SAP Certified Cluster On SUSE® Linux Enterprise Server For SAP Applications

Architecture

Best Practices

Components

Fabian Herschel, SUSE
SAP Architect LinuxLab
Fabian.herschel@suse.com

Markus Gürtler, SUSE
Architect & Technical ISV Manager
mguertler@suse.com
Agenda

- Overview: SAP and High Availability
- SAP “Enqueue Replication” Scenario
- SAP Certified Cluster Architecture
- SAP in Virtualized Environments
- Outlook: SAP HANA in the cluster
Overview:
SAP and High Availability
What is SAP NetWeaver?

Application server platform for SAP Business Suite

Three application server flavors available:

**ABAP**: SAP's business programming language

**JAVA**: JAVA based SAP applications

**Dual-stack**: Combines ABAP and JAVA stack

Users use clients (web browser or SAP GUI) to connect to the SAP application servers
SAP Application Server Overview

Application Server ABAP
- ERP
- SCM
- PI / XI
- CRM

Application Server JAVA
- EP
- BI
- SolMan
- BW

SAP GUI
- Web / http(s)

Web / http(s)

Oracle
MaxDB

Sybase
DB2
Why High Availability for SAP?

Minimized downtimes
- in case of hardware and software failures
- in case of disasters
- for administrative tasks like patching of SAP and OS

High requirements on SLAs

SAP relies on 3rd party vendors for High Availability

<table>
<thead>
<tr>
<th>Availability</th>
<th>Outage times</th>
</tr>
</thead>
<tbody>
<tr>
<td>90%</td>
<td>5 weeks</td>
</tr>
<tr>
<td>98%</td>
<td>1 week</td>
</tr>
<tr>
<td>99%</td>
<td>3.7 days</td>
</tr>
<tr>
<td>99.8%</td>
<td>18 hours</td>
</tr>
<tr>
<td>99.9%</td>
<td>9 hours</td>
</tr>
<tr>
<td>99.99%</td>
<td>53 minutes</td>
</tr>
<tr>
<td>99.999%</td>
<td>5 minutes</td>
</tr>
</tbody>
</table>

99.99% and 99.999% requires an HA cluster
What Does SUSE Offer?

- **SAP certified solution**
  Industry proven and SAP certified High Availability solution

- **“Best Practice” guides**
  “Best Practice” guides describe detailed the most relevant SAP High Availability scenarios

- **Partner Ecosystem for HA integration**
  Strong partner ecosystem for HA integration projects

- **Continually improved**
  Our SAP HA solutions are continually improved and extended by new use-cases in a strong collaboration with SAP
High Availability Training

SAP offers a training course for SUSE HA

• Duration: 3 days
• Content: SLES for SAP + High Availability
• More details at http://training.sap.com
High Availability Software Components Shipped With SLES For SAP

✅ SUSE Linux Enterprise High Availability Extension

The most modern and complete open source solution for implementing highly available Linux clusters

✅ SAPInstance resource agent

Supports SAP NetWeaver stack with kernel >= 6.40

✅ SAPDatabase resource agent

Supports Sybase, Oracle, MaxDB, DB2
SAP HA Basis Architecture

- No single point of failure
- Split-brain avoidance
- Full redundant and high available
- Active / Active
- NFS for shared SAP directories
- Fast SAN for SAP database
Cluster Failover

Left node failed
Right node takes over
Split Brain

Split brain detection using fencing
SBD devices - Storage-based death
HA Stack for SAP

Oracle  SAP  IP  FS
SAPDatabase  SAPInstance  IPAddr2  SBD  MD  LVM  FS

pacemaker/OpenAIS

Network  Ext3  LVM
Network Bonding  Multipath  MD-Raid

Linux system / Kernel

Application Layer
Resource Agent Layer
Cluster Layer
I/O Layer
System Layer
High Availability “Best Practice” Guides

Jointly developed by SUSE, customers and partners
Addressing “real world” use-cases
High Availability – Use Cases

Simple Stack
- cluster node 1
  - SAP group 1
    - database
    - SAP instances
  - resource failover
    - active / active
  - shared storage
- cluster node 2
  - SAP group 2
    - database
    - SAP instances

