Best Practice for SAP

Simple Stack - SAP NetWeaver High Availability on SUSE Linux Enterprise 12

Best Practice for SAP

Simple Stack - SAP NetWeaver High Availability on SUSE Linux Enterprise 12

by Fabian Herschel and Lars Pinne

Publication Date: October 20, 2016

SUSE LLC
10 Canal Park Drive
Suite 200
Cambridge MA 02141
USA
https://www.suse.com/documentation

Contents

1 Executive Summary

SAP Business Suite is a sophisticated application platform for large enterprises and mid-size companies. Many critical business environments require the highest possible SAP application availability. SUSE Linux Enterprise Server for SAP Applications 12, when running on modern x86-64 hardware platforms, satisfies this requirement. Together with a redundant layout of the technical infrastructure, single points of failure can be eliminated.

SAP NetWeaver is a common stack of middleware functionality used to support the SAP business applications. This guide describes a SAP NetWeaver installation on SUSE Linux Enterprise Server for SAP Applications 12 with the included SUSE Linux Enterprise High Availability Extension. We will also describe possible failure scenarios and methods for avoiding them. The described concept has shown its maturity during several years of productive operations for customers of different size and branches.

The described storage stack and SAP configuration can be used with or without a high availability cluster. It is possible to add high availability functionality to an already running system, if the installation complies with the described solution.

This guide will show you how to:

- Plan a SUSE Linux Enterprise platform for SAP workload
- Set up a Linux high availability infrastructure for SAP
- Perform a basic SAP NetWeaver installation on SUSE Linux Enterprise Server for SAP Applications 12

This guide will also help you to install the following software components:

- SUSE Linux Enterprise Server for SAP Applications 12, including the high availability extension
- MaxDB (SAP HANA, SAP ASE, Oracle, and DB2 are supported as well)
- SAP NetWeaver 7.5 (other versions are supported, too)

For SAP HANA scenarios SUSE is publishing seperate best practices at: https://www.suse.com/products/sles-for-sap/resource-library/.

This guide is aimed at IT professionals with skills in:

- SAP basic operating
- Data center system concepts and configuration
- Linux knowledge at LPIC-1 or SCA (SUSE Certified Administrator) level

To follow this guide you need access to the following resources:

- SUSE Linux Enterprise Server for SAP Applications 12 installation media. To update the systems you must have either Internet access, or a local software management solution, such as SUSE Manager or SMT (Subscription Management Tool).
- SAP NetWeaver 7.5 Installation Media
- Appropriate hardware (two servers, network, storage, see below)

While this guide focuses on well known components, additional features of SUSE Linux Enterprise Server for SAP Applications 12 allow for enhanced concepts, like the use of local area network based Distributed Replicated Block Devices (DRBD) instead of storage area networks. Those concepts are described in additional documents.

2 Overview

2.1 What is this guide about?

SUSE Linux Enterprise Server for SAP Applications is the optimal platform to run SAP applications with high availability. Together with a redundant layout of the technical infrastructure, single points of failure can be eliminated.

The high availability cluster described in this document can be installed on physical hardware as well as in virtual machines. All major virtualization platforms are supported. Information about virtualization platform support for SAP can be found in SAP Note 1122387 - Linux: SAP Support in virtualized environments. Information about virtualization platform support for SAP NetWeaver and SUSE Linux Enterprise Server for SAP Applications could be found on http://scn.sap.com/docs/DOC-27321 and other web pages listed in the appendix.

This guide will show you how to:

- Plan a SUSE Linux Enterprise high availability platform for SAP NetWeaver.
- Set up a Linux high availability platform and perform a basic SAP NetWeaver installation. Central services will bw installed as a separate SAP instance. A database will be installed as well.

This guide will also help you to install the following software components:

- SUSE Linux Enterprise Server for SAP Applications 12 including the pacemaker cluster
- SAP NetWeaver 7.5 (other versions are supported, too).

This guide is aimed at IT professionals with skills in:

- SAP basic operating,
- Data center system concepts and configuration,
- Linux knowledge at LPIC-1 or SCA (SUSE Certified Administrator) level.

To follow this guide you need access to the following resources:

- SUSE Linux Enterprise Server for SAP Applications 12 installation media.
- To update the systems you must have either Internet access, SUSE Manager, or a local Subscription Management Tool (SMT).
- SAP NetWeaver 7.5 Installation Media.
- Appropriate hardware (two servers, network, storage, see below).

This guide is organized into two main parts:

- Part I - Introduction and Concept
  In a perfect world everyone who plans, installs, or runs SAP NetWeaver including central services and database with high availability on SUSE Linux Enterprise Server for SAP Applications has read this chapter.
- Part II - Planning, Installing, and Testing the Cluster
  Details to help perform the mentioned tasks or to learn more details about SAP NetWeaver with high availability on SLES for SAP.

2.2 What is not covered?

While this guide uses well known components, additional features of SUSE Linux Enterprise Server for SAP Applications 12 allow for enhanced concepts, like the use of local area network-based Distributed Replicated Block Devices (DRBD) instead of storage area networks. These concepts are described in additional documents.

SLES for SAP contains SUSE Linux Enterprise Server and SUSE Linux Enterprise High Availability Extension. For platforms where SLES for SAP is not available, it is possible to set up an high availability solution using SUSE Linux Enterprise Server and SUSE Linux Enterprise High Availability Extension instead of SLES for SAP. Such a setup will lack features of SLES for SAP, like page cache limit, and Expanded Service Pack Overlap Support. Nevertheless, SUSE Linux Enterprise Server and SUSE Linux Enterprise High Availability Extension would allow to run SAP NetWeaver with high availability on other hardware platforms like System z. Such platforms may need different configuration parameters that are not explained in this document.

I Introduction and Concept

3 Introduction

3.1 SAP on Linux

SUSE has multiple dedicated resources working at SAP headquarters and the SAP LinuxLab to ensure maximum interoperability between our products with SAP software and technologies.

SAP has built SAP LinuxLab to assist with the release of SAP software on Linux. LinuxLab supports other SAP departments in development of the Linux platform, processes Linux-specific support problems and acts as an information hub to all the SAP partners in the Linux ecosystem.

Where possible, SAP endorsed open standards and technologies. This allowed SAP to support a very wide range of operating systems and hardware platforms. Open-source-based Linux provides the maximum in “openness,” so it was only natural for SAP to start supporting it in 1999. SAP tries to be “operating system agnostic” and act neutral on the customer's chosen operating systems. Unlike other software vendors, SAP has clearly stated its policies toward open source and Linux. For instance, the usage of binary only (closed source) device drivers (kernel modules) is not supported. This helps the Linux and open source communities, since hardware vendors are encouraged to either publish the specifications and APIs of their hardware so the Linux community can write drivers, or make driver source code that can be included in the Linux kernel available (see SAP note 784391).

3.2 SUSE Linux Enterprise Server

3.2.1 SUSE Linux Enterprise Server Is Prepared for SAP

Since the very start of SAP involvement with Linux, SUSE has been the number one Linux platform for SAP. SUSE Linux Enterprise Server has become the reference platform for SAP's software development. SAP now actively suggests SUSE Linux Enterprise Server when customers want to run SAP workloads on Linux. Recently, SUSE Linux Enterprise Server became part of SAP products by being chosen as the operating system for SAP's all-in-one business solutions.

SUSE Linux Enterprise Server has been fine tuned to provide the best Linux platform for SAP software.

YaST, the SUSE Linux Enterprise Server main installation and administration tool, provides a SAP software pattern that installs the prerequisites needed for running SAP software. Also, some system parameters are fine-tuned in preparation for the SAP workload.

In this document, SUSE Linux Enterprise Server for SAP Applications 12 SP1 x86-64 with updates from SUSE Customer Center (SCC) is used.

3.2.2 SUSE Linux Enterprise Server for SAP Applications

SUSE Linux Enterprise Server for SAP Applications is the only operating system optimized for all mission-critical SAP software solutions, including appliances, and is recommended by SAP as a preferred Linux platform.

SUSE Linux Enterprise Server for SAP Applications is based on the newest SUSE Linux Enterprise technology, a highly reliable, scalable, secure and optimized server operating system that is built to power both physical and virtual mission-critical workloads. SUSE Linux Enterprise Server for SAP Applications is optimized for all mission-critical SAP software solutions. It is validated and certified by SAP and is, therefore, uniquely positioned to support SAP software solutions including appliances. Find more details about that product at http://www.suse.com/products/sles-for-sap/.

Installing a high available cluster using SUSE Linux Enterprise Server for SAP Applications is more comfortable, because all needed packages including the cluster packages and SAP related packages like the java JDK are already included in one single product.

3.2.3 Support and Certification

For SAP customers, SUSE offers the SUSE Linux Enterprise Server Priority Support for SAP Applications. This support offering provides customers with a single support point of entry, from the operating system through the application (http://www.novell.com/docrep/2007/05/4611143_f_en.pdf).

Currently the following SUSE Linux Enterprise Server systems including all service packs (SP) are certified as 64-bit version for SAP:

- SUSE Linux Enterprise Server for SAP Applications 12 (http://www.suse.com/products/sles-for-sap/)
- SUSE Linux Enterprise Server 12 (http://www.suse.com/products/server/)
- SUSE Linux Enterprise Server for SAP Applications 11 (http://www.suse.com/products/sles-for-sap/)
    - SUSE Linux Enterprise Server for SAP Applications 11 for x86_64 (AMD64 and Intel EM64T)
    - SUSE Linux Enterprise Server for SAP Applications 11 for IBM Power
- SUSE Linux Enterprise Server 11 (http://www.suse.com/products/server/)
    - SUSE Linux Enterprise Server 11 for AMD64 and Intel EM64T
    - SUSE Linux Enterprise Server 11 for IBM Power
    - SUSE Linux Enterprise Server 11 for IBM System z
    - SUSE Linux Enterprise Server 11 for Intel ia64
- SUSE Linux Enterprise Server 10
    - SUSE Linux Enterprise Server 10 for AMD64 and Intel EM64T
    - SUSE Linux Enterprise Server 10 for IBM Power
    - SUSE Linux Enterprise Server 10 for IBM System z
    - SUSE Linux Enterprise Server 10 for Intel ia64

SAP and SUSE are working together to ensure that SUSE Linux Enterprise Server service packs always match the certification of the respective product. In fact SAP recommends always using the latest available service pack.

SUSE will provide at least ten years of general support for platform and operating system products, including its revisions, starting at the date of a product's general availability. When general support ends, SUSE will offer extended support for a minimum of three years. This gives SAP customers a long installation run-time, ensuring a low TCO.

3.3 SUSE Linux Enterprise High Availability Extension

Your data is the most valuable asset that you have — it is what your business depends on. Robust, scalable and manageable storage is a top priority for your IT department. High availability storage, applications, and services are critical for your business to be competitive. The SLE High Availability Extension comes with SLES for SAP and satisfies these needs. It includes high availability service and application clustering, file systems/clustered file systems, network file systems (NAS), volume managers, networked storage systems and drivers (SAN), and the management of all these components working together.

Unlike proprietary solutions, SLE High Availability Extension keeps costs low by integrating open source, enterprise-class components. The key components of the extension are:

- Pacemaker, a high availability cluster manager that supports multinode failover.
- Resource Agents to monitor and manage resources, compliant to the Open Cluster Framework (OCF).
- Oracle Cluster File System 2 (OCFS2), a parallel cluster file system that offers scalability.
- Cluster Logical Volume Manager (cLVM2), a logical volume manager for the Linux kernel, which provides a method of allocating space on mass storage devices that is more flexible than conventional partitioning schemes.
- Distributed Replicated Block Devices (DRBD) provides fast data resynchronization capabilities over LAN, replicated storage area network (SAN) semantics, allowing cluster-aware file systems to be used without additional SANs.
- High Availability Web Konsole (HAWK) and various command line tools.

Availability is a result of the interaction of cluster software with application services on the front side and the operating system and hardware resources on the other side. Following this basic idea, cluster software like Pacemaker could not increase the availability on its own. It needs a lot of modules, such as services, resource agents, a messaging layer, network and file system availability, and a stable Linux kernel designed and configured for productive server systems in data centers.

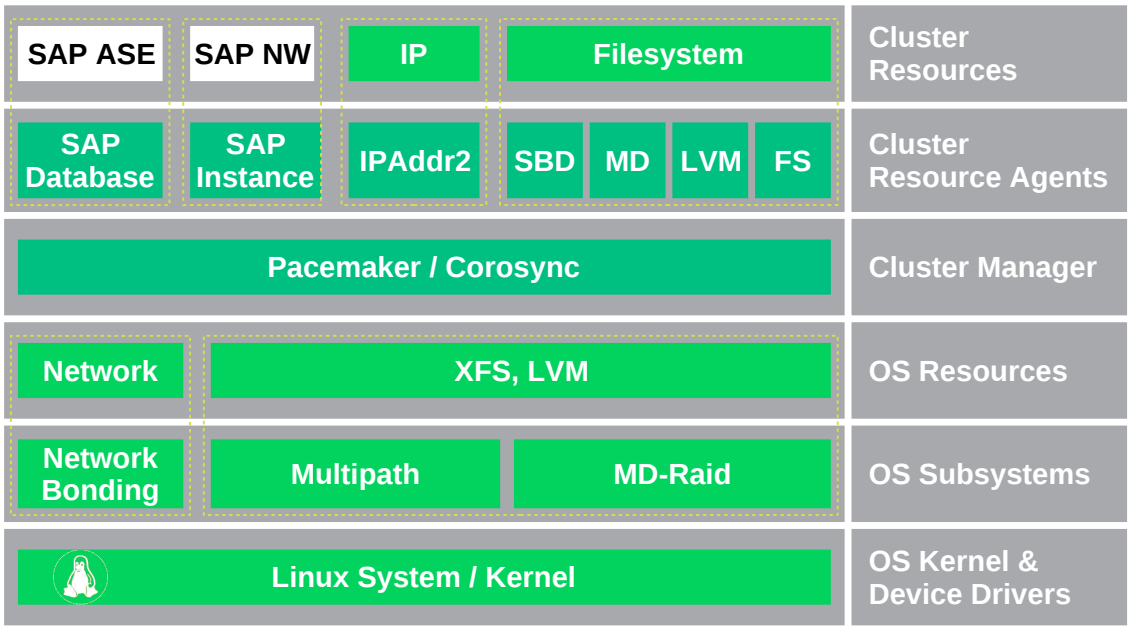

FIGURE 3.1: MODULES OF A HIGH AVAILABILITY SAP CLUSTER

The central application of our cluster is the SAP system itself. We need to provide the SAP database and the central SAP instance with high availability (white boxes). Operating system (light colored boxes) and cluster software (dark colored boxes) together give us the needed functionality.

In this document, SUSE Linux Enterprise Server for SAP Applications 12 with updates from SUSE Customer Center (SCC)

3.3.1 Storage Scenarios and I/O Stack for SAP

The file systems required by the operating system (i.e., mounted on `/` or `/var`) are usually stored on local hard drives using RAID-1 or RAID-5 disk arrays. Sometimes the LVM is also used to gain more flexibility; i.e., for online file system extensions. This implies that these file systems can only be accessed by the local OS or applications running locally on this server. Even if these file systems could be placed on LUNs in a storage area network (SAN) outside the server, they are treated as local.

Besides the usual OS file systems, SAP and the SAP databases require their own file systems. These file systems are not stored locally. Instead they are provided by NFS file servers or on LUNs exported by storage units in storage area networks. Please be aware that the storage which the database resides on is subject to specifications and recommendations of the database vendor.

Some file systems have to be shared across all cluster nodes such that every cluster node can access the file system at the same time. The recommended way of providing this service is the usage of a network file system (NFS). The NFS service also has to be highly available. In this document we use a dedicated NFS. This server could be either a third-party product or a SUSE Linux Enterprise High Availability Extension cluster that runs NFS resources. Under specific circumstances it also might be possible to integrate the NFS server into the SAP cluster. Both, the separate NFS cluster as well as the integrated NFS server will be described in separate documents (https://www.suse.com/products/sles-for-sap/resource-library/sap-best-practices.html).

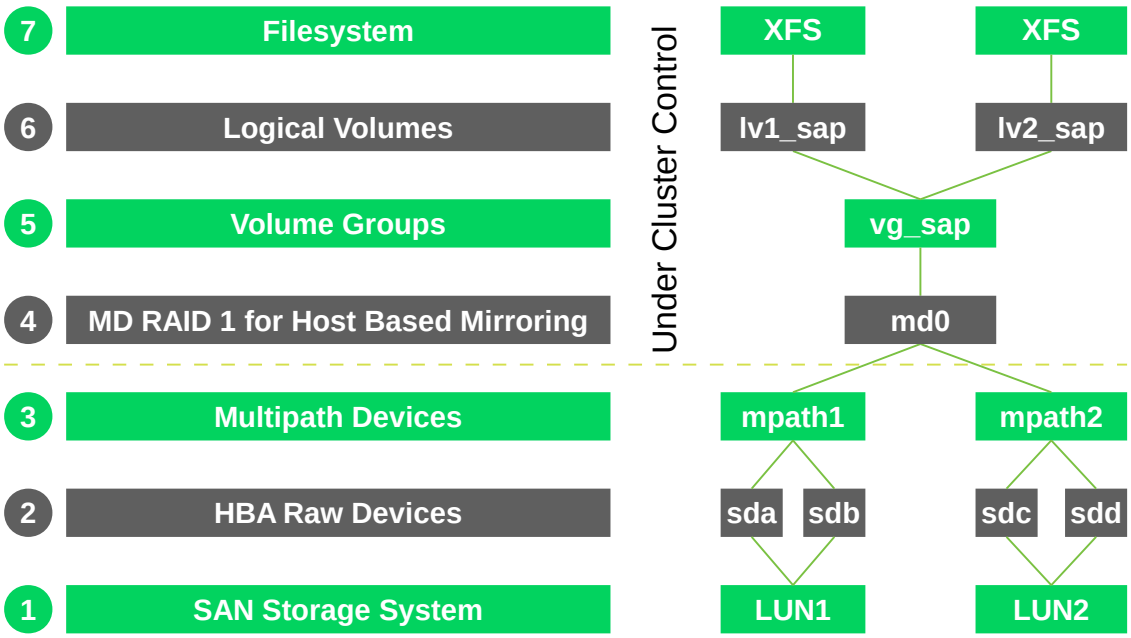

FIGURE 3.2: SAN IO LAYERS OF HIGH AVAILABILITY STACK FOR SAP

Other file systems, mounted to i.e. `/usr/sap/SID` or `/oracle/SID`, only have to be available on one cluster node at the same time. However, each cluster node must be able to access these file systems if the cluster manager decides to use them.

In our current concept we use LVM2 on top of MD RAID, that has shown its road capability for years. This storage stack can be used with or without cluster.

