoteInstance=10 \
     --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 10).

Now start the database instance again and verify the system replication status. On the secondary node, the mode should be one of "SYNC" or "SYNCMEM". "ASYNC" is also a possible replication mode **but not supported with automated cluster takeover**. The mode depends on the **sync** option defined during the registration of the secondary.

EXAMPLE 9: START SECONDARY AND CHECK SR CONFIGURATION

To start the new secondary use the command line tool *HDB*. Then check the SR configuration using `hdbnsutil -sr_stateConfiguration`.

```
suse02:~> HDB start
...
suse02:~> hdbnsutil -sr_stateConfiguration --sapcontrol=1
SAPCONTROL-OK: <begin>
mode=sync
site id=2
site name=ROT
active primary site=1
primary masters=suse01
SAPCONTROL-OK: <end>
done.
```

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

EXAMPLE 10: CHECKING SYSTEM REPLICATION STATUS DETAILS

The python script *systemReplicationStatus.py* provides details about the current system replication.

```
suse01:~> HDBSettings.sh systemReplicationStatus.py --sapcontrol=1
...
site/2/SITE_NAME=ROT1
site/2/SOURCE_SITE_ID=1
site/2/REPLICATION_MODE=SYNC
site/2/REPLICATION_STATUS=ACTIVE
site/1/REPLICATION_MODE=PRIMARY
site/1/SITE_NAME=WDF1
local_site_id=1
...
```

7.4 Manual Test of SAP HANA SR Takeover

Before you integrate your SAP HANA system replication into the cluster it is mandatory to do a manual takeover. Testing without the cluster helps to make sure that basic operation (takeover and registration) is working as expected.

- Stop SAP HANA on node 1
- Takeover SAP HANA to node 2
- Register node 1 as secondary
- Start SAP HANA on node 1
- Wait till sync state is active

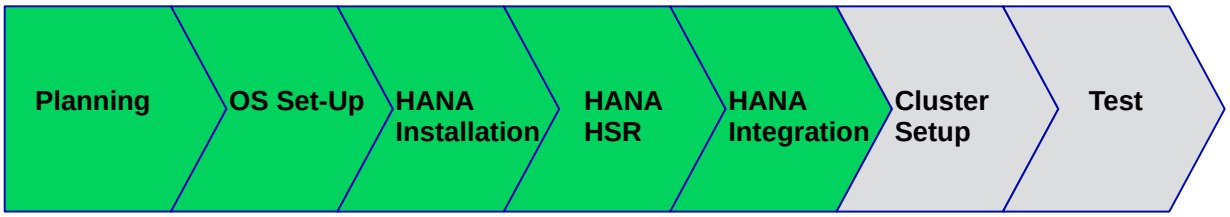
7.5 Optional: Manually Re-Establishing of SAP HANA SR to Original State

Bring the systems back to the original state:

- Stop SAP HANA on node 2
- Takeover SAP HANA to node 1
- Register node 2 as secondary

- Start SAP HANA on node2
- Wait until sync state is active

8 Set Up SAP HANA HA/DR providers



This step is mandatory to inform the cluster immediately if the secondary gets out of sync. The hook is called by SAP HANA using the HA/DR provider interface in point-of-time when the secondary gets out of sync. This is typically the case when the first commit pending is released. The hook is called by SAP HANA again when the system replication is back.

Procedure

1. Implement the python hook SAPHanaSR
2. Configure system replication operation mode
3. Allow <sidadm> to access the cluster
4. Start SAP HANA
5. Test the hook integration

8.1 Implementing the Python Hook SAPHanaSR

This step must be done on both sites. 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 package (available since version 0.153). See manual pages SAPHanaSR.py(7) and SAPHanaSR-manageProvider(8) for details.

- 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 package. If the scripts are moved or copied, regular SUSE package updates will not work.

EXAMPLE 11: STOP SAP HANA

Stop SAP HANA either with *HDB* or using *sapcontrol*.

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

EXAMPLE 12: ADDING SAPHANASR VIA GLOBAL.INI

```
[ha_dr_provider_SAPHanaSR]
provider = SAPHanaSR
path = /usr/share/SAPHanaSR
execution_order = 1

[trace]
ha_dr_saphanasr = info
```

8.2 Configuring System Replication Operation Mode

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

- *delta_datashipping*
- *logreplay*
- *logreplay_readaccess*

Until a takeover and re-registration in the opposite direction, the entry for the operation mode is missing on your primary site. The first operation mode which was available was *delta_datashipping*. Today the preferred modes for HA are *logreplay* or *logreplay_readaccess*. Using the operation mode *logreplay* makes your secondary site in the SAP HANA system replication a hot standby system. For more details regarding all operation modes check the available SAP documentation such as "How To Perform System Replication for SAP HANA ".

EXAMPLE 13: CHECKING THE OPERATION MODE

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

section

[system_replication]

entry

operation_mode = logreplay

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

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

8.3 Allowing <sidadm> to access the Cluster

The current version of the SAPHanaSR python hook uses the command `sudo` to allow the <sidadm> 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 <sidadm> must be able to set the cluster attributes hana_<sid>_site_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 (like `ha1`).

EXAMPLE 14: ENTRY IN SUDO PERMISSIONS /ETC/SUDOERS FILE

Basic sudoers entry to allow <sidadm> to use the srHook.

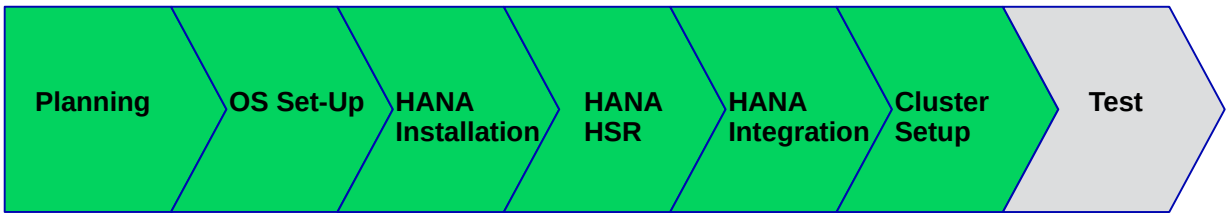
```
# SAPHanaSR-ScaleUp entries for writing srHook cluster attribute
<sidadm> ALL=(ALL) NOPASSWD: /usr/sbin/crm_attribute -n hana_<sid>_site_srHook_*
```

More specific sudoers entries 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. In our example we would have four separate lines with Cmnd_Alias entries, one line for the <sidadm> user and one or more lines for comments. The alias identifier (e.g. SOK_SITEA) needs to be in capitals.

```
# SAPHanaSR-ScaleUp entries for writing srHook cluster attribute
Cmnd_Alias SOK_SITEA   = /usr/sbin/crm_attribute -n hana_<sid>_site_srHook_<siteA> -
v SOK   -t crm_config -s SAPHanaSR
Cmnd_Alias SFAIL_SITEA = /usr/sbin/crm_attribute -n hana_<sid>_site_srHook_<siteA> -
v SFAIL -t crm_config -s SAPHanaSR
Cmnd_Alias SOK_SITEB   = /usr/sbin/crm_attribute -n hana_<sid>_site_srHook_<siteB> -
v SOK   -t crm_config -s SAPHanaSR
```

```
Cmnd_Alias SFAIL_SITEB = /usr/sbin/crm_attribute -n hana_<sid>_site_srHook_<siteB> -
v SFAIL -t crm_config -s SAPHanaSR
<sidadm> ALL=(ALL) NOPASSWD: SOK_SITEA, SFAIL_SITEA, SOK_SITEB, SFAIL_SITEB
```

9 Configuration of the Cluster



This chapter describes the configuration of the cluster software SUSE Linux Enterprise High Availability Extension, which is part of the SUSE Linux Enterprise Server for SAP Applications, and SAP HANA Database Integration.

ACTIONS

1. Basic Cluster Configuration.
2. Configure Cluster Properties and Resources.

9.1 Basic Cluster Configuration

The first step is to set up the basic cluster framework. For convenience, use YaST2 or the ha-cluster-init script. It is strongly recommended to add a second corosync ring, change to UCAST communication and adjust the timeout values to your environment.

9.1.1 Set up Watchdog for "Storage-based Fencing"

If you use the SBD fencing mechanism (disk-less or disk-based), you must also configure a watchdog. The watchdog is needed to reset a node if the system could not longer access the SBD (disk-less or disk-based). It is mandatory that to configure the Linux system to load a watchdog driver. It is strongly recommended to use a watchdog with hardware # assistance (as is available on most modern systems), such as hpwdt, iTCO_wdt, or others. As a fall-back, you can use the softdog module.