Enqueue Replication
- cluster node 1
  - SAP system
    - database
    - SAP (A)SCS
  - resource failover
    - active / active
  - lock table replication
- cluster node 2
  - SAP system
    - SAP Enq-Repl
  - shared storage

DRBD Data Sync
- cluster 1
  - DRBD master
    - SAP system
  - DRBD data replication
    - active / active
  - Data Center 1
- cluster 2
  - DRBD slave
  - Data Center 2

HA in Virtual Environments
- Cluster nodes on virtual systems
  - cluster B
  - cluster C
- Clustered Hyper Visor on physical systems
  - cluster A
SAP Enqueue Replication Scenario
SAP Enqueue-Replication
How does it work?

- DB transaction
- SAP transaction
- create lock
- SAP enqueue lock table in shared memory
- ENQ process
- cluster node 1
- replicate ENQ table
- ERS process
- cluster node 2
- release lock
- SAP enqueue lock table in shared memory
Enqueue Replication – Use Case

node 1

SAP system - SAP (A)SCS

node 2

SAP system - database - SAP Enq-Repl

shared storage

resource failover

active / active

node 1

node 2
Enqueue Replication - Take Over

The cluster registers that Node1 is failed.

With help of a STONITH action we ensure that Node1 is really off.

Node2 takes all resources, and the Enq.-Repl. Instance get promoted to the (A)SCS instance.

The locking table within the shared memory gets transferred and released.
Enqueue Replication - Node Re-Join

Node 1 comes back and joins the running cluster

The cluster starts the ERS resource on this “new” node

The SAP Enqueue – Replication synchronizes all lock entries
Bind IP addresses to M/S status but do not stop, if instance is down
Cluster Concept - Constraints
Order Constraints for SAP Instances and Groups

Each SAP application server needs the central services (ASCS)
Cluster Concept - Constraints
Order Constraints for SAP Instances and SAP Database

Advisory order constraints for all SAP application servers against Database and central services instance (ASCS)

Optional advisory order constraint between Database and central services instance (ASCS)
# Failures and Solutions

Some examples of “What happens when”

<table>
<thead>
<tr>
<th>Failure Description</th>
<th>Solution Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Server (blade) fails</td>
<td>Other node takes over</td>
</tr>
<tr>
<td>Network switch fails</td>
<td>2nd network interface takes over (network bonding)</td>
</tr>
<tr>
<td>Enqueue server dies</td>
<td>ASCS instance will fail over</td>
</tr>
<tr>
<td>Message server dies</td>
<td>Message server will be locally restarted</td>
</tr>
<tr>
<td>Network separation (split-brain)</td>
<td>SBD detects failure and fences (reboots) one node</td>
</tr>
<tr>
<td>Storage fails</td>
<td>RAID mirror leg gets lost 2nd mirror leg continues</td>
</tr>
<tr>
<td>Database fails</td>
<td>Database will be restarted and/or started on 2nd node</td>
</tr>
</tbody>
</table>
The Dos and Don'ts
Things you should consider

- Keep cluster configuration simple
- Use SBD for node fencing (STONITH)
- Define and perform tests for all failure scenarios
- Follow our best practices
The Dos and Don'ts
Things you should avoid

- Build Cluster cluster without node fencing (STONITH)
- Go live without tests planned and done
- Go live without proper operations manual
- Cluster resource (like SBD and STONITH) timings shorter than SAN timings
SAP Certified Cluster Architecture
SAP NetWeaver High Availability Cluster 7.30 Certification

Start/Stop infrastructure

certifies the **Start/Stop infrastructure** within HA-Setups (SAP NetWeaver Management Agents, Cluster-API).

Reference SAP HA setup scenario

Defines a **reference SAP HA setup scenario** that must be used by all HA vendors

HA setups for SAP NetWeaver 7.x

Unifies **HA setups for SAP NetWeaver 7.x**, based on 7.20 DCK availability for ABAP as well as for Java Application Servers.

http://scn.sap.com/docs/DOC-26718
SAP System in a SAP recommended HA Setup

http://scn.sap.com/docs/DOC-25453
How to Connect SAPSTARSRV and Cluster Frameworks using the Components saphascriptco.so and SAP_Vendor_Cluster_Connector

Applies to:
All SAP products controlled by sapstartsrv. For more information, visit the Application Management homepage.