The UNIX file system is the highest layer of a whole I/O stack consisting of multiple I/O layers. Each layer provides a certain kind of functionality. For all I/O critical tasks, we have configured an I/O stack that supports the following functions:

- Low latency: high I/O throughput and fast response times
- Host-based mirroring for storing data simultaneously on two separate storage units in a SAN
- Logical Volume Manager for a flexible management of file systems
- Multipath I/O for an additional level of redundancy for file systems, stored on LUNs in the SAN
- Online resizing (extending) of file systems, snapshots of file systems using LVM snapshots, moving or copying file system

While this guide focuses on well-known components, additional features of SUSE Linux Enterprise Server for SAP Applications 12 allow for enhanced concepts, like the use of local area network-based Distributed Replicated Block Devices (DRBD) instead of storage area networks. Those concepts are described in additional documents (https://www.suse.com/documentation/sle-ha-12/singlehtml/book_sleha_techguides/book_sleha_techguides.html).

3.3.2 SAP Resource Agents

The SAPInstance resource agent is responsible for starting, stopping, and monitoring the services in a SAP instance. The following services are monitoring in the resource agent: disp+work, msg_server, enserver, enrepserver, jcontrol, and jstart.

The SAPInstance resource agent can be used to manage the following SAP instances:

- SAP NetWeaver AS ABAP Release 6.20 – 7.5
- SAP NetWeaver AS Java Release 6.40 - 7.5 (for 6.40 please also read SAP Note 995116)
- SAP NetWeaver AS ABAP + Java Add-In Release 6.20 - 7.5 (Java is not monitored by the cluster)

The purpose of the SAPDatabase resource agent is to start, stop and monitor the database instance of an SAP system. Together with the relational database management systems (RDBMS) it will also control the related network service for the database. Like the Oracle Listener and

the xserver of MaxDB. The resource agent expects a standard SAP installation and therefore needs fewer parameters to configure. The monitor operation of the resource agent can test the availability of the database by using SAP tools (R3trans or jdbcconnect). This ensures that the database is really accessible for the SAP system.

The SAPDatabase resource agent supports the following databases in a SAP installation:

- SAP HANA DATABASE 1.x (Scale-Up)
- MaxDB 7.x
- SAP ASE 16.x
- IBM DB2 UDB for Windows and UNIX 9.x
- Oracle 11 and 12

The resource agents are part of SUSE Linux Enterprise Server for SAP Applications. Currently SAP HANA is also supported in system replication (Scale-Up) using the resource agent SAPHana (package SAPHanaSR). More information can be found in another document (https://www.suse.com/products/sles-for-sap/resource-library/sap-best-practices.html). SAP HANA will be supported in system replication (Scale-Out) using the resource agent SAPHanaController (package SAPHanaSR-ScaleOut).

3.3.3 The SAP SUSE cluster connector

The SAP SUSE cluster connector was designed to control single independent SAP NetWeaver instances. In the simple stack the SAP system is used and managed with dependencies between the SAP instances, so the SAP SUSE cluster connector should not used in this concept.

If you want to use the SAP SUSE cluster connector, it is highly recommended to implement the SAP Enqueue Replication Server scenario.

3.3.4 Virtualization

The SLE High Availability Extension used in this document is supported on all major virtualization platforms, particularly current releases of VMware ESX. Detailed information about virtualization platform support for SLE High Availability Extension could be found on web pages listed in the appendix.

3.4 Database for SAP

Three major RDBMS are supported with SAP on SUSE Linux Enterprise Server, either with or without the SUSE Linux Enterprise Server High Availability Extension. In addition, for SAP deployments with each of these databases, SAP priority support is available from SUSE.

- SAP HANA DATABASE 1.x
- MaxDB 7.x
- SAP ASE 16.x
- IBM DB2 UDB for Windows and UNIX 9.x
- Oracle 11 and 12

Supported hardware platforms for all RDBMS on SUSE Linux Enterprise Server 12 is x86-64 (also known as “x86 64-bit”, “AMD64”, “EM64T”). Nevertheless, some databases are also supported on other hardware platforms such as ppc64 or System z. Details can be found on the certification pages listed as links in the appendix.

SUSE Linux Enterprise Server ships with an installation pattern for Oracle database servers. This contains the RPM orarun which does several system settings according to Oracle server needs. If Oracle is installed together with SAP, this package should not be installed, to avoid conflicts with the settings from the RPM sapconf. Further information can be found on the web pages listed in the appendix.

MaxDB was chosen as RDBMS for this document. It was mainly chosen to shorten the planning and installation procedure. Since the document is targeting a basic installation of SAP in a high-availability cluster, the use cases are covered with MaxDB. This installation example does not imply any decision about how one of the database management systems fits for a given workload.

3.4.1 MaxDB

The SAP MaxDB is the database of choice for small business and midsize companies requiring a solid, affordable low-maintenance database. MaxDB is available for all installations of the SAP Business All-in-One solution and the SAP Business Suite family of business applications. MaxDB is bundled with the full license for the SAP NetWeaver technology platform, and the SAP Business by Design solution uses MaxDB as the default database for the host system. Designed for

online transaction processing and database sizes up to multiple terabytes, MaxDB is the preferred database for internal SAP installations on UNIX and Linux (http://www.sap.com/solutions/sme/businessallinone/kits/lowertco.epx).

MaxDB´s ancestor AdabasD was available on Linux in 1996. Between 1997 and 2004 the software was available as SAP-DB, and then it was named MaxDB. MaxDB as standalone product is supported for SUSE Linux Enterprise Server 11 on the hardware platforms x86-64, and ppc64 (http://maxdb.sap.com/documentation/). As an integrated database for the SAP NetWeaver technology platform, the respective certification applies for SUSE Linux Enterprise Server 11 on x86-64, and ppc64. For SAP systems, the appropriate product certification matrix should be applied (https://websmp201.sap-ag.de/pam).

MaxDB installation media can be obtained from the SAP portal along with NetWeaver. The installation of MaxDB is seamlessly integrated into the SAP installer. SAP offers several services around MaxDB for SAP applications. More information can be found on web pages listed in the appendix.

3.5 Sizing, Deployment and Migration Considerations

Running SAP workloads in SUSE Linux Enterprise Server can reduce the TCO considerably. Through the homogenization of the data center (and thus the consolidation of resources), cost savings can be achieved even for highly specialized SAP installations. SUSE Linux Enterprise Server is very well prepared for automated installation and deployment, making it easy to deploy many servers. SUSE provides tools for managing updates and releases to ease day-to-day administration.

SUSE Linux Enterprise Server is very capable of providing the base for small or large systems. Customers run smaller central instances or larger distributed systems all with the same system base. It is quite possible to run multiple SAP instances in parallel on one system even when using high availability clusters.

SAP requests that system sizing is done by the hardware vendor. SUSE has good relationships with many hardware vendors to make sure SUSE Linux Enterprise Server runs smoothly on a broad range of enterprise servers fit to run SAP workloads. SUSE and its partners are very active in providing customers with solutions to their specific needs when it comes to Linux deployment. SUSE consulting has been developing best practices for high availability SAP installations and provides this information to customers and partners.

Hundreds of successful SAP to Linux migrations have been made. The results regarding cost savings, performance and reliability have exceeded expectations in many instances.

Since most data centers have adopted a Linux strategy, the know-how for deploying and administrate Linux systems is often in place and available. SAP-specific configurations and administration experience is available through SUSE consultation and partners. This makes the operating system side of the migration less risky and an ROI can be seen within the first six months of migration.

SAP provides check lists and guidelines for the OS and database migration.

4 One Concept — Different Use Cases

SAP NetWeaver installation can be adapted to several architectures for the entire system. You can plan and implement a “simple stack” SAP system with instances and the database on the same machine without any methods to increase the availability. You can take this simple installation scenario, add high availability components and bring the SAP system into a high availability scenario. There might also be reasons to install a SAP system with additional application servers (like for horizontal scaling).

In a high availability scenario this could lead into a more advanced “enqueue replication” setup. In all previously mentioned cases, SAP is treated in a two-tier manner, that means SAP application and database processes could run on the same system.

Another common use case is to run enqueue replication server and database on different machines. In this case, high availability for the enqueue replication server is covered by the described concept as well. A best practice guide covering the "enqueue-replication" scenario is available at our SUSE website https://www.suse.com/products/sles-for-sap/resource-library/

This document covers the use case "simple stack".

4.1 High Availability Concept Overview

The concept was chosen to match the top-level goals:

- High availability
- Low complexity
- Flexible scalability
- Road capability

To fit these goals, we separate the SAP system into a clustered and an unclustered area. The clustered area holds all mandatory SAP components such as SAP database and needed SAP instances.

The unclustered area holds the optional and scalable SAP components such as additional SAP instances. This allows to scale the entire SAP system without increasing the cluster complexity. The horizontal scaling is just a purpose of the unclustered area.

The architecture is focused to one single SAP system, even if is possible to run more than one SAP system in the same cluster.

The concept uses SBD (STONITH Block Device) for node fencing, to protect the storage (node fencing via remote managemnt boards is decribed in another document).

The network file system (NFS) is used to share data between the nodes, for example for the SAP transport directory. In this concept we assume that a reliable NFS is provided by a service outside the cluster. Either a highly available NFS server based on SUSE Linux Enterprise Server for SAP Applications 12 or a third party product could be used. An NFS high availability cluster based on SUSE Linux Enterprise Server for SAP Applications 12 is described in another document. In some situations it might be desirable to have the NFS server in the same cluster as the SAP application. Such an embedded NFS scenario will be covered in a separate document.

In a complex, high availability SAP environment, several types of failures may occur. These failures range from software crashes up to a loss of the whole network or SAN infrastructure. The cluster must be able to safely handle all of these failures. Even in a split brain scenario, if the cluster communication between both nodes is broken, the cluster must ensure a proper continuation of all services.

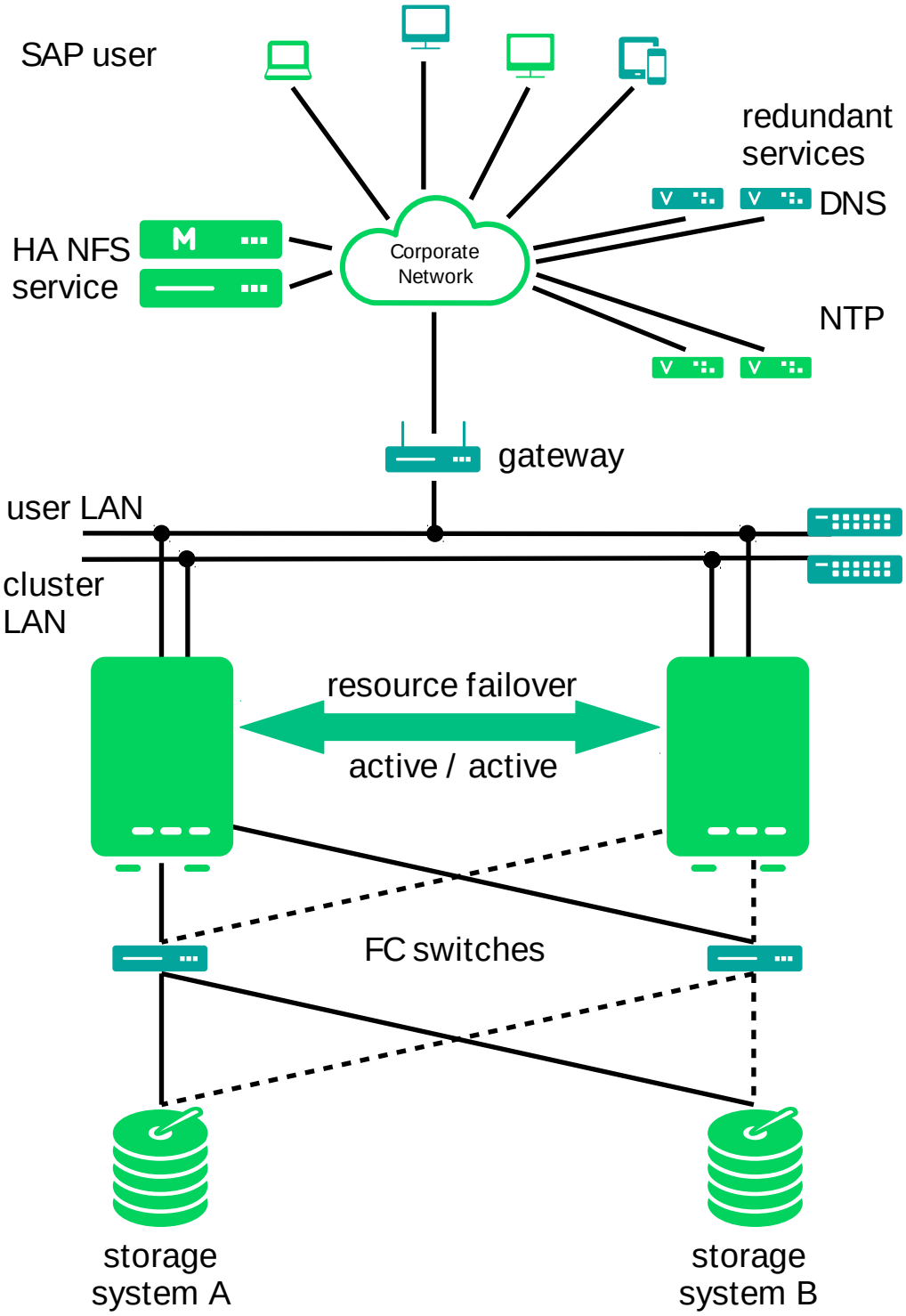

FIGURE 4.1: GENERAL ARCHITECTURE

The table below shows all possible failure scenarios and how the cluster deals with them.

TABLE 4.1: FAILURES AND MECHANISMS

Failure and Remarks	1st	2nd
Application crash on active node Local restart: Three times (could be tuned)	Resource restart	Resource failover
Active node crashes STONITH ensures that the node is really dead	STONITH - server fencing	Resource failover
Network outage, cluster inter-communication fails (Split-Brain)	STONITH - server fencing	Resource failover (occasionally)
Partial network outage on any node (one link fails) - Switch to 2nd link. The bonding mode defines, if the two links are used for parallel or alternative load	bond	
Partial network outage on active node (gateway not reachable) - Shutdown of active resources	Ping node	Resource failover
SAN outage on active node STONITH is triggered by file system monitoring and/ or SBD detects missing fencing disk	STONITH - server fencing	Resource failover
Partial SAN outage on any node (one link) Failover to second path	multipathing	
Power outage of the active node Automatically processed when using sbd as STONITH method. If not using sbd requires operator interaction if no external APC device for STONITH is used.	STONITH - server fencing	Resource failover
Split-Site The integration of SBD as STONITH device is recommended to solve this situation.	STONITH - server fencing via SBD	Resource failover (occasionally)

4.1.1 Server Fencing per STONITH Block Device

A STONITH Block Device (SBD) allows server fencing using shared storage. It prevents the cluster from corrupting the data. It also helps the cluster to handle servere failures on resources or nodes.

SBD uses a partition or LUN of the shared storage. The SBD device should be placed close to the data it should protect. Current implementations of the SBD are able to handle up to three devices, so a single failing SBD device will not lead into fencing of cluster nodes or a frozen cluster. Current implementations of the SBD also respect the cluster's quorum status. In case the SBD disk fails, the cluster will continue to work as long as it has the quorum. Thus, the impact of a failing SAN LUN is reduced. Second major advantage is, that server fencing works reliable in LAN-split, SAN-split, and complete-split scenarios.

The Simple Stack setup is decribed with two SAN storages, one at each site. So, two SBD devices are used here, one at each site. To set up SBD for minimum service interruption, the following should be taken into account:

- Configure the SAN boxes identically on both sites, having the exact same LUNs and partitions.
- Initialize the SBD devices in exact the same way on both SAN boxes.
- Configure the SBD according to Best Practices described in this document.

This gives:

- In a symmetrical LAN split scenario, the SBD device is not affected. The service could still run on each site. Nevertheless, one site will be fenced. Thus, a service will be migrated, if it was on the affected side.
- In a symmetrical SAN split scenario, the cluster will refuse to change anything. The service still is running, on each site.
- In a complete split brain scenario, one site will be fenced. Thus, a service will be migrated, if it was on that site.
- In a disaster scenario, a service will be migrated, if it was on the affected site.
- If the SBD devices are in-accesible for a certain time, while the LAN communication is still working, nothing happens. The service still is running, on each site.
- Data integrity is given in all scenarios.

SBD always has to be configured together with a watchdog. A third SBD might be used to meet additional requirements. More information on different SBD setups can be found in the manual pages sbd(8) and stonith_sbd(7).

4.2 Use Case 1 “Simple Stack Standalone”

To get started with SAP on SUSE Linux Enterprise Server for SAP Applications 12 the “Simple Stack standalone” use case shows the installation of an entire SAP system on SUSE Linux Enterprise Server for SAP Applications 12. All components (database, service instance and central instance) are installed on a single server.

The SUSE Linux Enterprise High Availability Extension is not needed for this use case. Nevertheless, everything is prepared to add high availability functionality later. It is important to follow the instructions for storage configuration carefully.

4.3 Use Case 2 “Simple Stack High Availability”

Use case 2, “Simple Stack High Availability” defines an entire SAP system (database and all cluster controlled instances) on a single cluster node, running within a single resource group.

The advantages of this cluster model:

- Less complex cluster design
- Easy to expand with additional SAP systems
- Avoids domino effects, if running one single SAP system in the cluster

Some disadvantages are:

- Less flexible in the view of SAP load balancing
- No enqueue replication server support
- Hierarchical dependencies within the resource group (Database, SCS, CI)

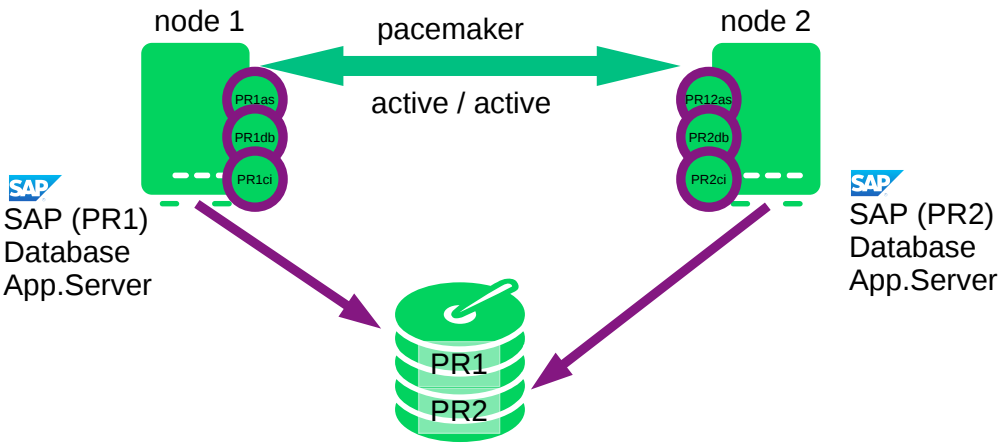

FIGURE 4.2: SIMPLE STACK HIGH AVAILABILITY WITH TWO SAP SYSTEMS ACTIVE/ACTIVE

II Installation of Simple Stack with MaxDB

5 Installation Overview

This part describes the installation of a SAP NetWeaver with MaxDB “Simple Stack High Availability” on SUSE Linux Enterprise Server for SAP Applications 12 SP1 for a proof of concept. The procedure is divided into the following steps:

- Planning
- Check prerequisites
- Download SAP NetWeaver installation media
- Install SUSE Linux Enterprise Server for SAP Applications 12 SP1
- Prepare SAN storage
- Install SAP NetWeaver and MaxDB
- Install SUSE Linux Enterprise Server for SAP Applications 12 SP1 on second node
- Install SUSE Linux Enterprise High Availability Extension cluster packages on both nodes
- Integrate SAP NetWeaver and MaxDB into the high availability cluster
- Check final result

As described earlier, the installation procedure could also be used to set up a non-clustered SAP NetWeaver (“Simple Stack standalone”) if the cluster-related steps are omitted.