EXAMPLE 15: SET UP FOR WATCHDOG

Important

Access to the Watchdog Timer: No other software must access the watchdog timer; it can only be accessed by one process at any time. Some hardware vendors ship systems management software that use the watchdog for system resets (for example HP ASR daemon). Such software must be disabled if the watchdog is to be 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.

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
# ls -l /dev/watchdog
```

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 must implement the watchdog on all cluster members.

9.1.2 Initial Cluster Setup Using `ha-cluster-init`

For more information, see *Automatic Cluster Setup*, SUSE Linux Enterprise High Availability Extension.

Create an initial setup, using `ha-cluster-init` command and follow the dialogs. This has only to be done on the first cluster node.

```
suse01:~> ha-cluster-init -u -s <sbddevice>
```

This command 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

Important

As requested by `ha-cluster-init`, change the password of the user hacluster.

9.1.3 Adapting the Corosync and SBD Configuration

It is recommended to add a second corosync ring. If you did not start ha-cluster-init with the *-u* option, you need to change corosync to use UCAST communication. To change to UCAST stop the already running cluster by using `systemctl stop pacemaker`. After the setup of the corosync configuration and the SBD parameters, start the cluster again.

9.1.3.1 Corosync Configuration

Check the following blocks in the file */etc/corosync/corosync.conf*. See also the example at the end of this document.

```
totem {
    ...

    interface {
        ringnumber: 0
        mcastport:  5405
        ttl:        1
    }
    #Transport protocol
    transport:      udpu

}
nodelist {
        node {
            ring0_addr:     192.168.1.11
            nodeid: 1
        }
        node {
            ring0_addr:     192.168.1.12
            nodeid: 2
        }
    }
```

9.1.3.2 Adapting SBD Config

You can skip this section if you do not have any SBD devices, but be sure to implement another supported fencing mechanism.

See man pages sbd.8 and stonith_sbd.7 for details.

TABLE 3: SBD OPTIONS

Parameter	Description
-W	Use watchdog. It is mandatory to use a watchdog. SBD does not work reliable without watchdog. Refer to the SLES manual and SUSE TIDs 7016880 for setting up a watchdog. This is equivalent to `SBD_WATCHDOG="yes"`
-S 1	Start mode. If set to one, sbd will only start if the node was previously shut down cleanly or if the slot is empty. This is equivalent to `SBD_STARTMODE="clean"`
-P	Check Pacemaker quorum and node health. This is equivalent to `SBD_PACEMAKER="yes"`

In the following, replace /dev/disk/by-id/SBDA and /dev/disk/by-id/SBDB by your real sbd device names.

```
# /etc/sysconfig/sbd
SBD_DEVICE="/dev/disk/by-id/SBDA;/dev/disk/by-id/SBDB"
SBD_WATCHDOG_DEV="/dev/watchdog"
SBD_PACEMAKER="yes"
SBD_STARTMODE="clean"
SBD_OPTS=""
```

9.1.3.3 Verifying the SBD Device

You can skip this section if you do not have any SBD devices, but make sure to implement a supported fencing mechanism.

It is a good practice to check if the SBD device can be accessed from both nodes and does contain valid records. Check this for all devices configured in */etc/sysconfig/sbd*.

```
suse01:~ # sbd -d /dev/disk/by-id/SBDA dump
==Dumping header on disk /dev/disk/by-id/SBDA
```

```
Header version     : 2.1
UUID               : 0f4ea13e-fab8-4147-b9b2-3cdcfff07f86
Number of slots    : 255
Sector size        : 512
Timeout (watchdog) : 20
Timeout (allocate) : 2
Timeout (loop)     : 1
Timeout (msgwait)  : 40
==Header on disk /dev/disk/by-id/SBDA is dumped
```

The timeout values in our sample are only start values, which need to be tuned to your environment.

To check the current SBD entries for the various cluster nodes, you can use `sbd list`. If all entries are `clear`, no fencing task is marked in the SBD device.

```
suse01:~ # sbd -d /dev/disk/by-id/SBDA list
0     suse01      clear
```

For more information on SBD configuration parameters, read the section *Storage-based Fencing*, SUSE Linux Enterprise High Availability Extension and TIDs 7016880 and 7008216.

Now it is time to restart the cluster at the first node again (`systemctl start pacemaker`).

9.1.4 Cluster Configuration on the Second Node

The second node of the two nodes 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. Than all needed configuration files are copied over. As a result the cluster is started on both nodes.

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

9.1.5 Checking the Cluster for the First Time

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

```
suse01:~ # systemctl status pacemaker
suse01:~ # systemctl status sbd
suse02:~ # systemctl status pacemaker
suse01:~ # systemctl start pacemaker
suse02:~ # systemctl status sbd
suse02:~ # systemctl start pacemaker
```

Check the cluster status with crm_mon. We 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 following screen output. The most interesting information for now is that there are two nodes in the status "online" and the message "partition with quorum".

```
Stack: corosync
Current DC: suse01 (version 1.1.19+20180928.0d2680780-1.8-1.1.19+20180928.0d2680780) -
 partition with quorum
Last updated: Fri Nov 29 12:41:16 2019
Last change: Fri Nov 29 12:40:22 2019 by root via crm_attribute on suse02
2 nodes configured
1 resource configured
Online: [ suse01 suse02 ]
Full list of resources:
 stonith-sbd    (stonith:external/sbd): Started suse01
```

9.2 Configuring Cluster Properties and Resources

This section describes how to configure constraints, resources, bootstrap and STONITH using the `crm configure` shell command as described in section *Configuring and Managing Cluster Resources (Command Line)* of the SUSE Linux Enterprise High Availability Extension documentation.

Use the command `crm` to add the objects to CRM. Copy the following examples to a local file, edit the file and then load the configuration to the CIB:

```
suse01:~ # vi crm-fileXX
suse01:~ # crm configure load update crm-fileXX
```

9.2.1 Cluster Bootstrap and More

The first example defines the cluster bootstrap options, the resource and operation defaults. The stonith-timeout should be greater than 1.2 times the SBD msgwait timeout.

```
suse01:~ # vi crm-bs.txt
# enter the following to crm-bs.txt
property $id="cib-bootstrap-options" \
```

```
              stonith-enabled="true" \
              stonith-action="reboot" \
              stonith-timeout="150s"
rsc_defaults $id="rsc-options" \
              resource-stickiness="1000" \
              migration-threshold="5000"
op_defaults $id="op-options" \
                 timeout="600"
```

Now we add the configuration to the cluster.

```
suse01:~ # crm configure load update crm-bs.txt
```

9.2.2 STONITH device

Skip this section if you are using disk-less SBD.

The next configuration part defines an SBD disk STONITH resource.

```
# vi crm-sbd.txt
# enter the following to crm-sbd.txt
primitive stonith-sbd stonith:external/sbd \
    params pcmk_delay_max="15"
```

Again we add the configuration to the cluster.

```
suse01:~ # crm configure load update crm-sbd.txt
```

For fencing with IPMI/ILO see section *Section 9.2.3, "Using IPMI as fencing mechanism"*.

9.2.3 Using IPMI as fencing mechanism

For details about IPMI/ILO fencing see our cluster product documentation (https://documentation.suse.com/sle-ha/12-SP4/html/SLE-HA-all/book-sleha.html). An example for an IPMI STONITH resource can be found in section *Section 13.4, "Example for the IPMI STONITH Method"* of this document.

To use IPMI the remote management boards must be compatible with the IPMI standard.

For the IPMI based fencing you need to configure a primitive per cluster node. Each resource is responsible to fence exactly one cluster node. You need to adapt the IP addresses and login user / password of the remote management boards to the STONITH resource agent. We recommend to create a special STONITH user instead of providing root access to the management board. Location rules must guarantee that a host should never run its own STONITH resource.

9.2.4 Using Other Fencing Mechanisms

We recommend to use SBD (best practice) or IPMI (second choice) as STONITH mechanism. The SUSE Linux Enterprise High Availability product also supports additional fencing mechanism not covered here.

For further information about fencing, see SUSE Linux Enterprise High Availability Guide.

9.2.5 SAPHanaTopology

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

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

```
# vi crm-saphanatop.txt
# enter the following to crm-saphanatop.txt
primitive rsc_SAPHanaTopology_HA1_HDB10 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="10"
clone cln_SAPHanaTopology_HA1_HDB10 rsc_SAPHanaTopology_HA1_HDB10 \
        meta clone-node-max="1" interleave="true"
```

Additional information about all parameters can be found with the command:

```
man ocf_suse_SAPHanaTopology
```

Again we add the configuration to the cluster.