Summary
Running SAP in a high availability cluster environment needs additional interfaces to communicate between the SAP program SAPSTARSERV and the high availability cluster. At least it is essential to inform the cluster, if a SAP instance is started or shutdown. This article describes the reference interface implementation sap_suse_cluster_connector.

Author: Fabian Herschel
Company: SUSE Linux Products GmbH
Created on: April 3, 2012
Version: 1.0

Author Bio
Fabian Herschel is SAP architect and employee of SUSE Linux Products GmbH in Nuremberg, Germany. He is currently working directly at the SAP LinuxLab in Rot and is leading, together with his colleague Markus Gürtler, the creation of the architectures of HA solutions for SAP products running on SUSE Linux Enterprise Server and the SUSE Linux Enterprise High Availability Extention.

http://scn.sap.com/docs/DOC-28875
SAP HA Cluster Interface
Interfaces to integrate our HA solution in SAP

http://scn.sap.com/docs/DOC-25453
http://www.sap.com/partners/directories/searchpartner.epx
→ Search for Company Name SUSE
The Certificate

- SLES HA 11 on architectures x86_64 and power64
- Certification valid 3 years till Sep 2015 / Aug 2016
- For SAP NetWeaver 7.30
- Certifies the SAP SUSE cluster integration

So we are...

SAP® Certified Integration with SAP NetWeaver®
<table>
<thead>
<tr>
<th>Master/Slave Set: rsc_sap_HA0_ASCS00_0</th>
<th>Slave</th>
</tr>
</thead>
<tbody>
<tr>
<td>cl2n01: Online</td>
<td>cl2n02: Online</td>
</tr>
<tr>
<td>rsc_sap_HA0_ASCS00_0: Master</td>
<td>rsc_sap_HA0_ASCS00_1: Slave</td>
</tr>
<tr>
<td>stonith-sbd: Started</td>
<td></td>
</tr>
<tr>
<td>rsc_ip_HA0_sapha0as: Started</td>
<td></td>
</tr>
<tr>
<td>rsc_ip_HA0_sapha0cr: Started</td>
<td></td>
</tr>
<tr>
<td>rsc_fs_HA0_dvebmgs01: Started</td>
<td></td>
</tr>
<tr>
<td>rsc_sap_HA0_DVEBMGS01: Started</td>
<td></td>
</tr>
<tr>
<td>rsc_ip_HA0_sapha0d2: Started</td>
<td></td>
</tr>
<tr>
<td>rsc_fs_HA0_d02: Started</td>
<td></td>
</tr>
<tr>
<td>rsc_sap_HA0_D02: Started</td>
<td></td>
</tr>
</tbody>
</table>
Sample Configuration

Primitive: rsc_sapinst_HA0_ASCS00_sapha0as

<table>
<thead>
<tr>
<th>Resource Agent</th>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>InstanceName</td>
<td>HA0_ASCS00_sapha0as</td>
</tr>
<tr>
<td></td>
<td>ERS_InstanceName</td>
<td>HA0_ERS10_sapha0as</td>
</tr>
<tr>
<td></td>
<td>AUTOMATIC_RECOVER</td>
<td>true</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Type</th>
<th>Interval</th>
<th>Timeout</th>
<th>Start delay</th>
<th>Role</th>
<th>On fail</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>monitor</td>
<td>120s</td>
<td>60s</td>
<td>120s</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>monitor</td>
<td>121s</td>
<td>60s</td>
<td>120s</td>
<td>slave</td>
<td></td>
</tr>
<tr>
<td></td>
<td>monitor</td>
<td>119s</td>
<td>60s</td>
<td>120s</td>
<td>master</td>
<td></td>
</tr>
<tr>
<td></td>
<td>start</td>
<td>0</td>
<td>180s</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>stop</td>
<td>0</td>
<td>240s</td>
<td></td>
<td></td>
<td>block</td>
</tr>
</tbody>
</table>
SAP in Virtualized Environments
Use Case: Cluster in Cluster

