Towards Zero Downtime
How to Maintain SAP HANA System Replication Clusters

Fabian Herschel
Senior Architect SAP
Fabian.Herschel@suse.com

Markus Gürtler
Senior Architect SAP
Markus.Guertler@suse.com
Agenda

SUSE Linux Enterprise Server for SAP Applications

Business Continuity with SLES for SAP Applications

SAP HANA System Replication Automation Scenarios

Maintenance for SAP HANA System Replication Clusters
SUSE + SAP

Unrivaled Relationship Making SUSE the Smart Choice for SAP Workloads

• 17+ years of joint testing and development at the SAP LinuxLab
• Joint collaboration on Cloud Foundry
• SUSE Linux Enterprise is the leading platform for SAP workloads on Linux
• Seamless support from SAP and SUSE
• SUSE Linux Server for SAP Applications delivers built-in high availability, superior performance and security
• First and leading OS for SAP HANA
• The platform powering SAP HANA Enterprise Cloud
• SUSE OpenStack Cloud powers SAP’s HANA Cloud platform
SUSE Linux Enterprise Server for SAP Applications 12 SP1

- 24x7 Priority Support for SAP
- Extended Service Pack Support
  18 Month Grace Period
- Page Cache Management
- Installation Wizard
- SAP HANA Firewall
- SAP HANA Resource Agents
- SUSE Linux Enterprise Server
- SLE High Availability
  SAP HANA & SAP NetWeaver
- SAP specific update channel

SAP specific update channel
Lifecycle Model / Extended Service Pack Support

13-year lifecycle (10 years general support, 3 years extended support)

Up to 5-year lifecycle per Service Pack (3 years general + 2 years extended support)

18 month migration period between two service packs

6 month window to support “skip service pack” functionality (e.g. SPn to SPn+2)

Long Term Service Pack Support (LTSS) available on top (x86-64 only)

More information available on: http://www.suse.com/lifecycle
Full System Rollback with One Click

Reduce downtime from service pack update errors
SUSE High Availability Solution for SAP HANA

nodeA

SAP HANA Primary

vIP

Cluster Communication

nodeB

SAP HANA Secondary

SAP Hana
Master/Slave Resource

Slave

Clone

SAP HanaTopology
Clone Resource

Fencing
Four Steps to Install and Configure

01101  Install SAP HANA

Configure SAP HANA System Replication

Install and initialize SUSE Cluster

Configure SR Automation using HAWK wizard
SAPHanaSR HAWK Wizard
What is the Delivery?
SUSE Linux Enterprise Server for SAP Applications

The package **SAPHanaSR**
- the two resource agents
  - *SAPHanaTopology*
  - *SAPHana*
- HAWK setup Wizard

The package **SAPHanaSR-doc**
- the important *SetupGuide*
SAPHanaSR Scale-Up Scenarios
### SAP HANA Scale-Up: Performance Optimized

<table>
<thead>
<tr>
<th><strong>Node 2 Usage:</strong></th>
<th>Dedicated</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Data pre-load on Secondary:</strong></td>
<td>Yes</td>
</tr>
<tr>
<td><strong>Take-over decision:</strong></td>
<td>Fully automated by SUSE cluster solution</td>
</tr>
<tr>
<td><strong>Take-over process:</strong></td>
<td>Fully automated by SUSE cluster solution</td>
</tr>
<tr>
<td><strong>Take-over reaction time:</strong></td>
<td>Fast due to pacemaker heartbeat</td>
</tr>
<tr>
<td><strong>Take-over speed:</strong></td>
<td>Fast since data pre-loaded</td>
</tr>
</tbody>
</table>

![Diagram showing SAP HANA System Replication](image)
<table>
<thead>
<tr>
<th><strong>Node 2 Usage:</strong></th>
<th>Shared with other system (e.g. QA1). Additional storage required.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Data pre-load on Secondary:</strong></td>
<td>No</td>
</tr>
<tr>
<td><strong>Take-over decision:</strong></td>
<td>Fully automated by SUSE cluster solution</td>
</tr>
<tr>
<td><strong>Take-over process:</strong></td>
<td>Fully automated by SUSE cluster solution</td>
</tr>
<tr>
<td><strong>Take-over reaction time:</strong></td>
<td>Fast due to pacemaker heartbeat</td>
</tr>
<tr>
<td><strong>Take-over speed:</strong></td>
<td>Slow: stop QA1 (meaning QA1 downtime) + completely load PR1 into memory</td>
</tr>
</tbody>
</table>