6 Planning

Proper planning is essential for a well performing SAP system. For planning and support for your SAP Installation, visit http://service.sap.com [http://service.sap.com/] to download installation guides, review installation media lists and to browse through the SAP notes. This section focuses on aspects of planning a SAP installation.

The first major step is to size your SAP system and then derive the hardware sizing to be used for implementation. Use the SAP benchmarks (http://www.sap.com/solutions/benchmark/index.epx) to estimate sizing for a proof of concept. If you plan to migrate an existing SAP system you should first obtain or estimate the system characteristics of the “old” SAP system. The key values of these characteristics include:

- SAPS of the old SAP system
- Memory (RAM) size and usage of the old hardware
- Disk size, performance and usage of the old SAP system
- Network performance and utilization of the old hardware
- Language support (including Unicode)

If you have valid key values you can adapt these to the characteristics of your “new” SAP system. If you plan a new installation instead of a migration, you might need to adapt experiences with other SAP installations or use some of the published benchmarks as mentioned above.

- Estimate the SAPS in the new SAP system. This includes planning additional capacities, if needed. The calculation should also include estimated growth calculations, such as a boost of SAPS per year. Typical SAP growth is be between 10% and 20% per year.
- Choose RAM size, as well as disk size and performance for the Linux system. Also include a boost of the usage. Depending on the usage, the disk space may grow 30% per year. The disk size must also include the export and r3trans areas if they are not provided by other servers.
- Check if Unicode support is necessary for the new system.

7 Prerequisites

This chapter describes what hardware and software is needed for a proof of concept. It also lines out how to gather all information that is necessary to succeed.

7.1 Hardware Requirements

The final hardware sizing for production systems is always the responsibility of the hardware vendor. At the time this document was written, a base installation of NetWeaver on top of SUSE Linux Enterprise Server for SAP Applications 12 SP1 for a proof of concept could be done on the following hardware equipment:

- 2* CPU x86-64 2GHz (AMD64 or EM64T)
- 8GB RAM
- 100GB disk space for OS and empty SAP system
- some hundred GB up to some TB additional disk space is needed for test data
- 3* Gbit/s ethernet
- IPMI or ILO remote management board (optionally)

Although the cluster offers High Availability across two servers, each single server should already have the maximum possible availability using redundant devices:

- self correcting ECC RAM
- two or three redundant power supplies (connected to two circuits)
- several redundant cooling fans
- two or more internal disks with RAID(1/5/6/10) controller
- redundant LAN network controllers
- redundant LAN network links (connected to two switches)
- redundant SAN host bus controllers
- redundant SAN FC links (connected to two switches)

Check to use certified hardware. Information about certified hardware can be found in the SUSE YES database (http://developer.novell.com/yessearch/Search.jsp), in the SAP notes and on the hardware manufacturer's pages. Use certification notes from the SUSE YES database and the hardware manufacturer to select appropriate hardware components.

7.2 Software Requirements, Connection Data, and all the rest

Before you can start with the installation, make sure you meet the prerequisites.

- SAP documentation: Make sure that you have the latest version of the Master Guide by checking SAP Service Marketplace immediately before starting the installation. The Master Guide is regularly updated on SAP Service Marketplace at http://service.sap.com/instguidesnw70.
- Sizing data. After planning the installation, you should have this data.
- Configuration data
  - IP addresses, DNS names, Gateway
  - SAP SIDs
  - Linux Usernames (like **<sid>adm**), UIDs
- Connection data
  - SAN LUNs (names, LUN numbers) and multipath configuration parameters. There are some special parameters settings for multipath and SAN-HBA kernel modules depending on the hardware setup (SAN storage model and SAN setup). Please refer to configuration guides from SUSE and hardware vendors.
  - In addition to the network that connects the SAP servers to the clients; we recommend two additional dedicated network links between the two servers for cluster intercommunication. At least one additional dedicated network link is mandatory.

- Infrastructure such as DNS server, NTP server and a ping-able high available network node. This network node could be the gateway between the SAP system and the clients who need to access the service. If the gateway is no longer available, the service is not available. The cluster can determine which cluster node has (ping) connection to the pingnode and can migrate a service if needed.

- SAP Installation media (for details see the table in the next section). The SAP installation media can either be ordered as a physical CD/DVD or downloaded from http://service.sap.com/swdc. The next section describes the procedure for downloading the SAP media.

- SAP S-User (partner user) to download the media and installation guides and to browse through the SAP notes system. To download media, the S-User must have permission to download the installation media. Ask your company's SAP partner manager to create an S-User and to grant the proper rights.

- During the installation of the central instance of SAP NetWeaver you will be asked to provide a Solution Manager Key. You will need to create such a key for your combination of hostname (DNS name of the virtual IP address for high availability installations), SAP system ID (SID) and SAP instance number (like 00, 01, 02). The key could be created using your company´s Solution Manager, an additional SAP program. This document does not cover the installation of the Solution Manager. If you do not have access to your company´s Solution Manager, ask your internal SAP partner manager how to get a Solution Manager key.

- To download the SAP installation media, you will need the SAP download manager. A short description of the installation is integrated in the next section.

- To run the download manager you need a matching Java version. In former PoCs, SUN Java 1.6.0 (package `java-1_6_0-sun-1.6.0.u1-26`) worked very well. Have a look at the installation notes presented during the procedure to download the SAP download manager.

- An up-to-date patch level of the "SUSE Linux Enterprise Server for SAP Applications 12 SP1 installation. You will need:

  - A SUSE Customer Center (SCC) account

  - "SUSE Linux Enterprise Server for SAP Applications 12 SP1" installation media for x86-64 hardware platform

    - Or "SUSE Linux Enterprise Server 12 SP1" and "SUSE Linux Enterprise High Availability Extension 12 SP1" installation media for x86-64 hardware platform
    - You might need some additional hardware-specific driver updates
    - A software management software such as the subscription management tool (optional)
- To test the SAP system you either need to have a previously installed SAP client (`guilogon`, `guistart`) or you need to install this software on at least one of your workstations.

8 Download the Needed SAP Installation Media

This section describes in a more general view how to download the SAP installation media. As an example, we download the media for SAP NetWeaver 7.5 with database engine MaxDB RDBMS 7.9 for SUSE Linux Enterprise Server 12 x86-64.

We recommend you use a SUSE Linux Enterprise Server for SAP Applications 12 or newer to download the media. This section only describes the download using such a system.

1. Access the SAP download site https://support.sap.com/swdc
2. You are asked to login with your S-User and password (see prerequisites).
3. From the menu on the left, select Download → Installations and Updates → Entry by Application group.
4. From the Installation and Upgrades menu on the right, select SAP NetWeaver → SAP NetWeaver (yes, twice).
5. Now you have to select your NetWeaver version. In our example, we select SAP NetWeaver 7.5 → Installation and Upgrade → Linux → MaxDB. The version for your use case may differ, the example will follow the names and paths of NW 7.5.
6. From the media list, select the following media and click an “Add to download basket”.

TABLE 8.1: SAP NETWEAVER INSTALLATION SOURCES

Title	Number / File	Type
SWPM 1.0 SP10 for NW higher than 7.0x	SWPM10SP10_1-20009701.SAR	SAR
SAP Kernel 7.45 Linux on x86_64 64bit - NW 7.5	51050082_3	ZIP
MaxDB 7.9 - SP8 Build 35 Linux on x86_64 64bit	51050545_8	ZIP
SAPCAR	SAPCAR_617-80000935.EXE	EXE

The total amount of installation sources is 10GB for the chosen NetWeaver 7.5 with MaxDB. To unpack the archives, roughly twice the disk space is needed. Other products might need more space.

7. After some time, a pop-up with two buttons appears. Press "Download Basket".
8. Your selected media is shown in your download basket. If you haven't installed the SAP download manager yet, you will have to download and install it now.
    - Click “get download manager” in this case. The SAP Download Manager Installation Guide is shown.
    - Check the section prerequisites and the SAP Download Manager installation guide.
    - You need a Java version that fits SAP needs. Download the Linux version.
    - You get a self extracting archive that starts after the download. Follow the installation steps. We have installed the Download Manager in the local home directory, “SAP_Download_Manager”.
9. Start the installed SAP Download Manager using the command **`~/SAP_Download/Manager/Download_Manager`**.
10. If you start the SAP Download Manager for the first time, you will need to provide some credentials such as the SAP Marketplace address (https://support.sap.com), your S-User, your S-User-Password and the Data Store (directory to place the downloaded files).
11. Press the "download all objects" button (the button with two right arrows).
12. Now it's time to be very patient.
13. After the download you will have to unpack the downloaded files using unzip (for ZIP type) and unrar (for EXE and RAR type). Unrar is able to skip the self extracting code in the EXE files and will include all files of a media set such as 1/2, 2/2.
14. Copy (rsync) the extracted files to your system to be installed or create NFS exports on the installation source and NFS mounts on the target systems. In our setup we use:
    - /sapcd/SWPM for the Software Provision Manager (SWPM)
    - /sapcd/Kernel for the NW 7.5 kernel

- /sapcd/MaxDB for the MaxDB engine
- /sapcd/InstExp NW 7.5 installation export

9 Install SUSE Linux Enterprise Server for SAP Applications 12

In this section we describe relevant configuration parameters of SUSE Linux Enterprise Server for SAP Applications 12 for a simple stack scenario. The installation procedure itself should be well known to the target audience, and thus is not described. Please refer to the product documentation for general information on SUSE Linux Enterprise Server for SAP Applications 12 (https://www.suse.com/documentation/sles-12/).

9.1 File System Layout

For a proof of concept the following partition and file system layout for the OS could be chosen:

TABLE 9.1: OS FILE SYSTEM LAYOUT

Mount Point	Size	File System
/	60GB	btrfs
swap	2*RAM (beyond 32GB RAM see SAP note 1597355)	swap

The swap space is 2 * RAM for smaller systems up to 32 GB RAM. A more detailed list for bigger systems is included in the SAP Note 1597355 - Swap-space recommendation for Linux.

SUSE Linux Enterprise Server for SAP Applications 12 needs ca. 4.5GB disk space. The size of `/boot` depends on the number of kernels that should be installed in parallel. Each kernel needs ca. 35MB disk space in /boot . The size of /var depends on the amount of log data and application specific usage, 5GB or more are appropriate. If the SAP NetWeaver installation sources should be put on the local disk, 20GB additional free space is needed. We use the directory link /sapcd in our examples.

Besides the usual OS file systems, SAP and the SAP databases require their own file systems. These file systems are not stored locally. Instead they are provided by NFS fileservers or on LUNs in Storage Area Networks (SAN). Typically we need for SAP:

- `/sapmnt`
- `/usr/sap/SID`
- `/sapdb` (for MaxDB. Of course, for Oracle and DB2 there are other necessary pathes.) File system sizes depend on the use case. The database file system can be from 100GB up to multiple TB. After a fresh installation, around 30GB are in the database.

9.2 Software Selection

Its a good practice to install as less software as needed or recommended. We de-select the „GNOME Desktop Environment“, while we plan to run the server in runlevel 3 and not to use local Gnome sessions. Since we install “X Window System” we still can start X-Window based applications local or remotely. We also de-select “AppArmor” here, because we do not plan to use it. On the other hand SAP recommends some software to be installed. The resulting pattern selection should be:

```
+ Base System pattern
+ X Window System pattern
+ 32-bit Runtime Environment pattern
+ Printing pattern
+ SAP NetWeaver Server Base pattern
+ SAP Application Server Base pattern
+ C/C++ Compiler and Tools pattern
- No AppArmor pattern
- No Gnome Desktop Environment pattern
```

Finally you get a pattern list as shown in the figure:

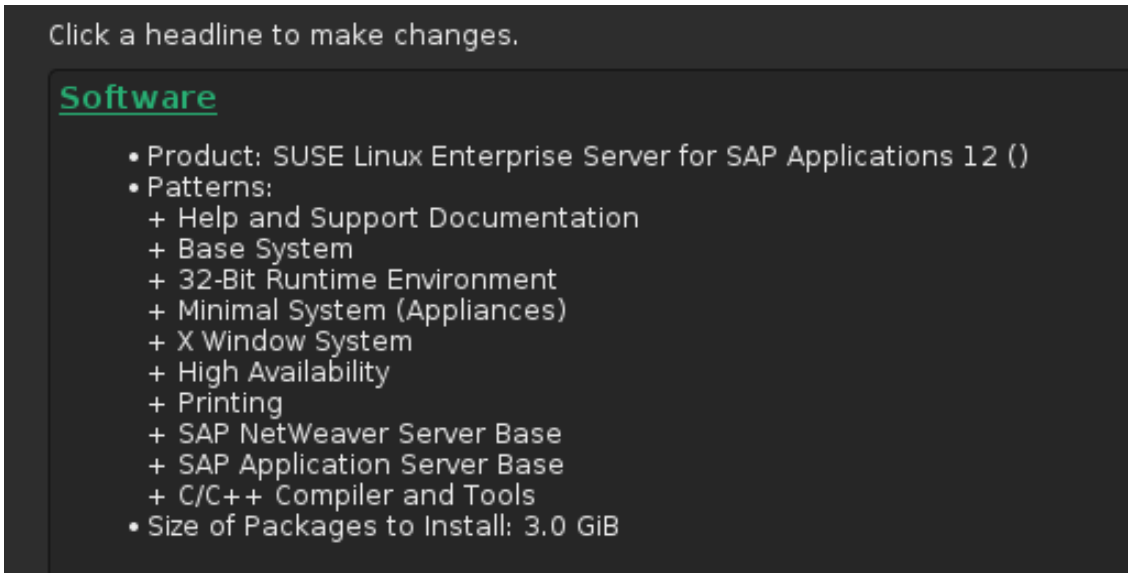

FIGURE 9.1: SUSE LINUX ENTERPRISE SERVER FOR SAP APPLICATIONS 12

- sapconf , prepares OS for SAP needs
- sap-locale, contains special code pages, only needed for non-unicode systems

If you plan to run SAP application servers on SUSE Linux Enterprise Server for SAP Applications 12 together with application servers on another OS, see SAP Notes 1069443 and 187864 on how to get the correct code pages.

If you plan to extract the SAP installation sources, you should also install the package `unrar`.

The package orarun conflicts with the SAP requirements and should not be installed.

It is recommended to update the complete operating system to an actual level. This could be done by either connecting to the SCC via internet, or by using a locally installed update proxy, f.e. SUSE Manager (https://www.suse.com/documentation/suse_manager/ or SMT (https://www.suse.com/solutions/tools/smt.html). The update procedure should be well known to the target audience, and thus is not described in this document. For information on SCC look into the SUSE documentation (https://scc.suse.com/docs/help).

As of the publication date on this document, you should have at least the following releases of the core operating system:

- kernel-default-3.12.53-60.30.1.x86_64
- lvm2-2.02.120-60.1.x86_64

- multipath-tools-0.5.0-46.1.x86_64
- dracut-037-66.2.x86_64
- device-mapper-1.02.97-60.1.x86_64
- glibc-2.19-35.1.x86_64
- nfs-client-1.3.0-18.1.x86_64
- libuuid1-2.25-30.1.x86_64
- uuidd-2.25-30.1.x86_64

You should also have at least the following releases of the additional software:

- sapconf-4.1-8.1.noarch
- sap-locale-1.0-92.4.x86_64
- gcc48-4.8.5-24.1.x86_64
- libgcc_s1-5.2.1 + r226025-4.1.x86_64

9.3 Default system target and system services

Recommended target is "`multi-user`". You can either use YaST (**`yast service-manager`**) or systemctr (**`systemctl set-default`** *`NAME`*) to set the default system target.

The services `ntpd`, `uuidd` and `multipathd` should be enabled. The Network Time Protocol (NTP) service has to be set up correctly.

9.4 Miscellaneous

System language has to be `en_US`. The IO scheduler should be set to either noop or deadline at boot time. If the cgroups memory controller is not used, it could be disabled at boot time.

9.5 Check SUSE Linux Enterprise Server for SAP Applications 12

The installed system should be checked. All selections and parameters mentioned in the chapters beforehand have to be correct. A script could help to collect the relevant data. An example is the script cs_precheck_for_sap from the package ClusterTools2. This package is included in the update channel of SUSE Linux Enterprise Server for SAP Applications 12.

For documentation of the general OS configuration, the supportconfig script from the supportutils RPM could be used. The supportutils RPM is part of SUSE Linux Enterprise Server for SAP Applications 12.

10 Prepare SAN Storage

In this step we will configure SAN access on the first cluster node for a proof of concept. We want to achieve:

- System has multipathed SAN access
- MD is prepared for cluster integration
- Logical Volumes and file systems on SAN are configured for SAP “Simple Stack” scenario

Please refer to the product documentation for general information on storage administration with SUSE Linux Enterprise Server for SAP Applications 12 (https://www.suse.com/documentation/sles-12/stor_admin/data/stor_admin.html).

10.1 Multipath Configuration

Enable multipathing service multipathd.

The multipathing drivers and tools support most storage arrays. Consult the vendor’s hardware documentation to determine what settings are required. When storage arrays are automatically detected, the default settings for multipathing apply. If you want non-default settings, you must manually configure the /etc/multipath.conf file.

Storage arrays that require special commands on failover from one path to the other or that require special nonstandard error handling might require hardware handlers. Consult the hardware vendor’s documentation to determine if its hardware handler needs to be installed.

To find the correct devices and blacklist section entries, talk to your SAN administrator.

In our setup we have decided to use no-path-retry to avoid service interruption in case of short-time path failures. Nevertheless path failure will result in read errors, which are detected by the upper layers of the I/O stack. Please read the Linux Multipath documentation to decide whether to use this option. We also set the failback time to a re-initiated path to value greater than zero. This setting reduces the chance to run into I/O errors if a single SAN path starts frequently to go up and down in a very short time (path flapping). This setting can be modified in the failback parameter.