Cluster Nodes on Virtual Systems

Cluster A

Cluster Hyper Visor on Physical Systems

node1 cluster B
node2 cluster B
node1 cluster C
node2 cluster C
node3 cluster C
VMware HA and SLE HAE

SUSE Linux Enterprise
High Availability Extension

* Both SLE HA Nodes running on ESX server 1
* ESX Server 3 is powered down

VMware HA and DRS Cluster

VMware ESX

(VMware ESX)
* VM is migrated to ESX server 2 without
  * SLE HA cluster interference
VMware HA and SLE HAE

SUSE Linux Enterprise High Availability Extension

* SLE HA cluster now runs on different ESX servers to have HA against Hardware failures...
VMware HA and SLE HAE

* SLE HA cluster now runs on different ESX servers to have HA against Hardware failures...
VMware HA and SLE HA

* ... This was just in time, because unfortunately a ESX hardware system fails
* SLE HA migrates the Database and optionally shutdown an Application Server

VMware HA and DRS Cluster

SUSE Linux Enterprise
High Availability Extension
VMware HA and SLE HA

- SUSE Linux Enterprise
  High Availability Extension

* ESX server 1 is now in hardware Maint.
* VMware DPM powers up ESX server 3
* Failed Virtual Machines get started by VMware HA

VMware HA and DRS Cluster

- SCS
- APP OS
- APP OS
- APP OS
- OS

(VMware ESX)

VMware ESX

VMware ESX

<table>
<thead>
<tr>
<th>SCS</th>
<th>APP OS</th>
<th>APP OS</th>
<th>APP OS</th>
<th>OS</th>
</tr>
</thead>
</table>

ZZ

43
VMware HA and SLE HA

SUSE Linux Enterprise
High Availability Extension

* One of the virtual machines with an SAP application server is migrated to ESX server 3
* SLE HA starts the SAP application Server on the cluster node

VMware HA and DRS Cluster

(VMware ESX)  VMware ESX  VMware ESX

SCS  APP  OS  APP  OS  APP  OS

vMotion

APP  OS  APP  OS
VMware HA and SLE HA

SUSE Linux Enterprise
High Availability Extension

* Migration is ready with complete business continuity

VMware HA and DRS Cluster

(VMware ESX)

VMware ESX

VMware ESX
Outlook: SAP HANA in the Cluster
Outlook: HANA in a Cluster

HANA Single Box – Cold Takeover

Productive HANA DB on node1

Take over takes quite long, because data need to be loaded into memory

Optional QA HANA DB on second node2

node 1

SAP HANA (PR1)

resource failover

active / active

node 2

SAP HANA (QA1)

shared storage

SBD
PR1
QA1
Outlook: HANA in a Cluster
HANA “Single Box – System Replication”

This cluster concept is still under scoping and research.
Outlook: HANA in a Cluster
HANA “Multi Node – System Replication”

This cluster concept is still under scoping and research.
Unpublished Work of SUSE. All Rights Reserved.

This work is an unpublished work and contains confidential, proprietary and trade secret information of SUSE. Access to this work is restricted to SUSE employees who have a need to know to perform tasks within the scope of their assignments. No part of this work may be practiced, performed, copied, distributed, revised, modified, translated, abridged, condensed, expanded, collected, or adapted without the prior written consent of SUSE. Any use or exploitation of this work without authorization could subject the perpetrator to criminal and civil liability.

General Disclaimer

This document is not to be construed as a promise by any participating company to develop, deliver, or market a product. It is not a commitment to deliver any material, code, or functionality, and should not be relied upon in making purchasing decisions. SUSE makes no representations or warranties with respect to the contents of this document, and specifically disclaims any express or implied warranties of merchantability or fitness for any particular purpose. The development, release, and timing of features or functionality described for SUSE products remains at the sole discretion of SUSE. Further, SUSE reserves the right to revise this document and to make changes to its content, at any time, without obligation to notify any person or entity of such revisions or changes. All SUSE marks referenced in this presentation are trademarks or registered trademarks of Novell, Inc. in the United States and other countries. All third-party trademarks are the property of their respective owners.