```
suse01:~ # crm configure load update crm-saphanatop.txt
```

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

9.2.6 SAPHana

Next we define the group of resources needed, before the HANA instances can be started. Edit the changes in a text file, for example *crm-saphana.txt*, and load it with the command:

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

TABLE 4: TYPICAL RESOURCE AGENT PARAMETER SETTINGS FOR DIFFERENT SCENARIOS

Parameter	Performance Optimized	Cost Optimized	Multi-Tier
PREFER_SITE_TAKEOVER	true	false	false / true
AUTOMATED_REGISTER	false / true	false / true	false
DUPLICATE_PRIMARY_TIMEOUT	7200	7200	7200

TABLE 5: DESCRIPTION OF IMPORTANT RESOURCE AGENT PARAMETERS

Parameter	Description
PREFER_SITE_TAKEOVER	Defines whether RA should prefer to takeover to the secondary instance instead of restarting the failed primary locally.
AUTOMATED_REGISTER	Defines whether a former primary should be automatically registered to be secondary of the new primary. With this parameter you can adapt the level of system replication automation. 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_TIMEOUT	Time difference needed between two primary time stamps if a dual-primary situation occurs. If the time difference is less than the time gap, than the cluster hold one or both instances in a "WAITING" status. This is to give an administrator the chance to react on a fail-over. 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 pri-

Parameter	Description
	mary 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 could be found with the command:

```
man ocf_suse_SAPHana
```

```
# vi crm-saphana.txt
# enter the following to crm-saphana.txt
primitive rsc_SAPHana_HA1_HDB10 ocf:suse:SAPHana \
        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="10" PREFER_SITE_TAKEOVER="true" \
        DUPLICATE_PRIMARY_TIMEOUT="7200" AUTOMATED_REGISTER="false"
ms msl_SAPHana_HA1_HDB10 rsc_SAPHana_HA1_HDB10 \
        meta clone-max="2" clone-node-max="1" interleave="true"
```

We add the configuration to the cluster.

```
suse01:~ # crm configure load update crm-saphana.txt
```

The most important parameters here are again SID and InstanceNumber. Beside these parameters the timeout values for the operations (start, promote, monitors, stop) are typical tuneables.

9.2.7 The virtual IP address for The Primary Site

The last resource to be added to the cluster is covering the virtual IP address.

```
# vi crm-vip.txt
# enter the following to crm-vip.txt

primitive rsc_ip_HA1_HDB10 ocf:heartbeat:IPaddr2 \
        op monitor interval="10s" timeout="20s" \
        params ip="192.168.1.20"
```

We load the file to the cluster.

```
suse01:~ # 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.

9.2.8 Constraints

Two 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.

```
# vi crm-cs.txt
# enter the following to crm-cs.txt

colocation col_saphana_ip_HA1_HDB10 2000: rsc_ip_HA1_HDB10:Started \
    msl_SAPHana_HA1_HDB10:Master
order ord_SAPHana_HA1_HDB10 Optional: cln_SAPHanaTopology_HA1_HDB10 \
    msl_SAPHana_HA1_HDB10
```

We load the file to the cluster.

```
suse01:~ # crm configure load update crm-cs.txt
```

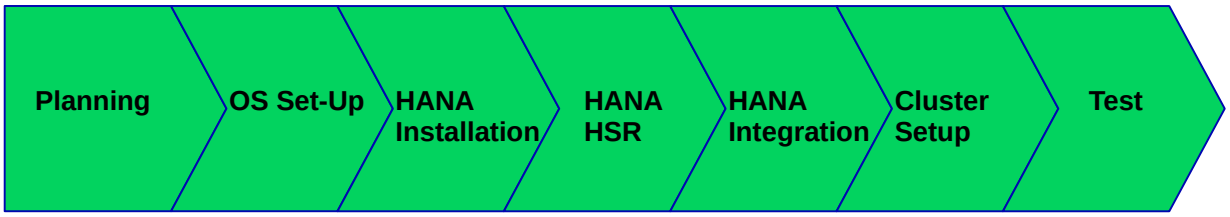
9.2.9 Active/Active Read-Enabled Scenario

This step is optional. If you have an active/active SAP HANA system replication with a read-enabled secondary, it is possible to integrate the needed second virtual IP address into the cluster. This is been done by adding a second virtual IP address resource and a location constraint binding the address to the secondary site.

```
# vi crm-re.txt
# enter the following to crm-re.txt

primitive rsc_ip_HA1_HDB10_readenabled ocf:heartbeat:IPaddr2 \
        op monitor interval="10s" timeout="20s" \
        params ip="192.168.1.21"
colocation col_saphana_ip_HA1_HDB10_readenabled 2000: \
    rsc_ip_HA1_HDB10_readenabled:Started msl_SAPHana_HA1_HDB10:Slave
```

10 Testing the Cluster



The lists of tests will be improved in the next update of this document.

As with any cluster testing is crucial. Make sure that all test cases derived from customer expectations are implemented and passed fully. Otherwise the project is likely to fail in production.

The test prerequisite, if not described differently, is always that both nodes are booted, normal members of the cluster and the HANA RDBMS is running. The system replication is in sync (SOK).

10.1 Test Cases for Semi Automation

In the following test descriptions we assume `PREFER_SITE_TAKEOVER="true"` and `AUTOMATED_REGISTER="false"`.

Note

The following tests are designed to be run in sequence and depend on the exit state of the proceeding tests.

10.1.1 Test: Stop Primary Database on Site A (Node 1)

EXAMPLE 16: TEST STOP_PRIMARY_SITE_A

COMPONENT:

- Primary Database

DESCRIPTION:

- The primary HANA database is stopped during normal cluster operation.

TEST PROCEDURE:

1. Stop the primary HANA database gracefully as *<sid>adm*.

```
suse01# HDB stop
```

RECOVERY PROCEDURE:

1. Manually register the old primary (on node 1) with the new primary after takeover (on node 2) as *<sid>adm*.

```
suse01# hdbnsutil -sr_register --remoteHost=suse02 --remoteInstance=10 \
```

```
          --replicationMode=sync --operationMode=logreplay \
          --name=WDF
```

2. Restart the HANA database (now secondary) on node 1 as root.

```
suse01# crm resource refresh rsc_SAPHana_HA1_HDB10 suse01
```

EXPECTED:

1. The cluster detects the stopped primary HANA database (on node 1) and marks the resource failed.
2. The cluster promotes the secondary HANA database (on node 2) to take over as primary.
3. The cluster migrates the IP address to the new primary (on node 2).
4. After some time the cluster shows the sync_state of the stopped primary (on node 1) as SFAIL.
5. Because AUTOMATED_REGISTER = "false" the cluster does not restart the failed HANA database or register it against the new primary.
6. After the manual register and resource refresh the system replication pair is marked as in sync (SOK).
7. The cluster "failed actions" are cleaned up after following the recovery procedure.

10.1.2 Test: Stop Primary Database on Site B (Node 2)

EXAMPLE 17: TEST STOP_PRIMARY_DB_SITE_B

Component:

Primary Database

Description:

The primary HANA database is stopped during normal cluster operation.

TEST PROCEDURE:

1. Stop the database gracefully as *<sid>adm*.

```
suse02# HDB stop
```

RECOVERY PROCEDURE:

1. Manually register the old primary (on node 2) with the new primary after takeover (on node 1) as *<sid>adm*.

```
suse02# hdbnsutil -sr_register --remoteHost=suse01 --remoteInstance=10 \
               --replicationMode=sync --operationMode=logreplay \
               --name=ROT
```

2. Restart the HANA database (now secondary) on node 2 as root.

```
suse02# crm resource refresh rsc_SAPHana_HA1_HDB10 suse02
```

EXPECTED:

1. The cluster detects the stopped primary HANA database (on node 2) and marks the resource failed.
2. The cluster promotes the secondary HANA database (on node 1) to take over as primary.
3. The cluster migrates the IP address to the new primary (on node 1).
4. After some time the cluster shows the sync_state of the stopped primary (on node 2) as SFAIL.
5. Because AUTOMATED_REGISTER="false" the cluster does not restart the failed HANA database or register it against the new primary.
6. After the manual register and resource refresh the system replication pair is marked as in sync (SOK).
7. The cluster "failed actions" are cleaned up after following the recovery procedure.

10.1.3 Test: Crash Primary Database on Site A (Node 1)

EXAMPLE 18: TEST CRASH_PRIMARY_DB_SITE_A

Component:

Primary Database

Description:

Simulate a complete break-down of the primary database system.