Configure multipath to use user friendly names. Definition of user friendly names should be done in a separate bindings file (`/etc/multipath.bind`), if you have more than a few devices. See man multipath.conf (bindings_file parameter) for details. It is mandatory to assign names to all LUNs. Mixing unassigned LUNs with assigned ones maybe leads to erratic behavior. A sample configuration for an HP EVA storage looks like this:

Details on multipath related Linux configuration can be found in the product documentation (https://www.suse.com/documentation/sles-12/stor_admin/data/cha_multipath.html).

```
#/etc/multipath.conf
defaults {
        dev_loss_tmo 10
        fast_io_fail_tmo 5
        user_friendly_names yes
        bindings_file /etc/multipath.bind
}
blacklist {
        devnode "^(ram|raw|loop|fd|md|dm-|sr|scd|st)[0-9]*"
        devnode "^cciss!c[0-9]d[0-9]*"
        devnode "^dcssblk[0-9]*"
}
devices {
     device {
        vendor         "HP|COMPAQ"
        product        "HSV1[01]1 (C)COMPAQ|HSV2[01]0|HSV300|HSV4[05]0"
        path_grouping_policy    group_by_prio
        getuid_callout          "/lib/udev/scsi_id -g -u /dev/%n"
        path_checker            tur
        path_selector           "round-robin 0"
        prio                    alua
        rr_weight               uniform
        failback                immediate
        hardware_handler        "0"
        no_path_retry           5
        rr_min_io_rq            1
     }
```

```
}
```

Note

This configuration is used for a particular environment only. Multipath configuration has to follow the hardware manufacturers recomendations and has to be aligned with the storage administrator's concepts. For hosts that are part of an HA clusters some settings will differ from stand-alone hosts. First, the common **`queue_if_no_path`** default sould be replaced by a certain value for **`no_path_retry`**. Second, the time multipathing needs to recover path failures should be shorter than the monitoring timeout of the storage stack resource agents. Otherwise a path failure could lead to node fencing in worst case. On the other hand, sporadic path flapping should not lead to permanently disabled pathes.

To fine-tune the multipath behaviour, the number of retries for a failed path (**`no_path_retry`**), the retry interval, the failback time to a re-initiated path, and the failback policy could be set. Details for specific hardware can be found in the multipath.conf man page (**`man 5 multipath.conf`**). Please read also the support TID on multipath timeouts (https://www.suse.com/support/kb/doc.php?id=7016305) .

Usually it is a good idea to start without any device section, but use the compiled-in defaults.

To make configuration changes or changes inside the SAN visible, you may have to flush the multipath tables. After you modify the `/etc/multipath.conf` file, you must use the command dracut to re-create the INITRD on your system. Please refer to the documentation mentioned above for details.

10.2 Partitioning

Some SAN storages require partition alignment for performance reasons. Usually this alignement should be done automatically when using fdisk, parted or YaST. In this document only one single LUN is used for each side of the mirror. So, we use the partition here for data.

```
# fdisk /dev/mapper/sapvol1
 ...
 Disk /dev/mapper/sapvol1: 214.7 GB, 214748364800 bytes
255 heads, 63 sectors/track, 26108 cylinders
Units = cylinders of 16065 * 512 = 8225280 bytes
```

```
Device         Boot Start   End     Blocks   Id System
/dev/mapper/sapvol1-part1           1    26108    1004031   83 Linux
# partprobe
```

Repeat this for the second LUN.

10.3 MD Configuration

- Mount file system in order to create sub-mountpoints beneath
- /etc/mdadm.conf must contain a line to disable scanning and automatic assembling of MD devices. The file should also contain an information, where the configuration files are placed and why:

```
# /etc/mdadm.conf
# Never add any devices to this file
# Cluster mdadm configuration can be found
# in /clusterconf/<sapinstance>/mdadm.conf
#
# MD-Devices, that are not under cluster control are stored
# in the file /etc/mdadm.conf.localdevices
#
# Prevent mdadm from finding devices by auto-scan:
DEVICE /dev/null
#
```

- Verifiy LUNs in `/dev/mapper` (names have to match exported names from storage systems)
- Create MD arrays using **`mdadm`**. Use **`mdadm –metadata=1.2`** . Use the suffix „p<n>“ f.e. „p2“ for the partition with **`mdadm`**. You have to use the metadata format 1.2 or above.

```
# mdadm --create /dev/md0 --metadata=1.2 --level=1 --raid-devices=2 \
--bitmap=internal \
       --force /dev/mapper/sapvol1_part1 /dev/mapper/sapvol2_part1
```

- Check the state of the array and note the RAID UUID.

```
# mdadm --misc --detail /dev/md0
```

- Create `/clusterconf/SID/mdadm.conf` with the following format. Each of the DEVICE and ARRAY definitions must be complete in one single line. In the example below the backslash is used only to show where the lines have to be concatenated.

```
# mkdir -p /clusterconf/SID
# vi /clusterconf/SID/mdadm.conf
DEVICE /dev/mapper/sapvol1_part1 /dev/mapper/sapvol2_part1
ARRAY /dev/md0 level=raid1 num-devices=2 \
UUID=dfa428a2:f74a4b42:48fd23c4:49003a8b
```

  The file must contain two lines. The first starts with DEVICE, the second starts with ARRAY. The backslash is printed here to show that the line is continued.
  Don't forget to edit the UUID! (Fetch it with **`mdadm --misc --detail /dev/md<x>`**). The UUID of a specific MD device is calculated when the MD device is created (**`mdadm --create ...`**). It changes for each new creation, even for the same MD device.
  Naming of partitions is handled somewhat inconsistant among several tools. Inside the config file, the suffix „part<n>“, is used to name the partition. If the MD device can not be assembled after the first reboot, check `/dev/mapper/` for the correct names.

- Check the status in `/proc/mdstat`.

10.4 LVM Configuration

- Create PVs using pvcreate on MDs.
- Create VGs using vgcreate on PVs.
- The logical extent size could be set to something larger than 8MB, f.e. 64MB.
- Create LVs using lvcreate on VGs.

```
# pvscan
# lvcreate -L 100G -n sapdb sapvg
# lvcreate -L 10G -n sapmnt sapvg
```

```
# lvcreate -L 10G -n usrsap sapvg
```

- Adapt LVM filter.
  The Linux Volume Manager automatically tries to find Volume Groups and Logical Volumes on all devices in `/dev`. To avoid this, make sure that you adjust the LVM filter in the file `/etc/lvm/lvm.conf` to something like this: **`filter = [ "a|/dev/sda[1-4]|", "a|/dev/md.*|", "r|/dev/.*|" ]`** . This filter avoids scanning for VGs in `/dev/disk*` directories. If you are using VGs for local file systems on your internal hard drives, make sure to add the local devices to this filter (**`a|/dev/<my_device>`**).

10.5 File System Creation

Note

XFS is a good choice for filesystems that contain data.

- Create xfs file system using mkfs.xfs on the LVs

```
# mkfs.xfs /dev/sapvg/sapdb
# mkfs.xfs /dev/sapvg/sapmnt
# mkfs.xfs /dev/sapvg/usrsap
```

  XFS supports online extention of file systems.

- Create mount points

```
# mkdir -p /sapdb /sapmnt /usr/sap
```

- Manually mount new file systems

```
# mount -onoatime /dev/sapvg/sapdb  /sapdb
# mount -onoatime /dev/sapvg/sapmnt /sapmnt
# mount -onoatime /dev/sapvg/usrsap /usr/sap
```

The option noatime reduces the number of IOs, because no access time entries are written for read access. See manual page mount(8) for details.

- Create sub-directories in the correct place

```
# mkdir /usr/sap/SID
# mkdir /var/sapcd
```

- If the file systems were already created before, do not forget to remove the `/etc/fstab` entries.

10.6 Check the SAN Storage

- Unmount the file systems, de-activate the VG, stop the MD RAID device, stop multipathing on the first node, and check the status.

```
# df -h; df -i
# umount /sapdb /usr/sap /sapmnt
# vgchange -a n sapvg
# mdadm --manage --stop /dev/md0
# systemctl stop multipathd
```

- Restart the multipathing and the MD RAID device, activate the VG, mount the file systems, and check the status.

```
# systemctl start multipathd
# multipath -ll
# mdadm --assemble --config /clusterconf/SID/mdadm.conf /dev/md0
# cat /proc/mdstat
# vgchange -a y sapvg
# lvs
# mount -onoatime /dev/sapvg/sapdb  /sapdb
# mount -onoatime /dev/sapvg/sapmnt /sapmnt
# mount -onoatime /dev/sapvg/usrsap /usr/sap
```

```
# df -h; df -i
```

- Finally, test if SAN access works with reasonable speed and without errors. The size of the test file should be at least 1.5 times the RAM size to get an reliable speed estimation. Do not forget to remove the test file.

  ```
  # dd if=/dev/zero of=/sapdb/test.dd bs=256M count=64 oflag=direct
  ... 245 MB/s
  # rm /sapdb/test.dd
  # grep "I/O error" /var/log/messages
  ```

  You should see no errors. Meaning of reasonable speed may vary. At the time this document was written, sustained linear write rate should be expected between 200 MB/s and 500 MB/s. Maybe this could be discussed with the SAN storage administrator.

Once the storage has been set up, the following mount points should exist:

- /usr/sap/
- /sapdb/
- /sapmnt/
- /etc/fstab

And the following storage related configuration files should be adapted:

- /etc/multipath.conf
- /etc/multipath.bind
- /etc/mdadm.conf
- /clusterconf/*SID*/mdadm.conf
- /etc/lvm/lvm.conf

For documentation of the storage related OS configuration, the supportconfig script from the supportutils RPM could be used. The supportutils RPM is part of SUSE Linux Enterprise Server for SAP Applications 12.

11 Install SAP NetWeaver 7.5

In this section we describe the installation of SAP NetWeaver 7.5 on SUSE Linux Enterprise Server for SAP Applications 12 SP1 in a “Simple Stack standalone” scenario. All components are placed on one single system. It is prepared to be integrated in a high availability cluster according to the SAP “Simple Stack High Availability” scenario. We need to install the SAP system components using several virtual host names to match the high availability installation needs.

11.1 SAP Installation Procedure

In this section we describe the installation process step-by-step for an easy, typical SAP installation. The installation itself will take multiple hours depending on the performance of your hardware. If you plan to disconnect your workstation (like a laptop) from the network during the installation, you should use an X-session manager like NoMachine NX or vncserver.

To cover High-Availability aspects and also to be able to place some of the components of the SAP system on different hosts later, we use the installation workflow “High Availability System”. This leads into a separate installation of the central services instance (ASCS), the database and the central instance (DVEBMGS).

- The three IP addresses for the SAP virtual hostnames have to be up and the names have to be resolved in `/etc/hosts`.
- You need a root session with X-forwarding. So use either **`ssh`** to login to the server system, including forwarding options (like -Y or -X).

```
# ssh -Y root@server
```

- Enter the mounted installation master directory. The correct path depends on your selections made above. If you are following our example, the path is: `/sapcd/InstMa`.

```
# cd /sapcd/InstMa
```

- The installation master for UNIX systems could be used for installations on AIX (PPC64), HPUX (PARISC), Linux (i386, ia64, PPC64, S390_64 and x86-64), OS390_32, OS400 (PPC64), Solaris (SPARC), Solaris (x86-64) and Windows (i386, ia64 and x86-64). In our sample we select Linux for x86-64 architecture and enter the directory IM_LINUX_X86_64.

```
# cd IM_LINUX_X86_64
```

11.2 Installation of the Central Services Instance ASCS

- Start the SAP Installation Master by starting the **sapinst** command: **./sapinst SAPINST_USE_HOSTNAME=sap<sid>as**. Instead of "**sap<sid>as**" you use your virtual hostname reserved for the ASCS. At the top of the screen you will see the 5 major phases of the installation procedure. The highlighted number represents the active step. The first step is named "1 - Choose Option".

- In the left area select "SAP NetWeaver 7.5 " → "SAP Application Server ABAP" → "Max DB" → "High-Availability System" → "Central Services Instance for ABAP (ASCS)". The screen shot shows the expanded service tree. If you have selected "Central Services Instance for ABAP (ASCS)" click at the "Next" button on the left side.

- In the next dialog you reach the major phase "2 - Define Parameters". The first dialog of this step is labeled "Parameter Mode > Default Settings". There are two radio buttons to select the parameter mode, "typical" and "custom". In this document we prefer the "typical" parameter mode to reduce the selection of prompts. Click "Next" to proceed with the following dialog.

- The dialog "SAP System > General Parameters" is used to define the SAP system identifier SID (we use "NA2" in our example) and the SAP system mount directory (we use /sapmnt). As we install an unicode system we activate the checkmark. Click "Next" to proceed.

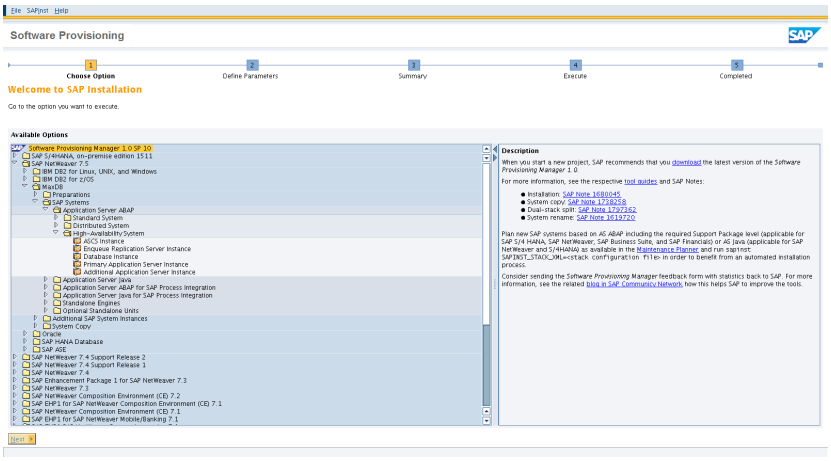

FIGURE 11.1: START DIALOG OF THE SAP INSTALLATION MANAGER

- The Dialog “SAP System > Administrator Password” provides the definition of the password of the Linux user <sid>adm. You may also define the unique user ID and the group ID for the Linux user group sapsys. Click “Next” to proceed.

- Now the installer asks for the two-digit instance number for the central servoces instance (ASCS). In our example we use 00. Click “Next” to proceed.

- In the dialog “SAP System > ASCS Instance” do not change the port numbers for the ASCS if you want to follow our example. Just click “Next” to proceed.

- The “Media Browser > Software Package” dialog asks for an additional path, the install media for the NetWeaver Kernel (ABAP). In our sample the path is /sapcd/kernel. You can either enter the path directly to the input filed or use the file browser. After you have provided the correct path, click “Next” to proceed.

- The installation is now in the phase “3 - Summary”. The last step before the installation of the central services instance (ASCS) has the dialog title “Paramater Summary”. Please double ckeck the settings. If everything is correct, click “Next” to proceed.

- The installer switches to phase “4 - Execute”. The dialog is named “Task Progress”. This dialog provides an overview about the installation progress and about the status of the scheduled tasks. The status bar at the bottom of the window may also give some detailed information for taks which are running very long.

- If the installation is successful, the installer switches to phase “5 - Completed”.

11.3 Installation of the Database

- Start the SAP Installation Master by starting the sapinst command: **`./sapinst SAPINST_USE_HOSTNAME=sap<sid>db`**. Instead of "**`sap<sid>db`**" you use your virtual hostname reserved for the database. At the top of the screen you will see again the 5 major phases of the installation procedure. The highlighted number represents the active step. The first step is named "Choose Service".

- In the left area select "SAP NetWeaver 7.5" → "SAP Application Server ABAP" → "Max DB" → "High-Availability System" → "Database Instance". The screen shot shows the expanded service tree. If you have selected "Database Instance" click at the "Next" button on the left side.

- The installer switches to phase "2 - Define Parameters". The dialog is titled "SAP System > General Parameters". Please check, if the checkmark "Profiles are available" is set and provide the "Profile Directory" path. Following out example "/sapmnt/<SID>/profile" is a good choice. Of course you have to replace "<SID>" with your SAP system ID. Click "Next" to proceed.

- The dialog "SAP System > Master Password" is used to set the start password for Linux, SAP, the database (such as **`<sidadm>`**, **`<sid>user`**, ...) and SAP system Users (such as DDIC). Click "Next" to proceed.

- The next dialog is "SAP System Database". You should provide a database ID (in our example we use **`<SID>`**. The "Database Host" should be the one, which you provided on the command line while starting sapinst (sap<SID>db in our example). Click "Next" to proceed.

- The dialog "MaxDB > Database Software Owner" is used to specify the name, group and IDs of the Linux user "owning" the software. Typically the values should be "sdb" and "sdba". If you do not intend to set specific user and group IDs let the installer choose those values. Click "Next" to proceed.

- The dialog "MaxDB > Database Instance Software Owner" is used to specify the name, password and IDs of the Linux user "owning" the database instance. Typically the values should be "**`sqd<sid>`**" and "**`sdba`**". If you do not intend to set specific user and group IDs let the installer choose those values. Click "Next" to proceed.

- The “Media Browser > Software Package” dialog asks for the full path of the installation media “Installation Export NW...”. In our sample the path is `/sapcd/InstExp`. Click “Next” to proceed.
- The installer asks for the target path for the database and database instance installation. In our example we choose “`/sapdb`”. Click “Next” to proceed.
- The “Media Browser > Software Package” dialog asks for an additional path, the install media for the MaxDB RDMBS. In our sample the path is `/sapcd/MaxDB`. Click “Next” to proceed.
- Now the installer asks for the passwords to be used for database users “**`superdba`**” and “`control`”. Click “Next” to proceed.
- The dialog “MaxDB > Database Parameters” could be used to provide some major installation and configurtion parameters. The most important one in the view of this document is the “Volume Medium Type”. This parameter must be set to “File System”, if you want to follow our example. Click “Next” to proceed.
- The “MaxDB > Log Volumes” dialog can be used to tune and size the database log area. In our example we do not change any values here. Click “Next” to proceed.
- The dialog “MaxDB > Data Volumes” let us tune and size the database files for objects like tables, indexes and so on. In our example we do not change any values here. Click “Next” to proceed.
- The installer show the dialog “MaxDB > ABAP Database Schema Password” and asks for the password of **`SAP<SID>`** scheme. Click “Next” to proceed.
- The dialog “SAP System > Database Import” you could define the codpage and the number of maximum parallel import jobs. In our example we do not change any values here. Click “Next” to proceed.
- The “Media Browser > Software Package” dialog asks for an additional path, the install media for the NetWeaver Kernel (ABAP). In our sample the path is `/sapcd/kernel`. You can either enter the path directly to the input filed or use the file browser. After you have provided the correct path, click “Next” to proceed.
- The dialog “SAP System > Unpack Archives” should show a list of archives to be unpacked. Normally you do not need to change anything here. Click “Next” to proceed.