![Diagram of SAP HANA System Replication](image)

A => B, Q
SAP HANA Multitenant Database Containers (MDC)

MDC Considerations:

• Can apply “Performance Optimized” or “Cost Optimized” scenarios

• A take-over acts on the parent HANA Database.

• All tenant database containers and associated services and therefore affected by a take-over.

• For new installations with SAP HANA rev > 120 MDC is the default and any installation results into a system and a data tenant.

%A => %B
SAP HANA Scale-Up: Multi Tier

Node 2 Usage: Dedicated

Data pre-load on Secondary: Yes

Take-over decision: Fully automated by SUSE cluster solution

Take-over process: Fully automated by SUSE cluster solution

Take-over reaction time: Fast due to pacemaker heartbeat

Take-over speed: Fast since data pre-loaded

A => B → C
SAPHanaSR Scale-Out Scenario
SAP HANA PR1 – site WDF

SLES for SAP Applications - pacemaker cluster

Majority maker

SR sync

SAP HANA PR1 – site ROT
A SAP HANA scale-out database consists of multiple nodes and SAP HANA instances.

Each worker node \textbf{W} has its own data partition.

Standby nodes \textbf{S} do not have a data partition.
A SAP HANA scale-out database consists of several services such as master name server \( M \).

The active master name server takes all client connections and redirects the client to the proper worker node. It always has data partition 1.

Master candidates \( (M) \) could be worker or standby nodes. Typically there are 3 nodes which could get active master name server.
SAP HANA Scale-Out System replication

SAP HANA PR1 – site WDF

SR channels per service

Majority maker

 Overall status SOK or SFAIL

client

VIP

NodeA1 NodeA2 NodeA3 NodeA4 NodeA5...

NodeB1 NodeB2 NodeB3 NodeB4 NodeB5...

1 2 3 ... N

1 2 3 ... N

primary

secondary

SAP HANA PR1 – site ROT

vIP

1 2 3 N

1 2 3 N

client
A lot of different failures must be detected and processed by SAP Hana SR-Scale-Out:
- outage of the majority maker
- outage of single or multiple nodes and instances
- outage of a complete SAP HANA SR site (primary or secondary)
- outage and recovery of system replication channels
- the vIP must “follow” the master name server of the primary replication site
## SAPHanaSR Scale-Out Conducting
### Typical Failures and Reactions

<table>
<thead>
<tr>
<th>Failure</th>
<th>SAPHanaSR</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Worker</strong> fails - node or instance</td>
<td>SAP HA processes failover. If SAP HA fails, SAPHanaSR processes a takeover or restart.</td>
</tr>
<tr>
<td><strong>Active master name server</strong> fails - node or instance</td>
<td>Like the worker failure. In addition SAPHanaSR migrates the virtual IP address to the new active master name server.</td>
</tr>
<tr>
<td><strong>Standby</strong> fails - node or instance</td>
<td>SAPHanaSR processes an instance restart to re-establish the full SAP HA capacity.</td>
</tr>
<tr>
<td><strong>Primary site</strong> fails</td>
<td>SAPHanaSR processes a takeover on secondary or restart of the failed primary depending on configuration and system replication status.</td>
</tr>
<tr>
<td><strong>Standby site</strong> fails</td>
<td>SAPHanaSR processes a database system restart to re-establish SAP HANA system replication.</td>
</tr>
</tbody>
</table>
Let us start with the Maintenance
A problem has been detected and your system has been shutdown to prevent damage of your computer.

**DRIVER_ERR_NEITHER_DIFFERENT_NOR_EQUAL**

If this is the first time you have seen this blue screen, restart your computer using key F13. If this screen appears again, follow these steps:
* Check to make sure any new hardware or software is properly installed.
* If this is a new installation, ask your software manufacturer for any updates you might need.
* Feel free to re-install the current OS as often as you like or have time to do that.