TEST PROCEDURE:

1. Kill the primary database system using signals as *<sid>adm*.

```
suse01# HDB kill-9
```

RECOVERY PROCEDURE:

1. Manually register the old primary (on node 1) with the new primary after takeover (on node 2) as *<sid>adm*.

```
suse01# hdbnsutil -sr_register --remoteHost=suse02 --remoteInstance=10 \
               --replicationMode=sync  --operationMode=logreplay \
               --name=WDF
```

2. Restart the HANA database (now secondary) on node 1 as root.

```
suse01# crm resource refresh rsc_SAPHana_HA1_HDB10 suse01
```

EXPECTED:

1. The cluster detects the stopped primary HANA database (on node 1) and marks the resource failed.
2. The cluster promotes the secondary HANA database (on node 2) to take over as primary.
3. The cluster migrates the IP address to the new primary (on node 2).
4. After some time the cluster shows the sync_state of the stopped primary (on node 1) as SFAIL.
5. Because AUTOMATED_REGISTER="false" the cluster does not restart the failed HANA database or register it against the new primary.
6. After the manual register and resource refresh the system replication pair is marked as in sync (SOK).
7. The cluster "failed actions" are cleaned up after following the recovery procedure.

10.1.4 Test: Crash Primary Database on Site B (Node 2)

EXAMPLE 19: TEST CRASH_PRIMARY_DB_SITE_B

Component:

Primary Database

Description:

Simulate a complete break-down of the primary database system.

TEST PROCEDURE:

1. Kill the primary database system using signals as *<sid>adm*.

```
suse02# HDB kill-9
```

RECOVERY PROCEDURE:

1. Manually register the old primary (on node 2) with the new primary after takeover (on node 1) as *<sid>adm*.

```
suse02# hdbnsutil -sr_register --remoteHost=suse01 --remoteInstance=10 \
               --replicationMode=sync  --operationMode=logreplay \
               --name=ROT
```

2. Restart the HANA database (now secondary) on node 2 as root.

```
suse02# crm resource refresh rsc_SAPHana_HA1_HDB10 suse02
```

EXPECTED:

1. The cluster detects the stopped primary HANA database (on node 2) and marks the resource failed.
2. The cluster promotes the secondary HANA database (on node 1) to take over as primary.
3. The cluster migrates the IP address to the new primary (on node 1).
4. After some time the cluster shows the sync_state of the stopped primary (on node 2) as SFAIL.
5. Because AUTOMATED_REGISTER="false" the cluster does not restart the failed HANA database or register it against the new primary.
6. After the manual register and resource refresh the system replication pair is marked as in sync (SOK).
7. The cluster "failed actions" are cleaned up after following the recovery procedure.

10.1.5 Test: Crash Primary Node on Site A (Node 1)

EXAMPLE 20: TEST CRASH_PRIMARY_NODE_SITE_A

Component:

Cluster node of primary site

Description:

Simulate a crash of the primary site node running the primary HANA database.

TEST PROCEDURE:

1. Crash the primary node by sending a 'fast-reboot' system request.

```
suse01# echo 'b' > /proc/sysrq-trigger
```

RECOVERY PROCEDURE:

1. If SBD fencing is used, pacemaker will not automatically restart after being fenced. In this case clear the fencing flag on all SBD devices and subsequently start pacemaker.

```
suse01# sbd -d /dev/disk/by-id/SBDA message suse01 clear
suse01# sbd -d /dev/disk/by-id/SBDB message suse01 clear
...
```

2. Start the cluster framework

```
suse01# systemctl start pacemaker
```

3. Manually register the old primary (on node 1) with the new primary after takeover (on node 2) as *<sid>adm*.

```
suse01# hdbnsutil -sr_register --remoteHost=suse02 --remoteInstance=10 \
               --replicationMode=sync  --operationMode=logreplay \
               --name=WDF
```

4. Restart the HANA database (now secondary) on node 1 as root.

```
suse01# crm resource refresh rsc_SAPHana_HA1_HDB10 suse01
```

EXPECTED:

1. The cluster detects the failed node (node 1) and declares it UNCLEAN and sets the secondary node (node 2) to status "partition with quorum".
2. The cluster fences the failed node (node 1).

3. The cluster declares the failed node (node 1) OFFLINE.
4. The cluster promotes the secondary HANA database (on node 2) to take over as primary.
5. The cluster migrates the IP address to the new primary (on node 2).
6. After some time the cluster shows the sync_state of the stopped primary (on node 2) as SFAIL.
7. If SBD fencing is used, then the manual recovery procedure will be used to clear the fencing flag and restart pacemaker on the node.
8. Because AUTOMATED_REGISTER = "false" the cluster does not restart the failed HANA database or register it against the new primary.
9. After the manual register and resource refresh the system replication pair is marked as in sync (SOK).
10. The cluster "failed actions" are cleaned up after following the recovery procedure.

10.1.6 Test: Crash Primary Node on Site B (Node 2)

EXAMPLE 21: TEST CRASH_PRIMARY_NODE_SITE_B

Component:

Cluster node of secondary site

Description:

Simulate a crash of the secondary site node running the primary HANA database.

TEST PROCEDURE:

1. Crash the secondary node by sending a 'fast-reboot' system request.

```
suse02# echo 'b' > /proc/sysrq-trigger
```

RECOVERY PROCEDURE:

1. If SBD fencing is used, pacemaker will not automatically restart after being fenced. In this case clear the fencing flag on all SBD devices and subsequently start pacemaker.

```
suse02# sbd -d /dev/disk/by-id/SBDA message suse02 clear
```

```
suse02# sbd -d /dev/disk/by-id/SBDB message suse02 clear
...
```

2. Start the cluster Framework

```
suse02# systemctl start pacemaker
```

3. Manually register the old primary (on node 2) with the new primary after takeover (on node 1) as *<sid>adm*.

```
suse02# hdbnsutil -sr_register --remoteHost=suse01 --remoteInstance=10 \
               --replicationMode=sync  --operationMode=logreplay \
               --name=ROT
```

4. Restart the HANA database (now secondary) on node 2 as root.

```
suse02# crm resource refresh rsc_SAPHana_HA1_HDB10 suse02
```

EXPECTED:

1. The cluster detects the failed secondary node (node 2) and declares it UNCLEAN and sets the primary node (node 1) to status "partition with quorum".
2. The cluster fences the failed secondary node (node 2).
3. The cluster declares the failed secondary node (node 2) OFFLINE.
4. The cluster promotes the secondary HANA database (on node 1) to take over as primary.
5. The cluster migrates the IP address to the new primary (on node 1).
6. After some time the cluster shows the sync_state of the stopped secondary (on node 2) as SFAIL.
7. If SBD fencing is used, then the manual recovery procedure will be used to clear the fencing flag and restart pacemaker on the node.
8. Because AUTOMATED_REGISTER = "false" the cluster does not restart the failed HANA database or register it against the new primary.
9. After the manual register and resource refresh the system replication pair is marked as in sync (SOK).
10. The cluster "failed actions" are cleaned up after following the recovery procedure.

10.1.7 Test: Stop the Secondary Database on Site B (Node 2)

EXAMPLE 22: TEST STOP_SECONDARY_DB_SITE_B

Component:

Secondary HANA database

Description:

The secondary HANA database is stopped during normal cluster operation.

TEST PROCEDURE:

1. Stop the secondary HANA database gracefully as *<sid>adm*.

```
suse02# HDB stop
```

RECOVERY PROCEDURE:

1. Refresh the failed resource status of the secondary HANA database (on node 2) as root.

```
suse02# crm resource refresh rsc_SAPHana_HA1_HDB10 suse02
```

EXPECTED:

1. The cluster detects the stopped secondary database (on node 2) and marks the resource failed.
2. The cluster detects the broken system replication and marks it as failed (SFAIL).
3. The cluster restarts the secondary HANA database on the same node (node 2).
4. The cluster detects that the system replication is in sync again and marks it as ok (SOK).
5. The cluster "failed actions" are cleaned up after following the recovery procedure.

10.1.8 Test: Crash the Secondary Database on Site B (Node 2)

EXAMPLE 23: TEST CRASH_SECONDARY_DB_SITE_B

Component:

Secondary HANA database

Description:

Simulate a complete break-down of the secondary database system.

TEST PROCEDURE:

1. Kill the secondary database system using signals as *<sid>adm*.

```
suse02# HDB kill-9
```

RECOVERY PROCEDURE:

1. Clean up the failed resource status of the secondary HANA database (on node 2) as root.

```
suse02# crm resource refresh rsc_SAPHana_HA1_HDB10 suse02
```

EXPECTED:

1. The cluster detects the stopped secondary database (on node 2) and marks the resource failed.
2. The cluster detects the broken system replication and marks it as failed (SFAIL).
3. The cluster restarts the secondary HANA database on the same node (node 2).
4. The cluster detects that the system replication is in sync again and marks it as ok (SOK).
5. The cluster "failed actions" are cleaned up after following the recovery procedure.