- The installation is now in the phase “3 - Summary”. The last step before the installation of the central services instance (ASCS) has the dialog title “Paramater Summary”. Please double ckeck the settings. If everything is correct, click “Next” to proceed.
- The installer switches to phase “4 - Execute”. The dialog is named “Task Progress”. This dialog provides an overview about the installation progress and about the status of the scheduled tasks. The status bar at the bottom of the window may also give some detailed information for taks which are running very long. The installation of the database software is quite fast, but the step “Import ABAP” could take multiple hours depending on the performance of your hardware. The installer GUI should stay open to get either the final success message or an error information.
- If the installation is successful, the installer switches to phase “5 - Completed”.

11.4 Installation of the Central Instance

- Start the SAP Installation Master by starting the sapinst command. **`./sapinst SAPINST_USE_HOSTNAME=sap<sid>ci`** Instead of “**`sap<sid>ci`**” you use your virtual hostname reserved for the central instance. At the top of the screen you will see the 5 major phases of the installation procedure. The highlighted number represents the active step. The first step is named “Choose Service”.
- In the left area select “SAP NetWeaver 7.5" → “SAP Application Server ABAP” → “Max DB” → “High-Availability System” → “Central Instance”. The screen shot shows the expanded service tree. If you have selected “Central Instance” click at the “Next” button on the left side.
- The dialog “SAP System > General Parameters” is used to define “Profile Directory” (we use **`/sapmnt/<SID>/profile`**). Click “Next” to proceed.
- In the dialog “SAP System > Master Password” you again provide the already defined password for all users which are created during the installation. Click “Next” to proceed.
- The dialog “SAP System > Central Instance” provides a list of already installed SAP instances. Please specify the “Central Instance Number”. In our example we use “01”. Click “Next” to proceed.

- The installer will show the dialog “SAP System > DDIC Users”. In our example the ceckmark “DDIC user has a password different from default” should not be set. This tells the installer to use the master password provided earlier. Click “Next” to proceed.
- The installer needs to kow the password used during the “Database Instance” installation. Provide the master password here, if you have used the same password.Click “Next” to proceed.
- The dialog “MaxDB > ABAP Database Schema Password” is used to provide the password defined during the “Database Instance” installation. Type in the master password here, if you have used the same password.Click “Next” to proceed.
- The “Media Browser > Software Package” dialog asks for an additional path, the install media for the NetWeaver Kernel (ABAP). In our sample the path is /sapcd/kernel. You can either enter the path directly to the input filed or use the file browser. After you have provided the correct path, click “Next” to proceed.
- The dialog “SAP System > Unpack Archives” should show a list of archives to be unpacked. Normally you do not need to change anything here. Click “Next” to proceed.
- The installation is now in the phase “3 - Summary”. The last step before the installation of the central services instance (ASCS) has the dialog title “Paramater Summary”. Please take time to double check the settings. If everything is correct, click “Next” to proceed.
- The installer switches to phase “4 - Execute”. The dialog is named “Task Progress”. This dialog provides an overview about the installation progress and about the status of the scheduled tasks. The status bar at the bottom of the window may also give some detailed information for taks which are running very long.
- During this phase the is one more interactive task: The installer needs a valid “Solution Manager Key” (which has been named in the prerequisites). You need to create such a “Solution Manager Key” using your local “Solution Manager”. This is an own SAP product which is used for central SAP system maintenance. This document does not include the installation of this product. Please use the SAP installation documentation, if you do not already have installed your “Solution Manager”.
- If the installation is successful, the installer switches to phase “5 - Completed”.

11.5 Check the SAP Installation

The command „**dpmon -pf=<instance_profile>**“ can be used to check the status of the SAP processes for each application instance (ASCS, CI/PAS) of an ABAP system. It allows to see the result of transaction SM50 without logging on to the SAP system.

- Use the SAP client to access the installed SAP system. Depending on your SAP client installation (which has been named in the prerequisites), you can start the SAP session either with guilogon to define the SAP client connections. With guistart you can access your SAP system directly if you already know the SAP logon string:

```
# export PLATIN_JAVA=/usr/lib/jvm/java-1_5_0-ibm-1.5.0_sr8a/bin/java;
# guistart  /H/<server>/S/<port>
```

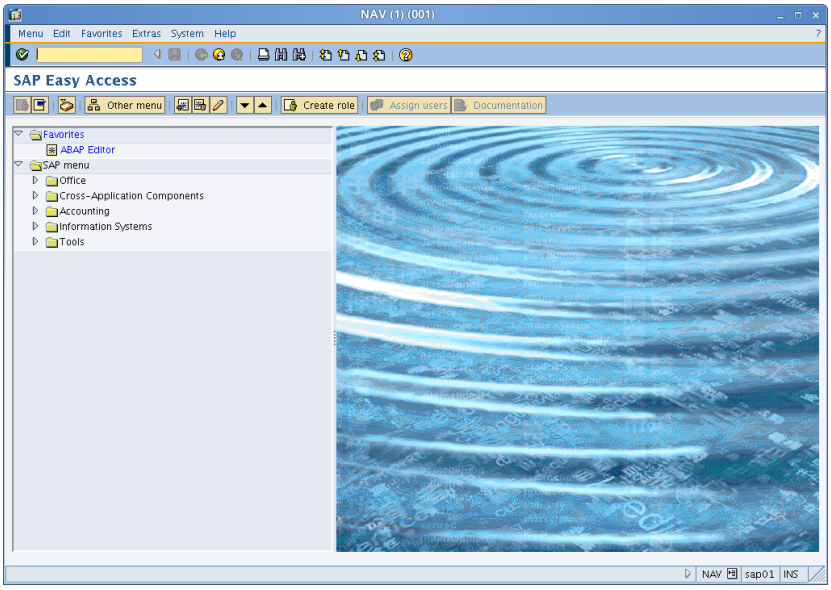

FIGURE 11.2: SAP CLIENT

Where server is your (virtual) server name and port is the SAP Port. The SAP Port can be calculated as: port = 3200 + instance_number. In case the SAP central instance has the instance number 02 and our SAP server is named sap01. So the correct access string is: /H/sap01/S/3202. Login as user DDIC and use the password given during the installation.

- Check stop of SAP and database processes. Login as user **<sid>adm** (here: **na2adm**) and stop ASCS, DB, and CI.

```
# stopsap r3 DVEBMGS01 sapna2ci
```

```
# stopsap r3 ASC00 sapna2as
# stopdb
```

- Check start of SAP and database processes. Double check for the three virtual IP addresses needed by the SAP processes. Login as user **<sid>adm** and start ASCS, DB, and CI.

```
# startsap r3 ASC00 sapna2as
# startdb
# startsap r3 DVEBMGS01 sapna2ci
```

Of course, you should check for the processes again.

12 Install the Second Node

This section gives hints for installing SUSE Linux Enterprise Server for SAP Applications 12 on the second node and describes how to duplicate the files and directories needed on the local disk to run SAP NetWeaver. It is important to install the second node correctly and identically to the first one.

12.1 Install SUSE Linux Enterprise Server for SAP Applications 12 on the Second Node

You can either repeat the manual installation for the second node or speed up some of the tasks by using AutoYaST. An AutoYaST control file was stored on the first node as part of the normal installation process. That control file can easily be adapted for the second node. Two or three items have to be changed, depending on the particular setup. Usually IP adress, disk ID, and MAC address need to be adpated. Besides speedup, using AutoYaST will eliminate the risk of mis-configuration and differences between the two cluster nodes. Do not forget to apply the exact same software updates on both nodes.

Check which LUNs are available for both nodes. Do not rely on looking at the first and last four digits. Use grep instead, to check for exact matching.

```
node1:~ # multipath -ll
node2:~ # multipath -ll | grep <LUN from first node>
```

12.2 Duplicate SAP Related Files to the Second Node

Create users and groups on the other node as they were created by the SAP installation on the first node. Use the same user ID and group ID. Be careful to use the same user-IDs and group-IDs in /etc/passwd and /etc/group.

Check login user profiles of **<sid>adm** and **sqd<sid>** on **<node1>**, they should be similar. Depending on the Installation Master CD that was used for the SAP installation, the logon profiles for the SAP Administrator user (**<sid>adm**) and the database administrator user might

be different. In older and non-high availability installations the user logon profiles look similar to .sapenv_hostname.csh . Optionally repair profiles. Ensure that the same user login profiles for the **<sid>adm** and **sqd<sid>** user are on all nodes. You can do so by copying the /home/<sid>adm to **<node2>**.

Copy the /etc/services or its values which were adjusted by SWPM (**sapinst**; see SAP related entries at the end of the file) to all nodes.

There are other directories within the SAP file system which has to be configured. These directories belong to specific SAP services, so their configuration depends on the particular SAP landscape. To set up systems quickly in no greater context or SAP landscape, it's sufficient to just copy them to the other node.

Create all needed mount points on the second node (/sapmnt, /sapdb, /usr/sap).

In case of choosing MaxDB as database, files and directories have to be synchronized, too. Copy the file /etc/opt/sdb to the other node.

Make the directory (with content) /sapdb/programs/lib available even if the file system /sapdb is not mounted. To do so, mount /sapdb, copy the directory to a temporary directory, **unmount /sapdb** and copy them locally to /sapdb/programs/lib. Do so on every node.

An empty work directory (/usr/sap/*SID*/<Instance><Number>/work) of an SAP instance leads to a monitoring error of the SAPInstance resource agent. Every instance has to be started manually once, so the correct entries will be written to the work directory. After that you can do a manual shutdown of the instances and then the cluster is ready to control them.

SWPM (sapinst) will create a script **/etc/rc.d/sapinit** and a configuration file /usr/sap/sapservices. Both are not used by the SAPInstance Ressource Agent. Make sure /etc/rc.d/sapinit is consistent on both nodes, but not activated in a runlevel for startup of the operating system.

12.3 Check the Second Node

Make sure that no file system will be mounted on both nodes at the same time. Special caution is necessary as long as we do not have the cluster configured to protect the file systems from human errors. Before you attempt to access the SAN from node2, disconnect it from node1.

The installed system should be checked, as on the first node. All selections and parameters mentioned in the chapters beforehand have to be correct. The configuration should be similar on both nodes. Disconnect the SAN storage from node1 and check if node2 has correct SAN configuration (XFS, LVM, multipath, MD-RAID). A selfmade script could help to collect the data.

If two nodes have been set up to build an HA cluster for the current scenario, the following mount points should exist, and the following configuration files should be identical on both nodes:

- /usr/sap/*SID*/
- /sapdb/
- /sapmnt/
- /etc/multipath.conf
- /etc/multipath.bind
- /etc/mdadm.conf
- /clusterconf/*SID*/mdadm.conf
- /etc/lvm/lvm.conf
- /etc/hosts
- /etc/ntp.conf
- /etc/sysctl.conf
- /etc/hosts
- /etc/products.d/baseproduct
- /etc/services
- /etc/opt/sdb
- /etc/rc.d/sapinit
- /usr/sap/sapservices
- /sapdb/programs/lib/*
- /home/*SID*adm/.*

Beside files, other parameters have to match as well. A basic check might include the commands mentioned below. Not all of the results will be absolutely identical, but all should give a reasonable similar output.

- free
- grep MHz /proc/cpuinfo

- ip a s
- uname -a
- cat /proc/cmdline
- find /sys/devices/ -name scheduler -exec cat {} \; | grep noop
- rpm -qa | sort
- zypper lr | sort
- systemctl list-unit-files --type=service | sort
- lsmod | sort
- multipath -ll
- mount
- df -h; df -i
- grep -e *SID* adm -e sqd *SID* /etc/passwd
- ls /usr/sap/*SID*/<InstNr>/work/

For documentation of the OS configuration, the supportconfig script from the supportutils RPM could be used. The supportutils RPM is part of SUSE Linux Enterprise Server for SAP Applications 12.

13 Install the High Availability Extension software pattern

This section covers the installation of the High Availability Extension software pattern on top of the already running SUSE Linux Enterprise Server for SAP Applications 12.

The SUSE Linux Enterprise High Availability Extension Software could be installed by the following command:

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

The installation procedure should be well known to the target audience, and thus is not described in detail.

It is recommended to update the cluster management software to the most current level. This could be done by either connecting to the SCC via the Internet, or by using a locally installed proxy, f.e. SUSE Manager or SMT (https://www.suse.com/solutions/tools/smt.html). The update procedure should be well known to the target audience, and thus is not described. As of the publication date of this document, you should have at least the following releases:

- corosync-2.3.5-2.2.x86_64
- cluster-glue-1.0.12-20.2.x86_64
- libcorosync4-2.3.5-2.2.x86_64
- libglue2-1.0.12-20.2.x86_64
- libpacemaker3-1.1.13-10.4.x86_64
- ldirectord-3.9.6+git.1442374860.7f3628a-3.1.x86_64
- pacemaker-1.1.13-10.4.x86_64
- resource-agents-3.9.6+git.1442374860.7f3628a-3.1.x86_64
- sbd-1.2.1-12.1.x86_64

14 Configuration of the High Availability Extension Cluster and SAP NetWeaver integration

14.1 Basic Cluster Configuration

The first step is to setup the base cluster framework. For convenience, use YaST2 or the ha-cluster-init script. Depending on your needs, you should later add a second ring and also change to UCAST communication.

Create an initial setup, using ha-cluster-init:

```
# ha-cluster-init
  Enabling sshd.service
  Generating ssh key
  Configuring csync2
  Generating csync2 shared key (this may take a while)...done
  Enabling csync2.socket
  csync2 checking files

Configure Corosync:
  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), a multicast address and multicast port.

  Network address to bind to (e.g.: 192.168.1.0) [10.20.88.0] 192.168.124.0
  Multicast address (e.g.: 239.x.x.x) [239.249.189.242]
  Multicast port [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/...) [] /dev/disk/by-id/lun-sbdA
  All data on /dev/disk/by-id/lun-sbdA will be destroyed
  Are you sure you wish to use this device [y/N] y
  Initializing SBD......done
  Enabling hawk.service
    HA Web Konsole is now running, to see cluster status go to:
      https://10.20.88.51:7630/
    Log in with username 'hacluster', password 'linux'
WARNING: You should change the hacluster password to something more secure!
  Enabling pacemaker.service
  Waiting for cluster........done
  Loading initial configuration
  Done (log saved to /var/log/ha-cluster-bootstrap.log)


 Change the hacluster password
```

So far we have configured 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.

As requested by **`ha-cluster-init`**, we change the passwords of the user **`hacluster`**.

Infos on ha-cluster-init can be found on https://www.suse.com/documentation/sle-ha-12/book_sleha/data/sec_ha_installation_setup_auto.html

14.2 Adapting the Configuration Manually

14.2.1 Change the corosync configuration to UCAST and two rings

The section is giving you an example configuration of '/etc/corosync/corosync.conf' and is explaining the necessary changes

In the section quorum you need to check or change *transport* and you should add a second interface (ring) like shown below.

The section totem should look similar to:

```
totem {
        ...
        #HMAC/SHA1 should be used to authenticate all message
        secauth:        on
        interface {
                #Network Address to be bind for this interface setting
                bindnetaddr:    192.168.124.0
                #The multicast port to be used
                mcastport:      5405
                #The ringnumber assigned to this interface setting
                ringnumber:     0
                #Time-to-live for cluster communication packets
                ttl:    1
        }
        interface {
                #Network Address to be bind for this interface setting
                bindnetaddr:    192.168.125.0
                #The multicast port to be used
                mcastport:      5407
                #The ringnumber assigned to this interface setting
                ringnumber:     1
                #Time-to-live for cluster communication packets
```

```
                ttl:    1
        }
        ...
        #Transport protocol
        transport:      udpu
}
```

In the section nodelist you should configure two nodes, each node uses two ring addresses. The number of ring addresses depends on the available interfaces

The section nodelist should look similar to:

```
nodelist {
        node {
        #ring0 address
        ring0_addr:     192.168.124.1
        #ring1 address
        ring1_addr:     192.168.125.1
        }
        node {
        #ring0 address
        ring0_addr:     192.168.124.2
        #ring1 address
        ring1_addr:     192.168.125.2
        }
}
```

In the quorum section check the values for *expected_votes* and *two_node*.

The section quorum should look similar to:

```
quorum {
        #votequorum requires an expected_votes value to function
        expected_votes: 2
        #Enables two node cluster operations
        two_node:       1
        #Enable and configure quorum subsystem
        provider:       corosync_votequorum
```

```
}
```

The example above might need changes to work in certain environments. For example the value for ttl (Time To Live) or the token_retransmits_before_loss_const need to be choosen greater.

14.2.2 Adapt watchdog and SBD setup for real-life conditions

You can skip this section, if you do not have any sbd devices, but be sure to implement an other supported fencing mechanism. If you use the newest updates of the pacemaker packages from the SUSE maintenance channels, you can also use the -P option (*Check Pacemaker quorum and node health*) , which enables the cluster nodes not to self-fence if SBDs are lost, but pacemaker communication is still available.

Please see the sbd man page for further details of the parameters -S <n>, -P, -W.

- Activate correct watchdog
- Use two SBD devices
- Adapt SBD general settings
- Adapt SBD timings to match SAN storage settings

SBD server fencing needs an hardware watchdog to be save. In virtual machines and in lab environments, sometimes the Linux kernel's softdog is used instead of a real harware watchdog. Unfortunately, auto-probing for the correct watchdog fails sometimes. The appropriate kernel module could be defined manually. If a watchdog is activated, the respective kernel module and a device node has to show up.

```
# lsmod | grep -e wdt -e dog
softdog
# ls -l /dev/watchdog
crw-rw---- 1 root root 10, 130 Mar 08 08:08 /dev/watchdog
# grep -e wdt -e dog /etc/modules-load.d/watchdog.conf
softdog
```

If something looks different you could follow the SUSE documentation on how to activate the right watchdog (https://www.suse.com/support/kb/doc.php?id=7016880, https://www.suse.com/documentation/sle-ha-12/singlehtml/book_sleha/book_sleha.html#pro.ha.storage.protect.watchdog).

If we use two SAN storages to gain data redundancy, it makes sense to also use two SBD devices, one on each SAN storage. To set up two SBD devices, we add the other device to the variable `SBD_DEVICE` in `/etc/sysconfig/sbd`. The device paths are separated by semi-colon (;), but no trailing semicolon is used. We use short paths to device nodes in this example. In real life, always the `/dev/disk/by-id/...` has to be used.