If problems continue, disable the current OS. We *strongly recommend* to switch to SUSE(R) Linux Enterprise for SAP Applications 12 SP2.

**Technical information:**
*** STOP: 0x00008A8A (0x00000003,0x00000002,0x00000001,0x00000000)***
*** goodby3.sys - Address 000B1E00 base at 000B1E00, DateStamp DEEEDEEE***

To continue or un-lock this session please shout "SUSE".
Wasn't that session about Towards zero down time?
About Maintenance

Why do I need special maintenance procedures for clusters?

What could be typical pitfalls?

Please check our best practices for most current maintenance procedures – these slides only provide some top-level ideas.

Our best practices are available at www.suse.com/products/sles-for-sap
Generic Maintenance With Running Cluster

**PRE1** Set cluster to be in maintenance mode

**<YOUR MAINTENANCE PROCEDURE>**
like Updating SAP HANA

**POST1** Set the m/s SAPHana to unmanaged
**POST2** Set the cluster to be ready again
**POST3** Cleanup the m/s SAPHanaController
**POST4** Set m/s SAPHanaController to be managed
Generic Maintenance With **Stopped** Cluster

**PRE1** Set cluster to be in maintenance mode

**PRE2** Stop the cluster on node2 then node1

**<YOU-MAINTENANCE-PROCEDURE>**
(like Updating SAP HANA; Manually exchanging the SAP HANA SR roles)

**POST1** Start the cluster on node1 then node2

**POST2** Set the m/s SAPHanaController to unmanaged

**POST3** Set the cluster to be ready again

**POST4** Cleanup the m/s SAPHanaController

**POST5** Set m/s SAPHanaController to be managed again
Updating SAP HANA in System Replication
Steps 1 and 2

SUSE PRE-STEPS
(See “Generic Maintenance with Running/Stopped Cluster”)

Please always follow the SAP documentation to update SAP HANA. This is only an example procedure.

Update Secondary (nodeB)
- Stop Secondary on nodeB and unregister the former Secondary on nodeB
- Run the SAP HANA update procedure on nodeB
- Re-register the former Secondary on nodeB

Takeover production to nodeB
- Wait till the Secondary is completely in Sync (all services are “ACTIVE”)
- Stop the former Primary on nodeA
- Start the SAP HANA takeover on nodeB
Updating SAP HANA in System Replication
Steps 3 and 4

Update former Primary (nodeA)
- Disable system replication on nodeA
- Start the SAP HANA on nodeA
- Run the SAP HANA update procedure on nodeA
- STOP SAP HANA on nodeA, register the former Primary on nodeA and start SAP HANA on nodeA

Optionally re-migrate the primary to nodeA
- Wait till the system replication is in sync (all channels are in status active)
- Stop SAP HANA on nodeB
- Start the SAP HANA takeover on nodeA
- Register SAP HANA on nodeB and start SAP HANA on nodeB

SUSE POST-STEPS
(See “Generic Maintenance with Running/Stopped Cluster”)
Migrating the Primary using SAP HANA Tools

SUSE PRE-STEPS
(See “Generic Maintenance with Running/Stopped Cluster”)
Takeover and Register: SAP command line tool (hdbnsutil) or SAP HANA STUDIO
• Stop the primary
• Takeover to secondary: hdbnsutil -sr_takeover
• Register the former primary: hdbnsutil -sr_register ....
• Start the new secondary

SUSE POST-STEPS
(See “Generic Maintenance with Running/Stopped Cluster”)
Migrating the Primary using Pacemaker

Cmd line vs. HAWK

Always use “migrate-away from here”:

```bash
crm resource migrate <ms> force
```

Never use the “migrate to nodeX”

Do not forget to **un-migrate** after the primary is taken over

Depending on AUTOMATED_REGISTER you need manual registration of the “old” primary
Migration in the Future

We plan to get SAPHanaSR version 0.153.1 “migration aware”.

Do not forget to **un-migrate** after the primary is taken over – alternatively use time-limited migration rule.