10.1.9 Test: Crash the Secondary Node on Site B (Node2)

EXAMPLE 24: TEST CRASH_SECONDARY_NODE_SITE_B

Component:

Cluster node of secondary site

Description:

Simulate a crash of the secondary site node running the secondary HANA database.

TEST PROCEDURE:

1. Crash the secondary node by sending a 'fast-reboot' system request.

```
suse02# echo 'b' > /proc/sysrq-trigger
```

RECOVERY PROCEDURE:

1. If SBD fencing is used, pacemaker will not automatically restart after being fenced. In this case clear the fencing flag on **all** SBD devices and subsequently start pacemaker.

```
suse02# sbd -d /dev/disk/by-id/SBDA message suse02 clear
suse02# sbd -d /dev/disk/by-id/SBDB message suse02 clear
...
```

2. Start the cluster framework.

```
suse02# systemctl start pacemaker
```

EXPECTED:

1. The cluster detects the failed secondary node (node 2) and declares it UNCLEAN and sets the primary node (node 1) to status "partition with quorum".
2. The cluster fences the failed secondary node (node 2).
3. The cluster declares the failed secondary node (node 2) OFFLINE.
4. After some time the cluster shows the sync_state of the stopped secondary (on node 2) as SFAIL.
5. If SBD fencing is used, then the manual recovery procedure will be used to clear the fencing flag and restart pacemaker on the node.
6. When the fenced node (node 2) rejoins the cluster the former secondary HANA database is started automatically.
7. The cluster detects that the system replication is in sync again and marks it as ok (SOK).

10.1.10 Test: Failure of Replication LAN

EXAMPLE 25: TEST FAIL_NETWORK_SR

Component:

Replication LAN

Description:

Loss of replication LAN connectivity between the primary and secondary node.

TEST PROCEDURE:

1. Break the connection between the cluster nodes on the replication LAN.

RECOVERY PROCEDURE:

1. Re-establish the connection between the cluster nodes on the replication LAN.

EXPECTED:

1. After some time the cluster shows the sync_state of the secondary (on node 2) as SFAIL.
2. The primary HANA database (node 1) "HDBSettings.sh systemReplicationStatus.py" shows "CONNECTION TIMEOUT" and the secondary HANA database (node 2) is not able to reach the primary database (node 1).
3. The primary HANA database continues to operate as “normal”, but no system replication takes place and is therefore no longer a valid take over destination.
4. When the LAN connection is re-established, HDB automatically detects connectivity between the HANA databases and restarts the system replication process
5. The cluster detects that the system replication is in sync again and marks it as ok (SOK).

10.2 Test Cases for Full Automation

In the following test descriptions we assume **PREFER_SITE_TAKEOVER = "true"** and **AUTOMATED_REGISTER = "true".**

Note

The following tests are designed to be run in sequence and depend on the exit state of the proceeding tests.

10.2.1 Test: Stop Primary Database on Site A

EXAMPLE 26: TEST STOP_PRIMARY_DB_SITE_A

COMPONENT:

- Primary Database

DESCRIPTION:

- The primary HANA database is stopped during normal cluster operation.

TEST PROCEDURE:

- Stop the primary HANA database gracefully as *<sid>adm*.

```
suse01# HDB stop
```

RECOVERY PROCEDURE:

1. Not needed, everything is automated
2. Refresh the cluster resources on node 1 as root.

```
suse01# crm resource refresh rsc_SAPHana_HA1_HDB10 suse01
```

EXPECTED:

1. The cluster detects the stopped primary HANA database (on node 1) and marks the resource failed.
2. The cluster promotes the secondary HANA database (on node 2) to take over as primary.
3. The cluster migrates the IP address to the new primary (on node 2).
4. After some time the cluster shows the sync_state of the stopped primary (on node 1) as SFAIL.
5. Because AUTOMATED_REGISTER="true" the cluster does restart the failed HANA database and register it against the new primary.
6. After the automated register and resource refresh the system replication pair is marked as in sync (SOK).
7. The cluster "failed actions" are cleaned up after following the recovery procedure.

10.2.2 Test: Crash the Primary Node on Site B (Node 2)

EXAMPLE 27: TEST CRASH_PRIMARY_NODE_SITE_B

COMPONENT:

- Cluster node of site B

DESCRIPTION:

- Simulate a crash of the site B node running the primary HANA database.

TEST PROCEDURE:

- Crash the secondary node by sending a 'fast-reboot' system request.

```
suse02# echo 'b' > /proc/sysrq-trigger
```

RECOVERY PROCEDURE:

- If SBD fencing is used, pacemaker will not automatically restart after being fenced. In this case clear the fencing flag on **all** SBD devices and subsequently start pacemaker.

```
suse02# sbd -d /dev/disk/by-id/SBDA message suse02 clear
suse02# sbd -d /dev/disk/by-id/SBDB message suse02 clear
...
```

- Start the cluster framework.

```
suse02# systemctl start pacemaker
```

- Refresh the cluster resources on node 2 as root.

```
suse02# crm resource refresh rsc_SAPHana_HA1_HDB10 suse02
```

EXPECTED:

1. The cluster detects the failed primary node (node 2) and declares it UNCLEAN and sets the primary node (node 2) to status "partition with quorum".
2. The cluster fences the failed primary node (node 2).
3. The cluster declares the failed primary node (node 2) OFFLINE.
4. The cluster promotes the secondary HANA database (on node 1) to take over as primary.
5. The cluster migrates the IP address to the new primary (on node 1).
6. After some time the cluster shows the sync_state of the stopped secondary (on node 2) as SFAIL.
7. If SBD fencing is used, then the manual recovery procedure will be used to clear the fencing flag and restart pacemaker on the node.
8. When the fenced node (node 2) rejoins the cluster the former primary became a secondary.

9. Because AUTOMATED_REGISTER = "true" the cluster does restart the failed HANA database and register it against the new primary.
10. The cluster detects that the system replication is in sync again and marks it as ok (SOK).

11 Administration

11.1 Do’s and Don’ts

In your project, you should:

- Define STONITH before adding other resources to the cluster
- Do intensive testing.
- Tune the timeouts of operations of SAPHana and SAPHanaTopology.
- Start with PREFER_SITE_TAKEOVER = ”true”, AUTOMATED_REGISTER = ”false” and DUPLICATE_PRIMARY_TIMEOUT = ”7200”.

In your project, avoid:

- 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 this activities are completely forbidden.

11.2 Monitoring and Tools

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

11.2.1 HAWK – Cluster Status and more

You can use an Internet browser to check the cluster status.

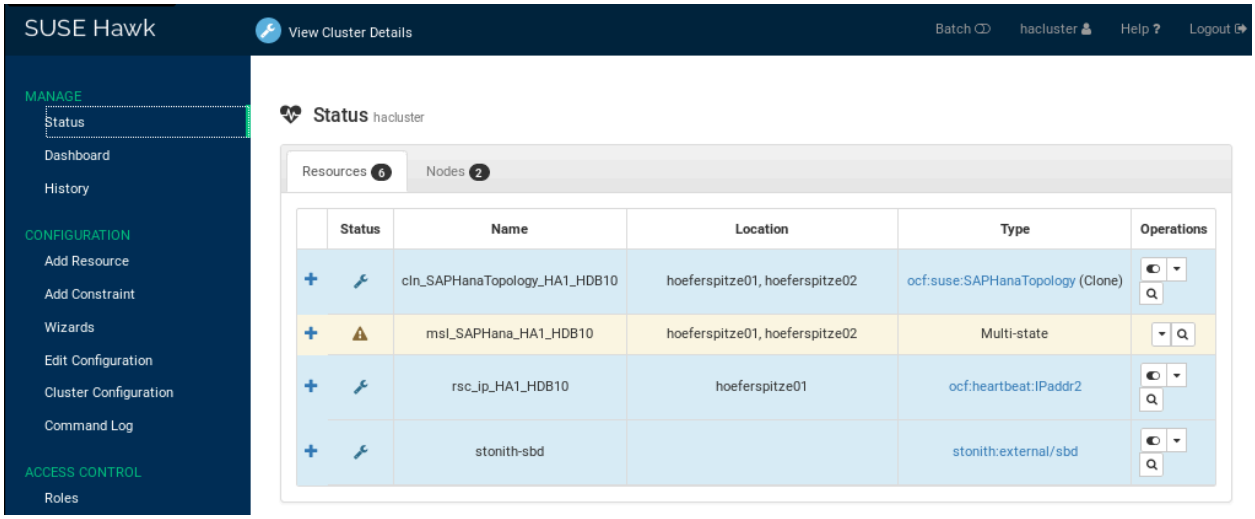


FIGURE 9: CLUSTER STATUS IN HAWK

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 configure resources and cluster bootstrap parameters. Read our product manuals for a complete documentation of this powerful user interface.