It is required to use a watchdog together with SBD. Thererfor the `-W` option is needed. Whit current pacemaker packages, you can also use the `-P` option (*Check Pacemaker quorum and node health*), which enables the cluster nodes not to self-fence if SBD devices are lost, but pacemaker communication is still available. If `-S` is set to `1` (one), the sbd daemon will only start if the node was previously shutdown cleanly (as indicated by an exit request message in the slot), or if the slot is empty.

TABLE 14.1: SBD OPTIONS

Parameter	Description
-W	Use watchdog. It is mandatory to use a watchdog. SBD does not work reliable without watchdog. Please 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 shutdown 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"`

```
# vi /etc/sysconfig/sbd
# /etc/sysconfig/sbd
# Pls. use /dev/disk/by-id/... here. No trailing ";".
SBD_DEVICE="/dev/disk/by-id/lun-sbdA;/dev/disk/by-id/lun-sbdB"
```

```
SBD_WATCHDOG="yes"
SBD_PACEMAKER="yes"
SBD_STARTMODE="clean"
SBD_OPTS=""
```

Note

This equates to the SUSE Linux Enterprise 11 settings **SBD_OPTS="-W -P -S 1"**.

To adapt the timings, it is necessary to initialise the SBD devices. Therefor the cluster should not be running. Usually it is recommended to increase the messagewait timeout to something that is longer than usual MPIO timeouts, f.e. 130 seconds. The watchdog timeout might also be slightly increased, f.e. 30 seconds. Please align this parameters with the actual environment. The initialisation of SBD devices is done one the first cluster node only.

```
# sbd -d /dev/disk/by-id/lun-sbdA -d /dev/disk/by-id/lun-sbdB -4 130 -1 30 create
```

With the following command, check what has been written to the device:

```
# sbd -d /dev/disk/by-id/lun-sbdA -d /dev/disk/by-id/lun-sbdB dump
```

You might the also look into the manual pages stonith_sbd(7) and sbd(8) for a more complete documentation of the available options. And you might read the SUSE documentation on how to set up SBD and estimate multipathing timeouts (https://www.suse.com/support/kb/doc.php?id=7016305 , https://www.suse.com/documentation/sle-ha-12/singlehtml/book_sleha/book_sleha.html#sec.ha.storageprotection.fencing.setup).

14.3 Join the 2nd Node to the Cluster

The appropriate watchdog kernel module could be defined manually.

```
# vi /etc/modules-load.d/watchdog.conf
softdog
# modprobe softdog
```

If the watchdog is activated, the respective kernel module and a device node has to show up.

```
# lsmod | grep -e wdt -e dog
softdog
# ls -l /dev/watchdog
crw-rw---- 1 root root 10, 130 Mar 08 08:08 /dev/watchdog
# grep -e wdt -e dog /etc/modules-load.d/watchdog.conf
softdog
```

The initialisation of SBD devices is done one the first cluster node only. With the following command, check what has been written to the device:

```
# sbd -d /dev/disk/by-id/lun-sbdA -d /dev/disk/by-id/lun-sbdB dump
```

If everything looks good, the second node can be joined to cluster. Use for HOST_IP the IP of an existing cluster node

```
# ha-cluster-join -c HOST_IP
```

After the second node has joined the cluster, a set of config files should be identical on both nodes.

- /etc/hosts
- /etc/ntp.conf
- /etc/modules-load.d/watchdog.conf
- /etc/sysconfig/sbd
- /etc/corosync/corosync.conf
- /etc/corosync/authkey
- /root/.ssh/authorized_keys

14.4 Start the Cluster for the first time

Now it's time to start the cluster for the first time on both nodes:

```
# systemctl start pacemaker
```

Check on both nodes, if the cluster is running:

```
# systemctl status 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 in the following screen output. The most interesting information for now is that there are two nodes in status "online" and the message "partition with quorum".

```
============
Last updated: Fri Sep  7 18:34:49 2012
Last change: Thu Sep  6 14:48:24 2012 by <sid>adm via cibadmin on cl2n01
Stack: openais
Current DC: cl2n01 - partition with quorum
Version: 1.1.6-b988976485d15cb702c9307df55512d323831a5e
2 Nodes configured, 2 expected votes
10 Resources configured.
============


Online: [ cl2n02 cl2n01 ]


rsc_stonith_sbd     (stonith:external/sbd): Started cl2n01
```

If the SBD resource isn't available, create it

```
  primitive rsc_stonith_sbd stonith:external/sbd
    params pcmk_delay_max="15"
```

14.5 Configure Bootstrap and STONITH Using crm Shell

There are multiple ways to configure resources in a pacemaker cluster. In the following we describe the method using the crm shell. For all of the following sections to configure parts of the cluster proceed the following steps:

- Start the crm shell and enter the configure action: **`crm configure`**
- Copy and paste and all prepared crm configuration sequences into the crm shell
- Quit the crm shell (**`quit`**). You might be asked, if you want to commit your changes.

An other method is to save the crm configuration sequences into a text file adding **`configure`** at the beginning of each object (like **`configure property ...`**. The created text file could then be loaded to the cluster using **`crm -f`** *crm-command-file*

We define some cluster wide properties. The most important definitions are to activate STONITH, to define the no-quorum-policy, and to adjust the STONITH timeout values. The values here are examples only, which work in our demo cluster. These values must be tuned to your hardware and SAN environment.

```
property $id="cib-bootstrap-options" \
         stonith-enabled="true" \
         no-quorum-policy="ignore" \
         stonith-action="reboot" \
         stonith-timeout="150s"
```

For resources, we define the default stickyness of 1000. Depending on other scoring values, this should tell the cluster not to move resources without having either a loss of a node or an explicit migration request.

```
rsc_defaults $id="rsc-options" \
         resource-stickiness="1000" \
         migration-threshold="5"
```

Now we define the operation defaults. We define 10 minutes (600 seconds) as default timeout and we set the cluster to show pending actions. This is needed by the sap_suse_cluster_connector for proper cluster communication.

```
op_defaults $id="op-options" \
        timeout="600" \
        record-pending="true"
```

14.6 Configure Resources Using crm Shell

The cluster design uses the following resource agents for the workload:

IPaddr2, Raid1, LVM, Filesystem, SAPDatabase, SAPInstance.

SAP System Group “grp_sap_*SID*”

- IPaddr2: Virtual IP addresses - You should adjust the resource names and the IP adresses.

```
    primitive rsc_ip_NA2_sapna2as ocf:heartbeat:IPaddr2
            params ip="172.16.12.172"
            op monitor interval="5s" timeout="20s"
    primitive rsc_ip_NA2_sapna2ci ocf:heartbeat:IPaddr2
            params ip="172.16.12.173"
            op monitor interval="5s" timeout="20s"
    primitive rsc_ip_NA2_sapna2db ocf:heartbeat:IPaddr2
            params ip="172.16.12.171"
            op monitor interval="5s" timeout="20s"
```

- Raid1: MD Raid arrays - You should at least adjust the resource names and the parameters raidconf and raiddev.

```
    primitive rsc_md_NA2_md0 ocf:heartbeat:Raid1
            params raidconf="/clusterconf/NA2/mdadm.conf" raiddev="/dev/md0"
            op monitor interval="120s" timeout="60s" on_fail="fence"
```

- LVM: LVM volume groups - You should at least adjust the resource names and the parameter volgrpname.

```
     primitive rsc_lvm_NA2_sapvg ocf:heartbeat:LVM
             params volgrpname="sapvg"
             op monitor interval="120s" timeout="60s"
             op stop interval="0" timeout="30s" on_fail="fence"
```

- Filesystem: XFS file systems. You should at least adjust the resource names and the parameters device, directory, fstype and options.

```
     primitive rsc_fs_NA2_sapdb ocf:heartbeat:Filesystem
             params device="/dev/sapvg/sapdb" directory="/sapdb" fstype="xfs"
options="noatime,defaults"
             op stop interval="0" timeout="300s"
             op monitor interval="30s" timeout="60s"
     primitive rsc_fs_NA2_sapmnt ocf:heartbeat:Filesystem
             params device="/dev/sapvg/sapmnt" directory="/sapmnt" fstype="xfs"
options="noatime,data=writeback,defaults"
             op stop interval="0" timeout="300s"
             op monitor interval="30s" timeout="60s"
     primitive rsc_fs_NA2_usrsap ocf:heartbeat:Filesystem
             params device="/dev/sapvg/usrsap" directory="/usr/sap" fstype="xfs"
options="noatime,defaults"
             op stop interval="0" timeout="300s"
             op monitor interval="30s" timeout="60s"
```

- SAPDatabase: MaxDB database - You should at least adjust the resource names and the parameters SID and DBTYPE.

```
     primitive rsc_sapdb_NA2 ocf:heartbeat:SAPDatabase
             params SID="NA2" DBTYPE="ADA"
             op monitor interval="120s" timeout="60s" start_delay="180s"
             op start interval="0" timeout="120s"
             op stop interval="0" timeout="180s"
```

```
```

- SAPInstance ASCS01: ABAP central services instance - You should at least adjust the resource names and the parameters InstanceName, AUTOMATIC_RECOVER and START_PROFILE. The START_PROFILE must be specified with full path.

```
      primitive rsc_sapinst_NA2_ASCS01_sapna2as ocf:heartbeat:SAPInstance
              params InstanceName="NA2_ASCS01_sapna2as" AUTOMATIC_RECOVER="true" START_PROFILE="/
sapmnt/NA2/profile/START_ASCS01_sapna2as"
              op monitor interval="120s" timeout="60s" start_delay="120s"
              op start interval="0" timeout="120s"
              op stop interval="0" timeout="180s" on_fail="block"
```

- SAPInstance DVEBMGS00: central instance - You should at least adjust the resource names and the parameters InstanceName, AUTOMATIC_RECOVER and START_PROFILE. The START_PROFILE must be specified with full path.

```
      primitive rsc_sapinst_NA2_DVEBMGS00_sapna2ci ocf:heartbeat:SAPInstance
              params InstanceName="NA2_DVEBMGS00_sapna2ci" AUTOMATIC_RECOVER="true" START_PROFILE="/
sapmnt/NA2/profile/START_DVEBMGS00_sapna2ci"
              op monitor interval="120s" timeout="60s" start_delay="240s"
              op start interval="0" timeout="240s"
              op stop interval="0" timeout="240s" on_fail="block"
```

- After we have defined all cluster resources we want to combine them in a cluster resource group. You should adjust the resource names.

```
      group grp_sap_NA2 rsc_ip_NA2_sapna2as rsc_ip_NA2_sapna2db rsc_ip_NA2_sapna2ci rsc_md_NA2_md0
 rsc_lvm_NA2_sapvg rsc_fs_NA2_usrsap rsc_fs_NA2_sapmnt rsc_fs_NA2_sapdb rsc_sapdb_NA2
 rsc_sapinst_NA2_ASCS01_sapna2as rsc_sapinst_NA2_DVEBMGS00_sapna2ci
              meta target-role="Started"
```

14.7 Configure Dependencies of the Resources using crm

14.7.1 Location Constraints

Both cluster nodes have equal hardware, so we do not need a location constraint to place the SAP system group “grp_sap_*SID*.” Operational procedures and standards may require a location rule for the entire SAP system. In this case we recommend setting location rules with a score of 1000.

```
location loc_grp_sap_NA2_on_hpn03 grp_sap_NA2 1000: hpn03

location loc_grp_sap_NA2_on_hpn04 grp_sap_NA2 0: hpn04
```

We need some additional location rules to be sure that the entire SAP system starts only on the node that fits all operational needs.

- Optional: Define a location rule. This rule will move the entire SAP system to the other node if the failcount of either the databases or a SAP instance is too high. This should help to increase the availability even if one of the nodes has a defect and can't run the SAP system, but is available from the cluster view. There should be four rules, one per database and SAP instance.

- Optional: Define a location rule. This rule will move the entire SAP system to the other node if the node cannot reach at least one of the ping nodes. This should help increase availability even if the cluster node has a partial network defect and cannot reach the ping node, but is still visible to the cluster. This mechanism also includes a split-brain scenario, where only one cluster node is able to reach the ping node. We decided to use the standard gateway as the ping node. If the standard gateway is used to help SAP clients communicate with the SAP system, this also shows that the correct node (which should also be reachable from the clients) is selected in a split-brain scenario. Of course, the standard gateway is a critical resource for client/server communication and should also be highly available. On the other hand, the standard gateway should be visible only on one site in case of a split side scenario.

14.7.2 Other Constraints

Collocations

All necessary collocations are implicitly defined by adding all SAP system components to one resource group.

Order Constraints

All necessary orders are implicitly defined by adding all SAP system components to one resource group.

Implicit Dependencies

Two components (MD device and LVM volume group) of the IO stack are configured to fence a node, if the action “stop” fails. This is done by explicit setting on_fail=”fence” for the stop operation. Thus, if a cluster node fails to release those IO stack resources, the cluster will fence the node to be able to continue operation with the remaining node.

14.7.3 Start resources

So, finally the SAP database resource has to be started. This could done by using crm, with the GUI, or with the ClusterService tool.

If everything went well, all resources of the SAP system are up and running.

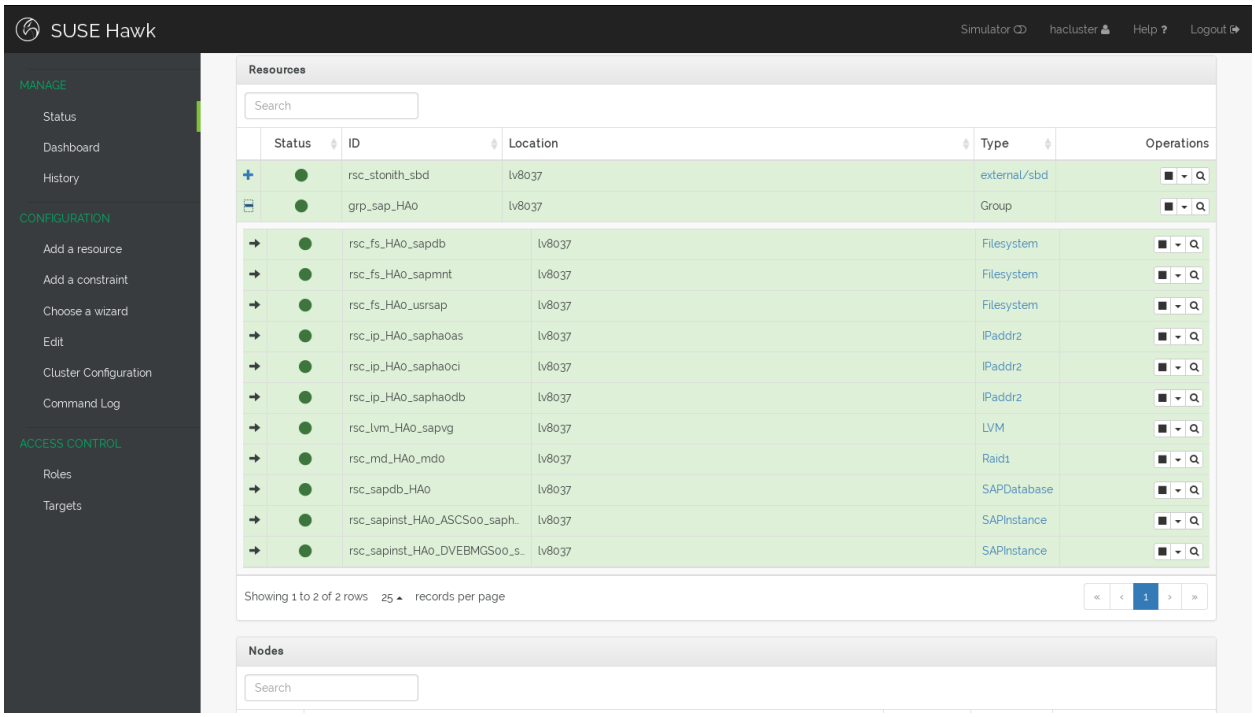

FIGURE 14.1: SAP CLUSTER UP AND RUNNING

Congratulations!

14.8 Check the SAP Integration

Use crm_mon or hawk to check, if the resources are running on one cluster node. hawk should show all resources GREEN, like above.

Check the needed default policies for stonith-enabled, default-resource-stickiness, and no-quorum-policy. F.e. use `cibadmin -Q | grep <string>` .

Use the hawk to migrate the whole SAP system to the other node. Don't forget to remove the migration constraint after the migration has finished successfully.

Login on the node that runs the SAP system and issue “systemctl stop pacemaker” to stop the cluster manager gracefully. Restart the cluster manager after all resources have been migrated to the other node.

Conduct some more tests from the test table below. Please discuss what tests are meaningful in your particular environment beforehand. For productive clusters a test plan for all possible failure scenarios is necessary.

TABLE 14.2: BASIC CLUSTER TESTS

Test Case	Expected Result
Set node with running resources to standby	All resources will start on remaining node.
Shutdown one node gracefully	The node leaves the cluster. All resources start on remaining node.
Turn off one node (power off)	The remaining node will STONITH the missing node. After the STONITH was successfull, the remaining node takes over all resources.
Turn on node again, start pacemaker	The node rejoins the cluster. Whether the resources “fall-back”, depends on the setting of the location scoring against the resource stickiness.
Plug out User LAN	The cluster does nothing. If resources are on this node, they are not reachable any more.
Plug out both SAN links	The monitoring of the SAN resources fail. The affected node gets STONITHed.
Plug out all network (split-brain)	One of the nodes will fence the other one. Exactly one node survive. If that is the one with teh active resources, the application is not affected. If the node that was runnning the resource gets STONITHed, the resources will be restarted on the surviving node.
Kill SAP CI instance	The CI instance restarts three times. After the third try, it fails over to the other node.
Kill database	The database instance restarts three times. After the third try, it fails over to the other node.
Shut down one SAN storage	The MD-Mirrors get degraded but continue to work.

III Appendix

A Software Downloads

Product	URL
SLES for SAP 12 SP1	http://cdn.microfocus.com/prot/1g-grUNNrBg~/SLE-12-SP1-SAP-DVD-x86_64-GM-DVD1.iso
SAP NetWeaver 7.5	http://service.sap.com/swdc
MaxDB	http://service.sap.com/swdc

B SUSE Products Online Documentation

Documentation for SUSE products is available at http://www.suse.com/documentation in HTML and PDF formats.

Product	URL
SLES for SAP 12	https://www.suse.com/products/sles-for-sap-12/
SLES for SAP 12	https://www.suse.com/documentation/sles-for-sap-12/
SLES 12	https://www.suse.com/documentation/sles-12/
SLE High Availability Extension 12	https://www.suse.com/documentation/sle-ha-12/
SLES for SAP 12	https://www.suse.com/releasenotes/x86_64/SLE-SAP/12-SP1/
SLES 12	https://www.suse.com/releasenotes/x86_64/SUSE-SLES/12-SP1/
SLE High Availability Extension 12	https://www.suse.com/releasenotes/x86_64/SLE-HA/12-SP1/

C SAP Notes

The general installation of SAP on Linux is described in the SAP Note 171356 - SAP software on Linux: Essential information. This SAP note also points to some SAP notes with more detailed information about hardware platforms and Linux enterprise distributions. A good entry point for installing SAP on SUSE Linux Enterprise Server 11 is SAP Note 1310037. SAP Notes are available at the SAP Service Marketplace (http://service.sap.com [http://service.sap.com/]). You need an account to access this information.

SAP Note	Title
1310037	SUSE LINUX Enterprise Server 11 Installation notes
171356	Install SAP software on Linux: Essential information
516716	Linux: Locale problems after updating glibc
1014480	SAP Management Console (SAP MC)
784391	SAP support terms and 3rd-party kernel drivers
875322	J2EE engine installation on heterogenous architectures
1275776	Linux: Preparing SLES for SAP environments
941595	Download J2SE 1.4.2 for the x64 platform
1172419	Linux: Supported Java versions on the x86_64 platform
1240081	Java Cryptography Extention Jurisdiction Policy
1164532	Release Restrictions for SAP EHP 1 for SAP NetWeaver

SAP Note	Title
864172	SAP NetWeaver 7.0 (2004s) Documentation
940420	FAQ: Database structure check (VERIFY)
785925	SAP Web AS 6.40 SR1 ABAP Installation on UNIX
790879	HA
1008828	ACC 7.1 PI/Adaptive Computing Controller
877795	Problems w/ sapstartsrv as of Release 7.00 & 6.40
995116	Backward porting of sapstartsrv for earlier releases
1122387	Linux: SAP Support in virtualized environments.
1398634	Oracle database 11g: Integration in SAP environment
129352	Homogeneous system copy with MaxDB (SAP DB)
869267	FAQ: SAP MaxDB Log area
790879	SAP Web AS 6.40 SR1 Installation on UNIX: MaxDB
936058	FAQ: SAP MaxDB Runtime Environment
936058	FAQ: SAP MaxDB Runtime Environment
873286	Unloading/loading MaxDB statistics data

SAP Note	Title
1013441	Update required: Advantages for MaxDB on 64-bit
767598	Available MaxDB documentation
820824	FAQ: SAP MaxDB/liceCache technology
1554717	SYB: Planning information for SAP on Sybase ASE.
1672366	SYB: SAP NetWeaver 7.3 EHP1 on Sybase ASE
1585981	SYB: Ensuring Recoverability for Sybase ASE
1588316	SYB: Configure automatic database and log backups
1618817	SYB: How to restore a Sybase ASE database server (UNIX)
1633491	SYB: Timestamp, BigDecimal: Problem during serializatioSYB: Timestamp, BigDecimal: Problem during serializationn

D Links to SUSE Linux Enterprise Server, SAP, Databases

SUSE Linux Enterprise Server:
http://www.novell.com/products/server/sap/matrix.html
http://www.novell.com/partner/isv/isvcatalog
http://developer.novell.com/yessearch/Search.jsp
http://support.novell.com/products/server/supported_packages
http://www.novell.com/documentation/ncc/ncc/?page=/documentation/ncc/ncc/data/bktitle.html
http://www.novell.com/coolsolutions/feature/17529.html
http://www.novell.com/partners/sap/
http://www.novell.com/docrep/2010/04/x2_Novell_SAP_Brochure_SinglePage%20Online.pdf
http://www.novell.com/docrep/2010/07/SLES_for_SAP_Applications_FAQ.pdf
http://www.novell.com/global/datacentertour/sv/pdf/Novell_SAP.pdf
http://www.novell.com/docrep/2009/09/RT_WP_Linux_Trends_in_SAP_DC_200908231_Final_English_en.doc
http://www.novell.com/docrep/2007/05/4611143_f_en.pdf
http://developer.novell.com/wiki/index.php/SAP_on_hasi_v2_Resource-SAPDatabase
http://www.novell.com/rc/index/index.jsp
http://www.novell.com/rc/docrepository/portal_skins/NovellSearch_public/SearchResults?keywords=migration&page=main&docstatus1=P&docstatus1=U&tab=1&x=0&y=0
http://www.novell.com/rc/docrepository/portal_skins/NovellSearch_public/SearchResults?keywords=sap&page=main&docstatus1=P&docstatus1=U&tab=1&x=0&y=0
http://www.novell.com/rc/docrepository/portal_skins/NovellSearch_public/SearchResults?id=NovellSearch_public&path=http%3A%2F%2Fwww.novell.com%2Frc%2Fdocrepository%2Fportal_skins%2FNovellSearch_public&page=advsearch&solution1=&solution2=&solution3=&keywords=high+availab&title=&description=&PublishDate1=&PublishDate2=&geography1=&x=0&y=0
http://www.novell.com/docrep/2009/05/SUSE-Linux-Enterprise-11_Technical-Presentation_en_en.odp

SUSE Linux Enterprise Server:
http://www.novell.com/docrep/2009/03/corebuild_playbook_v1.3_03262009_en.pdf
http://www.novell.com/docrep/2009/11/Enterprise_Linux_Servers_Solution_Presentation_f_110409_en.pdf
http://www.linux-ha.org/
http://www.openais.org/
http://www.clusterlabs.org/
https://raid.wiki.kernel.org/index.php

SAP:
https://www.sdn.sap.com/irj/sdn/nw-products
https://www.sdn.sap.com/irj/sdn/nw-70ehp1
http://www.sap.com/platform/netweaver/index.epx
http://help.sap.com/content/documentation/netweaver/docu_nw_70_design.htm#nw70ehp1
http://www.sap.com/linux
https://www.sdn.sap.com/irj/sdn/linux
https://www.sdn.sap.com/irj/scn/weblogs?blog=/pub/wlg/13603
http://service.sap.com/
https://websmp201.sap-ag.de/pam
http://sdn.sap.com/
http://service.sap.com/osdbmigration
http://www.sap.com/solutions/benchmark/index.epx
http://www.sapinsideronline.com/archive.cfm?session
http://help.sap.com/saphelp_sm32/helpdata/de/c4/3a6bff505211d189550000e829fbbd/content.htm

IBM DB2:
http://www.sap.com/about/newsroom/press.epx?pressid=4517
http://www.sap.com/community/showdetail.epx?itemID=10875
http://www.sap.com/services/servsuptech/smp

IBM DB2:

https://websmp105.sap-ag.de/~form/sapnet?_SHORTKEY=01200252310000063662

https://websmp201.sap-ag.de/pam

http://www.redbooks.ibm.com/abstracts/sg246899.html

http://publib.boulder.ibm.com/infocenter/db2luw/v9r5/index.jsp

http://www-01.ibm.com/support/docview.wss?rs=71&uid=swg27009474

http://www-01.ibm.com/software/sw-library/en_US/products/J441045L92289N69/#White%20papers

http://www.ibm.com/developerworks/wikis/display/im/SUSE+Linux+Enterprise+Server+%28SLES%29+10+-+DB2+9.5

http://www.ibm.com/developerworks/wikis/display/im/DB2+9.5+for+Linux+-+Supported+Environments

http://www.ibm.com/developerworks/data/library/techarticle/dm-0406qi/

http://en.wikipedia.org/wiki/IBM_DB2

MaxDB:

http://dev.mysql.com/downloads/maxdb/7.6.00.html

https://www.sdn.sap.com/irj/sdn/maxdb?rid=/webcontent/uuid/7001df5e-549f-2a10-4487-f818b3c52031

http://maxdb.sap.com/doc/7_6/e6/0e9640dc522f28e10000000a1550b0/content.htm

http://maxdb.sap.com/currentdoc/62/aba9a0444311d5992400508b6b8b11/frameset.htm

http://maxdb.sap.com/currentdoc/76/b0ed16f56c1d41915c70c87bf44f04/frameset.htm

http://en.wikipedia.org/wiki/MaxDB

Oracle:

http://www.novell.com/products/server/oracle/documents.html

http://www.novell.com/products/server/oracle/matrix.html

http://ftp.novell.com/partners/oracle/docs/11gR2_sles11_install.pdf

http://www.sap.com/ecosystem/customers/directories/technology/oracle/index.epx

http://www.sdn.sap.com/irj/sdn/ora

Oracle:
https://websmp201.sap-ag.de/pam
http://www.oracle.com/technology/tech/linux/install/index.html
http://www.oracle-base.com/articles/linux/ArticlesLinux.php
http://www.oracle.com/global/ro/Colaterale/ROCS_Linux/ORACLEONLINUX.pdf
http://en.wikipedia.org/wiki/Oracle_Corporation

SAP ASE:
http://www.sybase.com/products/databasemanagement/adaptiveserverenterprise
http://www.sybase.com/linux/ase
http://sybooks.sybase.com/sybooks/sybooks.xhtml
http://infocenter.sybase.com/help/index.jsp
http://www.sap.com/solutions/technology/database/adaptive-server-enterprise/index.epx
http://scn.sap.com/community/sybase-ase
http://en.wikipedia.org/wiki/Adaptive_Server_Enterprise
http://www.isug.com/Sybase_FAQ/ASE/index.html

E Sample Configuration Files for SAP Simple Stack High Availability

E.1 Sample multipath.conf

```
#/etc/multipath.conf

defaults {

        dev_loss_tmo 10

        fast_io_fail_tmo 5

        user_friendly_names yes

        bindings_file /etc/multipath.bind

}

blacklist {

        devnode "^(ram|raw|loop|fd|md|dm-|sr|scd|st)[0-9]*"

        devnode "^cciss!c[0-9]d[0-9]*"

        devnode "^dcssblk[0-9]*"

}

devices {

     device {

        vendor         "HP|COMPAQ"

        product        "HSV1[01]1 (C)COMPAQ|HSV2[01]0|HSV300|HSV4[05]0"

        path_grouping_policy    group_by_prio

        getuid_callout          "/lib/udev/scsi_id -g -u /dev/%n"

        path_checker            tur

        path_selector           "round-robin 0"

        prio                    alua

        rr_weight               uniform

        failback                manual

        hardware_handler        "0"

        no_path_retry           5

        rr_min_io_rq            1

     }

}
```

E.2 Sample corosync.conf

```
#/etc/corosync/corosync.conf

totem {

        #Used for mutual node authentication

        crypto_hash:    none

        #The mode for redundant ring. None is used when only 1 interface specified, otherwise, only active or

 passive may be choosen

        rrp_mode:       active

        #How long to wait for join messages in membership protocol. in ms

        join:   60

        #The maximum number of messages that may be sent by one processor on receipt of the token.

        max_messages:   20

        #The virtual synchrony filter type used to indentify a primary component. Change with care.

        vsftype:        none

        #HMAC/SHA1 should be used to authenticate all message

        secauth:        on

        #Used for mutual node authentication

        crypto_cipher:  none

        #This specifies the name of cluster

        cluster_name:   hacluster

        #Timeout for a token lost. in ms

        token:  5000

        #How many token retransmits should be attempted before forming a new configuration.

        token_retransmits_before_loss_const:    10

        #To make sure the auto-generated nodeid is positive

        clear_node_high_bit:    yes

        #The only valid version is 2

        version:        2

        interface {

                #Network Address to be bind for this interface setting

                bindnetaddr:    192.168.124.0

                #The multicast port to be used

                mcastport:      5405

                #The ringnumber assigned to this interface setting

                ringnumber:     0
```

```
                #Time-to-live for cluster communication packets
                ttl:    1
        }
        interface {
                #Network Address to be bind for this interface setting
                bindnetaddr:    192.168.125.0
                #The multicast port to be used
                mcastport:      5407
                #The ringnumber assigned to this interface setting
                ringnumber:     1
                #Time-to-live for cluster communication packets
                ttl:    1
        }
        #How long to wait for consensus to be achieved before starting a new round of membership
 configuration.
        consensus:      6000
        #Transport protocol
        transport:      udpu
}
nodelist {
        node {
        #ring0 address
        ring0_addr:     192.168.124.1
        #ring1 address
        ring1_addr:     192.168.125.1
        }
        node {
        #ring0 address
        ring0_addr:     192.168.124.2
        #ring1 address
        ring1_addr:     192.168.125.2
        }
}
logging {
        #Log to a specified file
        to_logfile:     no
```

```
        #Log to be saved in this specified file
        logfile:        /var/log/cluster/corosync.log
        #Log timestamp as well
        timestamp:      on
        #Facility in syslog
        syslog_facility:        daemon
        logger_subsys {
                #Enable debug for this logger.
                debug:  off
                #This specifies the subsystem identity (name) for which logging is specified
                subsys: QUORUM
        }
        #Log to syslog
        to_syslog:      yes
        #Whether or not turning on the debug information in the log
        debug:  off
        #Log to the standard error output
        to_stderr:      no
        #Logging file line in the source code as well
        fileline:       off
}
quorum {
        #votequorum requires an expected_votes value to function
        expected_votes: 2
        #Enables two node cluster operations
        two_node:       1
        #Enable and configure quorum subsystem
        provider:       corosync_votequorum
}
```

E.3 Sample CRM Configure

```
node hpn03
node hpn04
primitive rsc_fs_NA2_sapdb ocf:heartbeat:Filesystem
```

```
        params device="/dev/sapvg/sapdb" directory="/sapdb" fstype="xfs" options="noatime,defaults"
        op stop interval="0" timeout="300s"
        op monitor interval="30s" timeout="60s"
primitive rsc_fs_NA2_sapmnt ocf:heartbeat:Filesystem
        params device="/dev/sapvg/sapmnt" directory="/sapmnt" fstype="xfs"
 options="noatime,data=writeback,defaults"
        op stop interval="0" timeout="300s"
        op monitor interval="30s" timeout="60s"
primitive rsc_fs_NA2_usrsap ocf:heartbeat:Filesystem
        params device="/dev/sapvg/usrsap" directory="/usr/sap" fstype="xfs" options="noatime,defaults"
        op stop interval="0" timeout="300s"
        op monitor interval="30s" timeout="60s"
primitive rsc_ip_NA2_sapna2as ocf:heartbeat:IPaddr2
        params ip="172.16.12.172"
        op monitor interval="5s" timeout="20s"
primitive rsc_ip_NA2_sapna2ci ocf:heartbeat:IPaddr2
        params ip="172.16.12.173"
        op monitor interval="5s" timeout="20s"
primitive rsc_ip_NA2_sapna2db ocf:heartbeat:IPaddr2
        params ip="172.16.12.171"
        op monitor interval="5s" timeout="20s"
primitive rsc_lvm_NA2_sapvg ocf:heartbeat:LVM
        params volgrpname="sapvg"
        op monitor interval="120s" timeout="60s"
        op stop interval="0" timeout="30s" on_fail="fence"
primitive rsc_md_NA2_md0 ocf:heartbeat:Raid1
        params raidconf="/clusterconf/NA2/mdadm.conf" raiddev="/dev/md0"
        op monitor interval="120s" timeout="60s" on_fail="fence"
primitive rsc_sapdb_NA2 ocf:heartbeat:SAPDatabase
        params SID="NA2" DBTYPE="ADA"
        op monitor interval="120s" timeout="60s" start_delay="180s"
        op start interval="0" timeout="120s"
        op stop interval="0" timeout="180s"
primitive rsc_sapinst_NA2_ASCS01_sapna2as ocf:heartbeat:SAPInstance
        params InstanceName="NA2_ASCS01_sapna2as" AUTOMATIC_RECOVER="true" START_PROFILE="/sapmnt/NA2/
profile/START_ASCS01_sapna2as"
```

```
        op monitor interval="120s" timeout="60s" start_delay="120s"
        op start interval="0" timeout="120s"
        op stop interval="0" timeout="180s" on_fail="block"
primitive rsc_sapinst_NA2_DVEBMGS00_sapna2ci ocf:heartbeat:SAPInstance
        params InstanceName="NA2_DVEBMGS00_sapna2ci" AUTOMATIC_RECOVER="true" START_PROFILE="/sapmnt/NA2/
profile/START_DVEBMGS00_sapna2ci"
        op monitor interval="120s" timeout="60s" start_delay="240s"
        op start interval="0" timeout="240s"
        op stop interval="0" timeout="240s" on_fail="block"
primitive rsc_stonith_sbd stonith:external/sbd
 params pcmk_delay_max="15"
group grp_sap_NA2 rsc_ip_NA2_sapna2as rsc_ip_NA2_sapna2db rsc_ip_NA2_sapna2ci rsc_md_NA2_md0
 rsc_lvm_NA2_sapvg rsc_fs_NA2_usrsap rsc_fs_NA2_sapmnt rsc_fs_NA2_sapdb rsc_sapdb_NA2
 rsc_sapinst_NA2_ASCS01_sapna2as rsc_sapinst_NA2_DVEBMGS00_sapna2ci
        meta target-role="Started"
location loc_grp_sap_NA2_on_hpn03 grp_sap_NA2 1000: hpn03
location loc_grp_sap_NA2_on_hpn04 grp_sap_NA2 0: hpn04
property $id="cib-bootstrap-options"
        dc-version="1.1.2-ecb1e2ea172ba2551f0bd763e557fccde68c849b"
        cluster-infrastructure="openais"
        expected-quorum-votes="2"
        stonith-enabled="true"
        stonith-action="reboot"
 stonith-timeout="180s"
        default-action-timeout="120s"
        no-quorum-policy="ignore"
 default-resource-stickiness="1000"
        rsc_defaults $id="rsc_defaults-options"
        migration-threshold="5"
```

F Licenses

GNU Free Documentation License

Version 1.2, November 2002

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.

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.

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.

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.

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.

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”.

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.

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.

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.

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.

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

To use this License in a document you have written, include a copy of the License in the document and put the following copyright and license notices just after the title page:

```
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.

GNU General Public License

Version 2, June 1991

Copyright (C) 1989, 1991 Free Software Foundation, Inc. 59 Temple Place - Suite 330, Boston, MA 02111-1307, USA

Everyone is permitted to copy and distribute verbatim copies of this license document, but changing it is not allowed.

Preamble

The licenses for most software are designed to take away your freedom to share and change it. By contrast, the GNU General Public License is intended to guarantee your freedom to share and change free software--to make sure the software is free for all its users. This General Public License applies to most of the Free Software Foundation’s software and to any other program whose authors commit to using it. (Some other Free Software Foundation software is covered by the GNU Library General Public License instead.) You can apply it to your programs, too.

When we speak of free software, we are referring to freedom, not price. Our General Public Licenses are designed to make sure that you have the freedom to distribute copies of free software (and charge for this service if you wish), that you receive source code or can get it if you want it, that you can change the software or use pieces of it in new free programs; and that you know you can do these things.

To protect your rights, we need to make restrictions that forbid anyone to deny you these rights or to ask you to surrender the rights. These restrictions translate to certain responsibilities for you if you distribute copies of the software, or if you modify it.

For example, if you distribute copies of such a program, whether gratis or for a fee, you must give the recipients all the rights that you have. You must make sure that they, too, receive or can get the source code. And you must show them these terms so they know their rights.