11.2.2 SAP HANA Studio

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

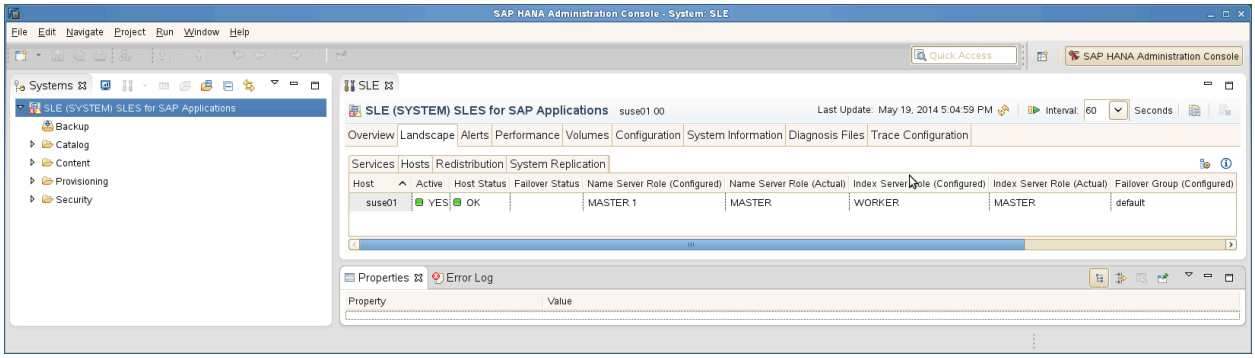


FIGURE 10: SAP HANA STUDIO – LANDSCAPE

11.2.3 Cluster Command Line Tools

A simple overview can be obtained by calling `crm_mon`. Using 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: suse01 (version 1.1.19+20180928.0d2680780-1.8-1.1.19+20180928.0d2680780) -
 partition with quorum
Last updated: Fri Nov 29 13:37:12 2019
Last change: Fri Nov 29 13:37:06 2019 by root via crm_attribute on suse02

2 nodes configured
6 resources configured

Online: [ suse01 suse02 ]

Full list of resources:

 stonith-sbd    (stonith:external/sbd): Started suse01
 Clone Set: cln_SAPHanaTopology_HA1_HDB10 [rsc_SAPHanaTopology_HA1_HDB10]
     Started: [ suse01 suse02 ]
 Master/Slave Set: msl_SAPHana_HA1_HDB10 [rsc_SAPHana_HA1_HDB10]
     Masters: [ suse01 ]
     Slaves: [ suse02 ]
 rsc_ip_HA1_HDB10       (ocf::heartbeat:IPaddr2):       Started suse01
```

See the manual page crm_mon(8) for details.

11.2.4 SAPHanaSR Command Line Tools

To show some SAPHana or SAPHanaTopology resource agent internal values, you can call the program `SAPHanaSR-showAttr`. The internal values, the 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.

Do not use cluster commands like `crm_attribute` to fetch the values directly from the cluster. If you use such commands, your methods will be broken when you need to move an attribute to a different storage place or even out of the cluster. At first *SAPHanaSR-showAttr* is a test program only and should not be used for automated system monitoring.

```
suse01:~ # SAPHanaSR-showAttr
Host \ Attr clone_state remoteHost roles       ... site    srmode sync_state ...
--------------------------------------------------------------------------------
```

```
 suse01      PROMOTED    suse02     4:P:master1:... WDF      sync  PRIM       ...
 suse02      DEMOTED     suse01     4:S:master1:... ROT      sync  SOK        ...
```

`SAPHanaSR-showAttr` also supports other output formats such as **script**. The script format is intended to allow running filters. The SAPHanaSR package beginning with version 0.153 also provides a filter engine `SAPHanaSR-filter`. In combination of `SAPHanaSR-showAttr` with output format script and `SAPHanaSR-filter` you can define effective queries:

```
suse01:~ # SAPHanaSR-showAttr --format=script | \
   SAPHanaSR-filter --search='remote'
Mon Nov 11 20:55:45 2019; Hosts/suse01/remoteHost=suse02
Mon Nov 11 20:55:45 2019; Hosts/suse02/remoteHost=suse01
```

`SAPHanaSR-replay-archive` can help to analyze the SAPHanaSR attribute values from `hb_report` (`crm_report`) archives. This allows post mortem analyzes.

In our example, the administrator killed the primary SAP HANA instance using the command `HDB kill-9`. This happened around 9:10 pm.

```
suse01:~ # hb_report -f 19:00
INFO: suse01# The report is saved in ./hb_report-1-11-11-2019.tar.bz2
INFO: suse01# Report timespan: 11/11/19 19:00:00 - 11/11/19 21:05:33
INFO: suse01# Thank you for taking time to create this report.
suse01:~ # SAPHanaSR-replay-archive --format=script \
    ./hb_report-1-11-11-2019.tar.bz2 | \
    SAPHanaSR-filter --search='roles' --filterDouble
Mon Nov 11 20:38:01 2019; Hosts/suse01/roles=4:P:master1:master:worker:master
Mon Nov 11 20:38:01 2019; Hosts/suse02/roles=4:S:master1:master:worker:master
Mon Nov 11 21:11:37 2019; Hosts/suse01/roles=1:P:master1::worker:
Mon Nov 11 21:12:43 2019; Hosts/suse02/roles=4:P:master1:master:worker:master
```

In the above example the attributes indicate that at the beginning suse01 was running primary (4:P) and suse02 was running secondary (4:S).

At 21:11 (CET) suddenly the primary on suse01 died - it was falling down to 1:P.

The cluster did jump-in and initiated a takeover. At 21:12 (CET) the former secondary was detected as new running master (changing from 4:S to 4:P).

11.2.5 SAP HANA LandscapeHostConfiguration

To check the status of an SAPHana database and to find out if the cluster should react, you can use the script **landscapeHostConfiguration** to be called as Linux user *<sid>adm*.

```
suse01:~> HDBSettings.sh landscapeHostConfiguration.py
| Host   | Host   | ... NameServer   | NameServer  | IndexServer | IndexServer |
```

```
|        | Active | ... Config Role  | Actual Role | Config Role | Actual Role |
| ------ | ------ | ... ------------ | ----------- | ----------- | ----------- |
| suse01 | yes    | ... master 1     | master      | worker      | master      |

overall host status: ok
```

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

TABLE 6: INTERPRETATION OF RETURN CODES

Return Code	Interpretation
4	SAP HANA database is up and OK. The cluster does interpret this as a 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 down by intention, this could trigger a takeover.
0	Internal Script Error – to be ignored.

11.3 Maintenance

To receive updates for the operating system or the SUSE Linux Enterprise High Availability Extension, it is recommended to register your systems to either a local SUSE Manager or SMT or remotely with SUSE Customer Center.

11.3.1 Updating the OS and Cluster

For an update of SUSE Linux Enterprise Server for SAP Applications packages including cluster software follow the rolling update procedure defined in the product documentation of SUSE Linux Enterprise High Availability Extension *Upgrading Your Cluster and Updating Software Packages* High Availability Administration Guide.

11.3.2 Updating SAP HANA - Seamless SAP HANA Maintenance

For updating SAP HANA database systems in system replication you need to follow the defined SAP processes. This section describes the steps to be done before and after the update procedure to get the system replication automated again.

SUSE has optimized the SAP HANA maintenance process in the cluster. The improved procedure only sets the master-slave-resource to maintenance and keep the rest of the cluster (SAPHanaTopology clones and IPaddr2 vIP resource) still active. Using the updated procedure allows a seamless SAP HANA maintenance in the cluster as the virtual IP address can automatically follow the running primary.

Prepare the cluster not to react on the maintenance work to be done on the SAP HANA database systems. Set the master/slave resource to be unmanaged and the cluster nodes in maintenance mode.

EXAMPLE 28: MAIN SAP HANA UPDATE PROCEDURE

Pre Update Task

For the master-slave-resource set the maintenance mode:

```
crm resource maintenance <master-slave-resource>
```

The <master-slave-resource> in the given guide is `msl_SAPHana_HA1_HDB10`.

Update

Process the SAP Update for both SAP HANA database systems. This procedure is described by SAP.

Post Update Task

Expect the primary/secondary roles to be exchanged after the maintenance. Therefore, tell the cluster to forget about these states and to reprobe the updated SAP HANA database systems.

```
crm resource refresh <master-slave-resource>
```

After the SAP HANA update is complete on both sites, tell the cluster about the end of the maintenance process. This allows the cluster to actively control and monitor the SAP again.

```
crm resource maintenance <master-slave-resource> off
```

11.3.3 Migrating an SAP HANA Primary

In the following procedures we assume the primary to be running on node1 and the secondary on node2. The goal is to "exchange" the roles of the nodes, so finally the primary should run on node2 and the secondary should run on node1.

There are different methods to get the exchange of the roles done. The following procedure shows how to tell the cluster to "accept" a role change via native HANA commands.

EXAMPLE 29: MIGRATING AN SAP HANA PRIMARY USING SAP TOOLSET

Pre move

Set the master-slave-resource to be maintenance. This could be done on any cluster node.

```
crm resource maintenance <master-slave-resource-name>
```

Manual Takeover Process

- Stop the primary SAP HANA database system. Enter the command in our example on node1 as user *<sid>adm*.

  ```
  HDB stop
  ```

- Start the takeover process on the secondary SAP HANA database system. Enter the command in our example on node2 as user *<sid>adm*.

  ```
  hdbnsutil -sr_takeover
  ```

- Register the former primary to become the new secondary. Enter the command in our example on node1 as user *<sid>adm*.

  ```
  hdbnsutil -sr_register --remoteHost=suse02 --remoteInstance=10 \
   --replicationMode=sync --name=WDF \
   --operationMode=logreplay
  ```

- Start the new secondary SAP HANA database system. Enter the command in our example on node1 as user *<sid>adm*.

  ```
  HDB start
  ```

Post Migrate

- Wait some time till SAPHanaSR-showAttr shows both SAP HANA database systems to be up again (field roles must start with the digit 4). The new secondary should have role "S" (for secondary).
- Tell the cluster to forget about the former master-slave roles and to re-monitor the failed master. The command could be submitted on any cluster node as user root.

```
crm resource refresh master-slave-resource-name
```

- Set the master/slave resource to the status managed again. The command could be submitted on any cluster node as user root.

```
crm resource maintenance <master-slave-resource-name> off
```

Now we explain how to use the cluster to partially automate the migration. For the described attribute query using *SAPHanaSR-showAttr* and *SAPHanaSR-filter* you need at least SAPHanaSR with package version 0.153.

EXAMPLE 30: MOVING AN SAP HANA PRIMARY USING THE CLUSTER TOOLSET

- Create a "move away" from this node rule by using the **force** option.

```
crm resource move <master-slave-resource-name> force
```

Because of the "move away" (**force**) rule the cluster will **stop** the current primary. After that, run a **promote** on the secondary site if the system replication was in sync before. You should not migrate the primary if the status of the system replication is not in sync (SFAIL).

Important

Migration without the **force** option will cause a takeover without the former primary to be stopped. Only the migration with **force** option is supported.

Note

The crm resource command **move** was previously named **migrate**. The **migrate** command is still valid but already known as obsolete.

- Wait till the secondary has completely taken over to be the new primary role. You see this using the command line tool *SAPHanaSR-showAttr* and check for the attributes "roles" for the new primary. It must start with "**4:P**".

```
suse01:~ # SAPHanaSR-showAttr --format=script | \
   SAPHanaSR-filter --search='roles'
Mon Nov 11 20:38:50 2019; Hosts/suse01/roles=1:P:master1::worker:
Mon Nov 11 20:38:50 2019; Hosts/suse02/roles=4:P:master1:master:worker:master
```

- If you have set up `AUTOMATED_REGISTER="true"`, you can skip this step. In other cases you now need to register the old primary. Enter the command in our example on node1 as user *<sid>adm*.

```
hdbnsutil -sr_register --remoteHost=suse02 --remoteInstance=10 \
    --replicationMode=sync --operationMode=logreplay \
    --name=WDF
```

- Clear the ban rules of the resource to allow the cluster to start the new secondary.

```
crm resource clear <master-slave-resource-name>
```

Note

The crm resource command **clear** was previously named **unmigrate**. The **unmigrate** command is still valid but already known as obsolete.

- Wait till the new secondary has started. You see this using the command line tool *SAPHanaSR-showAttr* and check for the attributes "roles" for the new primary. It must start with "**4:S**".

```
suse01:~ # SAPHanaSR-showAttr --format=script | \
   SAPHanaSR-filter --search='roles'
Mon Nov 11 20:38:50 2019; Hosts/suse01/roles=4:S:master1::worker:
Mon Nov 11 20:38:50 2019; Hosts/suse02/roles=4:P:master1:master:worker:master
```

12 Useful Links, Manuals, and SAP Notes

12.1 SUSE Best Practices and More

Blog series #towardsZeroDowntime

https://www.suse.com/c/tag/towardszerodowntime/

Best Practices for SAP on SUSE Linux Enterprise

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

Blog in 2014 - 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

12.2 SUSE Product Documentation

SUSE product manuals and documentation

https://documentation.suse.com/

Current online documentation of SLES for SAP

https://documentation.suse.com/sles-sap/12-SP4/

Current online documentation of SUSE Linux Enterprise High Availability Extension

https://documentation.suse.com/sle-ha/12-SP4/

Tuning guide for SUSE Linux Enterprise Server

https://documentation.suse.com/sles/12-SP4/html/SLES-all/book-sle-tuning.html

Storage admin guide for SUSE Linux Enterprise Server

https://documentation.suse.com/sles/12-SP4/single-html/SLES-storage/#stor-admin

Release notes

https://www.suse.com/releasenotes

TID Estimate correct multipath timeout

http://www.suse.com/support/kb/doc.php?id=7008216

TID How to load the correct watchdog kernel module

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

TID Addressing file system performance issues on NUMA machines

http://www.suse.com/support/kb/doc.php?id=7008919

TID Overcommit Memory in SLES

https://www.suse.com/support/kb/doc.php?id=7002775

TID Troubleshooting the SAPHanaSR python hook

https://www.suse.com/support/kb/doc.php?id=000019865

TID SAPHanaController running in timeout when starting SAP Hana

https://www.suse.com/support/kb/doc.php?id=000019899

TID SAP HANA monitors timed out after 5 seconds

https://www.suse.com/support/kb/doc.php?id=000020626

TID HA cluster takeover takes too long on HANA indexserver failure

https://www.suse.com/support/kb/doc.php?id=000020845

TID Cluster node fence as SAPHanaTopology fails ... during a normal cluster stop

https://www.suse.com/support/kb/doc.php?id=000020964

TID Basic health check for two-node SAP HANA performance based model

https://www.suse.com/support/kb/doc.php?id=7022984

TID HANA SystemReplication doesn't provide SiteName to Corosync Cluster

https://www.suse.com/support/kb/doc.php?id=000019754

TID SUSE Cluster Support for SAP HANA System Replication Active / Active ...

https://www.suse.com/support/kb/doc.php?id=7023884

TID SAPHanaSR-showAttr fails with error "Error: NIECONN_REFUSED ..."

https://www.suse.com/support/kb/doc.php?id=000020548

TID The vIP cluster resource does not follow the SAP HANA master ...

https://www.suse.com/support/kb/doc.php?id=000019769

TID HANA fail-over ... fail due to excessive os.system() execution times

https://www.suse.com/support/kb/doc.php?id=000020835

TID Address space monitoring and HANA DB performance

https://www.suse.com/support/kb/doc.php?id=000020746

TID HANA nodes end up having the same LPT values

https://www.suse.com/support/kb/doc.php?id=000020690

TID XFS metadata corruption and invalid checksum on SAP Hana servers

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

TID Handling failed NFS share in SUSE HA cluster for HANA system replication

https://www.suse.com/support/kb/doc.php?id=000019904

TID Indepth HANA Cluster Debug Data Collection (PACEMAKER, SAP)

https://www.suse.com/support/kb/doc.php?id=7022702

SLES 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/

12.3 Manual Pages

corosync.conf

corosync.conf.5

crm

crm.8

crm_mon

crm_mon.8

crm_simulate

crm_simulate.8

cs_clusterstate

cs_clusterstate.8

cs_show_sbd_devices

cs_show_sbd_devices.8

cs_wait_for_idle

cs_wait_for_idle.8

ha_related_sap_notes

ha_related_sap_notes.7

ha_related_suse_tids

ha_related_suse_tids.7

ocf_heartbeat_IPaddr2

ocf_heartbeat_IPaddr2.7

ocf_suse_SAPHana

ocf_suse_SAPHana.7

ocf_suse_SAPHanaTopology

ocf_suse_SAPHanaTopology.7

sbd

sbd.8

stonith_sbd

stonith_sbd.7

susChkSrv.py

susChkSrv.py.7

susTkOver.py

susTkOver.py.7

SAPHanaSR

SAPHanaSR.7

SAPHanaSR.py

SAPHanaSR.py.7

SAPHanaSR-manageProvider

SAPHanaSR-manageProvider.8

SAPHanaSR-filter

SAPHanaSR-filter.8

SAPHanaSR-showAttr

SAPHanaSR-showAttr.8

SAPHanaSR-replay-archive

SAPHanaSR-replay-archive.8

SAPHanaSR_basic_cluster

SAPHanaSR_basic_cluster.7

SAPHanaSR_manitenance_examples

SAPHanaSR_maintenance_examples.7

votequorum

votequorum.5

12.4 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

12.5 SAP Notes

1984787 - SUSE LINUX Enterprise Server 12: Installation notes

https://launchpad.support.sap.com/#/notes/1984787

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

https://launchpad.support.sap.com/#/notes/2205917

1876398 - Network configuration for System Replication in HANA SP6

https://launchpad.support.sap.com/#/notes/1876398

611361 - Hostnames of SAP servers

https://launchpad.support.sap.com/#/notes/611361

1275776 - Preparing SLES for Sap Environments

https://launchpad.support.sap.com/#/notes/1275776

1514967 - SAP HANA: Central Note

https://launchpad.support.sap.com/#/notes/1514967

1523337 - SAP In-Memory Database 1.0: Central Note

https://launchpad.support.sap.com/#/notes/1523337

2380229 - SAP HANA Platform 2.0 - Central Note

https://launchpad.support.sap.com/#/notes/2380229

1501701 - Single Computing Unit Performance and Sizing

https://launchpad.support.sap.com/#/notes/1501701

1944799 - SAP HANA Guidelines for SLES Operating System Installation

https://launchpad.support.sap.com/#/notes/1944799

1890444 - Slow HANA system due to CPU power save mode

https://launchpad.support.sap.com/#/notes/1890444

1888072 - SAP HANA DB: Indexserver crash in strcmp sse42

https://launchpad.support.sap.com/#/notes/1888072

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

https://launchpad.support.sap.com/#/notes/1846872

13 Examples

13.1 Example `ha-cluster-init` Configuration