We protect your rights with two steps: (1) copyright the software, and (2) offer you this license which gives you legal permission to copy, distribute and/or modify the software.

Also, for each author's protection and ours, we want to make certain that everyone understands that there is no warranty for this free software. If the software is modified by someone else and passed on, we want its recipients to know that what they have is not the original, so that any problems introduced by others will not reflect on the original authors' reputations.

Finally, any free program is threatened constantly by software patents. We wish to avoid the danger that redistributors of a free program will individually obtain patent licenses, in effect making the program proprietary. To prevent this, we have made it clear that any patent must be licensed for everyone's free use or not licensed at all.

The precise terms and conditions for copying, distribution and modification follow.

GNU GENERAL PUBLIC LICENSE TERMS AND CONDITIONS FOR COPYING, DISTRIBUTION AND MODIFICATION

0. This License applies to any program or other work which contains a notice placed by the copyright holder saying it may be distributed under the terms of this General Public License. The "Program", below, refers to any such program or work, and a "work based on the Program" means either the Program or any derivative work under copyright law: that is to say, a work containing the Program or a portion of it, either verbatim or with modifications and/or translated into another language. (Hereinafter, translation is included without limitation in the term "modification".) Each licensee is addressed as "you".

Activities other than copying, distribution and modification are not covered by this License; they are outside its scope. The act of running the Program is not restricted, and the output from the Program is covered only if its contents constitute a work based on the Program (independent of having been made by running the Program). Whether that is true depends on what the Program does.

1. You may copy and distribute verbatim copies of the Program's source code as you receive it, in any medium, provided that you conspicuously and appropriately publish on each copy an appropriate copyright notice and disclaimer of warranty; keep intact all the notices that refer to this License and to the absence of any warranty; and give any other recipients of the Program a copy of this License along with the Program.

You may charge a fee for the physical act of transferring a copy, and you may at your option offer warranty protection in exchange for a fee.

2. You may modify your copy or copies of the Program or any portion of it, thus forming a work based on the Program, and copy and distribute such modifications or work under the terms of Section 1 above, provided that you also meet all of these conditions:

a). You must cause the modified files to carry prominent notices stating that you changed the files and the date of any change.

b). You must cause any work that you distribute or publish, that in whole or in part contains or is derived from the Program or any part thereof, to be licensed as a whole at no charge to all third parties under the terms of this License.

c). If the modified program normally reads commands interactively when run, you must cause it, when started running for such interactive use in the most ordinary way, to print or display an announcement including an appropriate copyright notice and a notice that there is no warranty (or else, saying that you provide a warranty) and that users may redistribute the program under these conditions, and telling the user how to view a copy of this License. (Exception: if the Program itself is interactive but does not normally print such an announcement, your work based on the Program is not required to print an announcement.)

These requirements apply to the modified work as a whole. If identifiable sections of that work are not derived from the Program, and can be reasonably considered independent and separate works in themselves, then this License, and its terms, do not apply to those sections when you distribute them as separate works. But when you distribute the same sections as part of a whole which is a work based on the Program, the distribution of the whole must be on the terms of this License, whose permissions for other licensees extend to the entire whole, and thus to each and every part regardless of who wrote it.

Thus, it is not the intent of this section to claim rights or contest your rights to work written entirely by you; rather, the intent is to exercise the right to control the distribution of derivative or collective works based on the Program.

In addition, mere aggregation of another work not based on the Program with the Program (or with a work based on the Program) on a volume of a storage or distribution medium does not bring the other work under the scope of this License.

3. You may copy and distribute the Program (or a work based on it, under Section 2) in object code or executable form under the terms of Sections 1 and 2 above provided that you also do one of the following:

a). Accompany it with the complete corresponding machine-readable source code, which must be distributed under the terms of Sections 1 and 2 above on a medium customarily used for software interchange; or,

b). Accompany it with a written offer, valid for at least three years, to give any third party, for a charge no more than your cost of physically performing source distribution, a complete machine-readable copy of the corresponding source code, to be distributed under the terms of Sections 1 and 2 above on a medium customarily used for software interchange; or,

c). Accompany it with the information you received as to the offer to distribute corresponding source code. (This alternative is allowed only for noncommercial distribution and only if you received the program in object code or executable form with such an offer, in accord with Subsection b above.)

The source code for a work means the preferred form of the work for making modifications to it. For an executable work, complete source code means all the source code for all modules it contains, plus any associated interface definition files, plus the scripts used to control compilation and installation of the executable. However, as a special exception, the source code distributed need not include anything that is normally distributed (in either source or binary form) with the major components (compiler, kernel, and so on) of the operating system on which the executable runs, unless that component itself accompanies the executable.
If distribution of executable or object code is made by offering access to copy from a designated place, then offering equivalent access to copy the source code from the same place counts as distribution of the source code, even though third parties are not compelled to copy the source along with the object code.

4. You may not copy, modify, sublicense, or distribute the Program except as expressly provided under this License. Any attempt otherwise to copy, modify, sublicense or distribute the Program 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.

5. You are not required to accept this License, since you have not signed it. However, nothing else grants you permission to modify or distribute the Program or its derivative works. These actions are prohibited by law if you do not accept this License. Therefore, by modifying or distributing the Program (or any work based on the Program), you indicate your acceptance of this License to do so, and all its terms and conditions for copying, distributing or modifying the Program or works based on it.

6. Each time you redistribute the Program (or any work based on the Program), the recipient automatically receives a license from the original licensor to copy, distribute or modify the Program subject to these terms and conditions. You may not impose any further restrictions on the recipients' exercise of the rights granted herein. You are not responsible for enforcing compliance by third parties to this License.

7. If, as a consequence of a court judgment or allegation of patent infringement or for any other reason (not limited to patent issues), conditions are imposed on you (whether by court order, agreement or otherwise) that contradict the conditions of this License, they do not excuse you from the conditions of this License. If you cannot distribute so as to satisfy simultaneously your obligations under this License and any other pertinent obligations, then as a consequence you may not distribute the Program at all. For example, if a patent license would not permit royalty-free redistribution of the Program by all those who receive copies directly or indirectly through you, then the only way you could satisfy both it and this License would be to refrain entirely from distribution of the Program.

If any portion of this section is held invalid or unenforceable under any particular circumstance, the balance of the section is intended to apply and the section as a whole is intended to apply in other circumstances.
It is not the purpose of this section to induce you to infringe any patents or other property right claims or to contest validity of any such claims; this section has the sole purpose of protecting the integrity of the free software distribution system, which is implemented by public license practices. Many people have made generous contributions to the wide range of software distributed through that system in reliance on consistent application of that system; it is up to the author/donor to decide if he or she is willing to distribute software through any other system and a licensee cannot impose that choice.
This section is intended to make thoroughly clear what is believed to be a consequence of the rest of this License.

8. If the distribution and/or use of the Program is restricted in certain countries either by patents or by copyrighted interfaces, the original copyright holder who places the Program under this License may add an explicit geographical distribution limitation excluding those countries, so that distribution is permitted only in or among countries not thus excluded. In such case, this License incorporates the limitation as if written in the body of this License.

9. The Free Software Foundation may publish revised and/or new versions of the General Public 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.

Each version is given a distinguishing version number. If the Program specifies a version number of this License which applies to it and “any later version”, you have the option of following the terms and conditions either of that version or of any later version published by the Free Software Foundation. If the Program does not specify a version number of this License, you may choose any version ever published by the Free Software Foundation.

10. If you wish to incorporate parts of the Program into other free programs whose distribution conditions are different, write to the author to ask for permission. For software which is copyrighted by the Free Software Foundation, write to the Free Software Foundation; we sometimes make exceptions for this. Our decision will be guided by the two goals of preserving the free status of all derivatives of our free software and of promoting the sharing and reuse of software generally.

NO WARRANTY

11. BECAUSE THE PROGRAM IS LICENSED FREE OF CHARGE, THERE IS NO WARRANTY FOR THE PROGRAM, TO THE EXTENT PERMITTED BY APPLICABLE LAW. EXCEPT WHEN OTHERWISE STATED IN WRITING THE COPYRIGHT HOLDERS AND/OR OTHER PARTIES PROVIDE THE PROGRAM “AS IS” WITHOUT WARRANTY OF ANY KIND, EITHER EXPRESSED OR IMPLIED, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF MERCHANTABILITY AND FITNESS FOR A PARTICULAR PURPOSE. THE ENTIRE RISK AS TO THE QUALITY AND PERFORMANCE OF THE PROGRAM IS WITH YOU. SHOULD THE PROGRAM PROVE DEFECTIVE, YOU ASSUME THE COST OF ALL NECESSARY SERVICING, REPAIR OR CORRECTION.

12. IN NO EVENT UNLESS REQUIRED BY APPLICABLE LAW OR AGREED TO IN WRITING WILL ANY COPYRIGHT HOLDER, OR ANY OTHER PARTY WHO MAY MODIFY AND/OR REDISTRIBUTE THE PROGRAM AS PERMITTED ABOVE, BE LIABLE TO YOU FOR DAMAGES, INCLUDING ANY GENERAL, SPECIAL, INCIDENTAL OR CONSEQUENTIAL DAMAGES ARISING OUT OF THE USE OR INABILITY TO USE THE PROGRAM (INCLUDING BUT NOT LIMITED TO LOSS OF DATA OR DATA BEING RENDERED INACCURATE OR LOSSES SUSTAINED BY YOU OR THIRD PARTIES OR A FAILURE OF THE PROGRAM TO OPERATE WITH ANY OTHER PROGRAMS), EVEN IF SUCH HOLDER OR OTHER PARTY HAS BEEN ADVISED OF THE POSSIBILITY OF SUCH DAMAGES.

END OF TERMS AND CONDITIONS

How to Apply These Terms to Your New Programs

If you develop a new program, and you want it to be of the greatest possible use to the public, the best way to achieve this is to make it free software which everyone can redistribute and change under these terms.
To do so, attach the following notices to the program. It is safest to attach them to the start of each source file to most effectively convey the exclusion of warranty; and each file should have at least the “copyright” line and a pointer to where the full notice is found.

```
one line to give the program’s name and an idea of what it does.

Copyright (C) yyyy name of author
```

```
This program is free software; you can redistribute it and/or

modify it under the terms of the GNU General Public License

as published by the Free Software Foundation; either version 2

of the License, or (at your option) any later version.
```

```
This program is distributed in the hope that it will be useful,

but WITHOUT ANY WARRANTY; without even the implied warranty of

MERCHANTABILITY or FITNESS FOR A PARTICULAR PURPOSE. See the
```

```
GNU General Public License for more details.
```

```
You should have received a copy of the GNU General Public License

along with this program; if not, write to the Free Software

Foundation, Inc., 59 Temple Place - Suite 330, Boston, MA 02111-1307, USA.
```

Also add information on how to contact you by electronic and paper mail.

If the program is interactive, make it output a short notice like this when it starts in an interactive mode:

```
Gnomovision version 69, Copyright (C) year name of author

Gnomovision comes with ABSOLUTELY NO WARRANTY; for details

type `show w'. This is free software, and you are welcome

to redistribute it under certain conditions; type `show c'

for details.
```

The hypothetical commands `show w' and `show c' should show the appropriate parts of the General Public License. Of course, the commands you use may be called something other than `show w' and `show c'; they could even be mouse-clicks or menu items--whatever suits your program.

You should also get your employer (if you work as a programmer) or your school, if any, to sign a “copyright disclaimer” for the program, if necessary. Here is a sample; alter the names:

```
Yoyodyne, Inc., hereby disclaims all copyright

interest in the program `Gnomovision'

(which makes passes at compilers) written

by James Hacker.
```

```
signature of Ty Coon, 1 April 1989

Ty Coon, President of Vice
```

This General Public License does not permit incorporating your program into proprietary programs. If your program is a subroutine library, you may consider it more useful to permit linking proprietary applications with the library. If this is what you want to do, use the GNU Lesser General Public License [http://www.fsf.org/licenses/lgpl.html] instead of this License.

Terminology

active/active, active/passive

A concept of how services are running on nodes. An active-passive scenario means that one or more services are running on the active node and the passive node waits for the active node to fail. Active-active means that each node is active and passive at the same time.

cluster

A high-performance cluster is a group of computers (real or virtual) sharing the application load in order to achieve faster results. A high-availability cluster is designed primarily to secure the highest possible availability of services.

cluster information base (CIB)

A representation of the whole cluster configuration and status (node membership, resources, constraints, etc.) written in XML and residing in memory. A master CIB is kept and maintained on the *designated coordinator (DC)* and replicated to the other nodes.

cluster partition

Whenever communication fails between one or more nodes and the rest of the cluster, a cluster partition occurs. The nodes of a cluster are split in partitions but are still alive. They can only communicate with nodes in the same partition and are unaware of the separated nodes. If the loss of the nodes on the other partition cannot be confirmed, a split brain scenario develops (see also *split brain*).

cluster resource manager (CRM)

The main management entity responsible for coordinating all non-local interactions. Each node of the cluster has its own CRM, but the one running on the DC is the one elected to relay decisions to the other non-local CRMs and process their input. A CRM interacts with a number of components: local resource managers, both on its own node and on the other nodes, non-local CRMs, administrative commands, the fencing functionality, and the membership layer.

consensus cluster membership (CCM)

The CCM determines which nodes make up the cluster and shares this information across the cluster. Any new addition and any loss of nodes or quorum is delivered by the CCM. A CCM module runs on each node of the cluster.

designated coordinator (DC)

The “master” node. This node is where the master copy of the CIB is kept. All other nodes get their configuration and resource allocation information from the current DC. The DC is elected from all nodes in the cluster after a membership change.

distributed lock manager (DLM)

DLM coordinates disk access for clustered file systems and administers file locking to increase performance and availability.

distributed replicated block device (DRBD)

DRBD is a block device designed for building high availability clusters. The whole block device is mirrored via a dedicated network and is seen as a network RAID-1.

failover

Occurs when a resource or node fails on one machine and the affected resources are started on another node.

fencing

Describes the concept of preventing access to a shared resource by isolated or failing cluster members. Should a cluster node fail, it will be shut down or reset to prevent it from causing trouble. This way, resources are locked out of a node whose status is uncertain.

Heartbeat resource agent

Heartbeat resource agents were widely used with Heartbeat version 1. Their use is deprecated, but still supported in version 2. A Heartbeat resource agent can perform `start`, `stop`, and `status` operations and resides under `/etc/ha.d/resource.d` or `/etc/init.d`. For more information about Heartbeat resource agents, refer to http://www.linux-ha.org/HeartbeatResourceAgent (see also *OCF resource agent*).

high availability

High availability is a system design approach and associated service implementation that ensures a prearranged level of operational performance will be met during a contractual measurement period.

Availability is a key aspect of service quality. Availability is usually calculated based on a model involving the Availability Ratio and techniques such as Fault Tree Analysis.

See also: http://en.wikipedia.org/wiki/High_availability/ http://www.itlibrary.org/index.php?page=Availability_Management

local resource manager (LRM)

The local resource manager (LRM) is responsible for performing operations on resources. It uses the resource agent scripts to carry out these operations. The LRM is "dumb" in that it does not know of any policy. It needs the DC to tell it what to do.

Linux Standard Base (LSB)

The goal of the Linux Standard Base is to develop and promote a set of standards that will increase compatibility among Linux distributions.

LSB resource agent

LSB resource agents are standard LSB init scripts. LSB init scripts are not limited to use in a high availability context. Any LSB-compliant Linux system uses LSB init scripts to control services. Any LSB resource agent supports the options `start`, `stop`, `restart`, `status` and `force-reload` and may optionally provide `try-restart` and `reload` as well. LSB resource agents are located in `/etc/init.d`. Find more information about LSB resource agents and the actual specification at http://www.linux-ha.org/LSBResourceAgent and http://www.linux-foundation.org/spec/refspecs/LSB_3.0.0/LSB-Core-generic/LSB-Core-generic/iniscrptact.html (see also *OCF resource agent* and *Heartbeat resource agent*).

node

Any computer (real or virtual) that is a member of a cluster and invisible to the user.

policy engine (PE)

The policy engine computes the actions that need to be taken to implement policy changes in the CIB. This information is then passed on to the transaction engine, which in turn implements the policy changes in the cluster setup. The PE always runs on the DC.

OCF resource agent

OCF resource agents are similar to LSB resource agents (init scripts). Any OCF resource agent must support `start`, `stop`, and `status` (sometimes called `monitor`) options. Additionally, they support a `metadata` option that returns the description of the resource agent type in XML. Additional options may be supported, but are not mandatory. OCF resource agents reside in `/usr/lib/ocf/resource.d/<provider>`. Find more information about OCF resource agents and a draft of the specification at http://www.linux-ha.org/OCFResourceAgent and http://www.opencf.org/cgi-bin/viewcvs.cgi/specs/ra/resource-agent-api.txt?rev=HEAD (see also *Heartbeat resource agent*).

quorum

In a cluster, a cluster partition is defined to have quorum (is “quorate”) if it has the majority of nodes (or votes). Quorum distinguishes exactly one partition. It is part of the algorithm to prevent several disconnected partitions or nodes from proceeding and causing data and service corruption (split brain). Quorum is a prerequisite for fencing, which then ensures that quorum is indeed unique.

resource

Any type of service or application that is known to Heartbeat. Examples include an IP address, a file system, or a database.

resource agent (RA)

A resource agent (RA) is a script acting as a proxy to manage a resource. There are three different kinds of resource agents: OCF (Open Cluster Framework) resource agents, LSB resource agents (Standard LSB init scripts), and Heartbeat resource agents (Heartbeat v1 resources).

Single Point of Failure (SPOF)

A single point of failure (SPOF) is any component of a cluster that, should it fail, triggers the failure of the entire cluster.

split brain

A scenario in which the cluster nodes are divided into two or more groups that do not know of each other (either through a software or hardware failure). STONITH prevents a split brain situation from badly affecting the entire cluster. Also known as a “partitioned cluster” scenario.

The term split brain is also used in DRBD but means that the two nodes contain different data.

STONITH

The acronym for “Shoot the other node in the head”, which refers to the fencing mechanism that shuts down a misbehaving node to prevent it from causing trouble in a cluster.

STONITH Block Device (SBD)

Implementation of STONITH, using a shared SAN LUN for issuing shutdown instruction to certain nodes. This implemnentation always needs an a watchdog running on each node to be save.

transition engine (TE)

The transition engine (TE) receives policy directives from the PE and carries them out. The TE always runs on the DC. From there, it instructs the local resource managers on the others nodes which actions to take.

watchdog

A watchdog timer, or simply a watchdog, is an electronic timer that is used to detect and recover from computer malfunctions. During normal operation, the cluster regularly restarts the watchdog timer to prevent it from elapsing, or "timing out". If, due to a hardware fault or program error, the cluster fails to restart the watchdog, the timer will shutdown the computer.