```
suse01:~ # ha-cluster-init -u
  Generating SSH key
  Configuring csync2
  Generating csync2 shared key (this may take a while)...done
  csync2 checking files...done

Configure Corosync (unicast):
  This will configure the cluster messaging layer.  You will need
  to specify a network address over which to communicate (default
  is eth0's network, but you can use the network address of any
  active interface).

  Address for ring0 [192.168.1.11]
  Port for ring0 [5405]

Configure SBD:
  If you have shared storage, for example a SAN or iSCSI target,
  you can use it avoid split-brain scenarios by configuring SBD.
  This requires a 1 MB partition, accessible to all nodes in the
  cluster.  The device path must be persistent and consistent
  across all nodes in the cluster, so /dev/disk/by-id/* devices
  are a good choice.  Note that all data on the partition you
  specify here will be destroyed.

Do you wish to use SBD (y/n)? y
  Path to storage device (e.g. /dev/disk/by-id/...), or "none" []/dev/disk/by-id/SBDA
```

```
All data on /dev/disk/by-id/SBDA will be destroyed!
Are you sure you wish to use this device (y/n)? y
  Initializing SBD......done
  Hawk cluster interface is now running. To see cluster status, open:
    https://192.168.1.11:7630/
  Log in with username 'hacluster', password 'linux'
You should change the hacluster password to something more secure!
  Waiting for cluster........done
  Loading initial cluster configuration

Configure Administration IP Address:
  Optionally configure an administration virtual IP
  address. The purpose of this IP address is to
  provide a single IP that can be used to interact
  with the cluster, rather than using the IP address
  of any specific cluster node.

Do you wish to configure a virtual IP address (y/n)? n
  Done (log saved to /var/log/ha-cluster-bootstrap.log)
```

13.2 Example Cluster Configuration

The following complete crm configuration is for a two-node cluster (suse01, suse02) and an SAP HANA database with SID HA1 and instance number 10. The virtual IP address in the example is 192.168.1.20

```
node suse01
node suse02

primitive rsc_SAPHanaTopology_HA1_HDB10 ocf:suse:SAPHanaTopology \
        op monitor interval=10 timeout=300 \
        op start interval=0 timeout=300 \
        op stop interval=0 timeout=300 \
        params SID=HA1 InstanceNumber=10
primitive rsc_SAPHana_HA1_HDB10 ocf:suse:SAPHana \
        op monitor interval=61 role=Slave timeout=700 \
        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 \
        params SID=HA1 InstanceNumber=10 PREFER_SITE_TAKEOVER=true \
               DUPLICATE_PRIMARY_TIMEOUT=7200 AUTOMATED_REGISTER=false
primitive rsc_ip_HA1_HDB10 ocf:heartbeat:IPaddr2 \
        op monitor interval=10 timeout=20 \
        params ip="192.168.1.20"
```

```
primitive stonith-sbd stonith:external/sbd \
        params pcmk_delay_max=15
ms msl_SAPHana_HA1_HDB10 rsc_SAPHana_HA1_HDB10 \
        meta clone-max=2 clone-node-max=1 interleave=true
clone cln_SAPHanaTopology_HA1_HDB10 rsc_SAPHanaTopology_HA1_HDB10 \
        meta clone-node-max=1 interleave=true
colocation col_saphana_ip_HA1_HDB10 2000: \
        rsc_ip_HA1_HDB10:Started msl_SAPHana_HA1_HDB10:Master
order ord_SAPHana_HA1_HDB10 Optional: \
        cln_SAPHanaTopology_HA1_HDB10 msl_SAPHana_HA1_HDB10
property cib-bootstrap-options: \
        dc-version="1.1.19+20180928.0d2680780-1.8-1.1.19+20180928.0d2680780" \
        cluster-infrastructure=corosync \
        stonith-enabled=true \
        stonith-action=reboot \
        stonith-timeout=150s \
        last-lrm-refresh=1398346620
rsc_defaults rsc-options: \
        resource-stickiness=1000 \
        migration-threshold=5000
op_defaults op-options \
        timeout=600 \
        record-pending=true
```

13.3 Example for /etc/corosync/corosync.conf

The following file shows a typical corosync configuration with one ring. Review the SUSE product documentation about details and about additional rings.

```
# Read the corosync.conf.5 manual page
totem {
        version: 2
        secauth: on
        crypto_hash: sha1
        crypto_cipher: aes256
        cluster_name: suse-ha
        clear_node_high_bit: yes
        token: 5000
        token_retransmits_before_loss_const: 10
        join: 60
        consensus: 6000
        max_messages: 20
        interface {
                ringnumber: 0
                mcastport: 5405
```

```
                ttl: 1
        }

        transport: udpu
}

logging {
        fileline: off
        to_stderr: no
        to_logfile: no
        logfile: /var/log/cluster/corosync.log
        to_syslog: yes
        debug: off
        timestamp: on
        logger_subsys {
                subsys: QUORUM
                debug: off
        }

}

nodelist {
        node {
                ring0_addr: 192.168.1.11
                nodeid: 1
        }

        node {
                ring0_addr: 192.168.1.12
                nodeid: 2
        }

}

quorum {
        provider: corosync_votequorum
        expected_votes: 2
        two_node: 1
}
```

13.4 Example for the IPMI STONITH Method

```
primitive rsc_suse01_stonith stonith:external/ipmi \
    params hostname="suse01" ipaddr="192.168.1.101" userid="stonith" \
    passwd="k1llm3" interface="lanplus" \
```

```
    op monitor interval="1800" timeout="30"
    ...
primitive rsc_suse02_stonith stonith:external/ipmi \
    params hostname="suse02" ipaddr="192.168.1.102" userid="stonith" \
    passwd="k1llm3" interface="lanplus" \
    op monitor interval="1800" timeout="30"
    ...
location loc_suse01_stonith rsc_suse01_stonith -inf: suse01
location loc_suse02_stonith rsc_suse02_stonith -inf: suse02
```

14 Reference

For more detailed information, have a look at the documents listed below.

14.1 Pacemaker

Pacemaker Project Documentation

https://clusterlabs.org/pacemaker/doc/

15 Legal notice

16 GNU Free Documentation License

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.

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
```

```
   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.