Depending on AUTOMATED_REGISTER you need manual registration of the “old” primary.
Starting a Cluster with Orphan Primary
Lost Secondary Scenario

Understanding Data Integrity vs. Availability
How to tell the cluster that it is OK to start the PRIMARY with lost/stale SECONDARY

- Start PRIMARY node
- WAIT till the node starts the cluster software (crm_mon)
- START SAP HANA manually to overrule start protection
- WAIT till SAP HANA is up and running
- You are fine but should add the SECONDARY node soon
RISK: Full log area → DATABASE STUCK
SAP HANA Express
SAP HANA, express edition
Run SAP HANA on your own laptop or desktop

**DOWNLOADABLE SAP HANA IS NOW AVAILABLE!!!**

**SAP HANA, express edition**, is a **free downloadable** SAP HANA edition with a **smaller footprint** that can run on a **personal computer with 16GB memory (Sapphire)** – with a goal of further reducing this footprint.

<table>
<thead>
<tr>
<th>What does it include?</th>
<th>How can you get it?</th>
</tr>
</thead>
<tbody>
<tr>
<td>• A SAP HANA Virtual Appliance for Windows and Mac environments.</td>
<td>• Register: JAM Group</td>
</tr>
<tr>
<td>• A SAP HANA installation package for Linux environments</td>
<td>• Get started: receive email with download link and FAQ</td>
</tr>
<tr>
<td>• Smaller footprint</td>
<td>• Install: follow documentation to install SAP HANA on your computer</td>
</tr>
<tr>
<td>• Based on SAP HANA SPS 12</td>
<td></td>
</tr>
<tr>
<td>• Pre-configured and ready to use</td>
<td></td>
</tr>
<tr>
<td>• <strong>Vision:</strong> Introduce productive use model</td>
<td></td>
</tr>
</tbody>
</table>

- ✓ **Free to download and use**
- ✓ **Online Community Support**
- ✓ **Developer License (non-production use)**
- ✓ **Free renewable license**
SAP HANA Express on Intel NUC 5i / 7i

Your test cluster?

Get ➔ Install ➔ Integrate

Graphics provided by Intel
Results – What to take with

SUSE Linux Enterprise Server for SAP Applications is well prepared to limit the downtime of your workload.

SAP HANA SR supports various SAP HANA Scale-Up scenarios and the SAP HANA SR-Scale-Out package supports the Scale-Out scenario.

Setup of all mentioned scenarios and typical maintenance procedures are documented in the best practices available at www.suse.com.

SAP HANA express allows you to start your cluster test project with SLES for SAP applications today even with limited hardware resources.
Further Information

Best Practices

SUSE Linux Enterprise Server for SAP Applications
www.suse.com/products/sles-for-sap/

Training
elearning.suse.com/product-line/sles12/
training.suse.com/training/suse-linux-enterprise-server-2/
We adapt. You succeed.
Backup Slides
A normal worker node fails. Clients could still connect to the SAP HANA database. However answers which needs data from the failed node could not be processed.

SAP HA tries to repair this situation using a standby node:

- First of all the SAP HANA HA storage API must guarantee, that the old node does not longer have access to the data
- After the data partition is free for fail-over
SAP HANA Scale-Out – **Worker Failure**

Recover

(continue)

- Any available standby node could take the orphan data partition and get the worker node
- The active master name server will now redirect clients to the new node

**SAPHanaSR**

- detects all fail-overs of worker nodes
- checks the over-all landscape status of the SAP HANA database
- follows the decision of the SAP HA and checks, if the fail-over is successful
The active master name server is failing. All client connections are blocked.

As the active master name server is also a worker node SAP HA needs to fail-over the active master role including the worker part. The data partition 1 needs to be released.

One of the master name server candidates try to fail-over the active master name server role. In best this should be a standby node, because otherwise it’s data partition would be need to fail-over.
The new master name server mounts the data partition 1 and loads the data.

In the SAP HANA landscape this new node is shown as active master name server.

**SAPHanaSR**
- detects the fail-over of the active master name server and migrates the virtual IP address to that node
- allows clients to process a transparent reconnect
- follows the decision of the SAP HA and checks, if the fail-over is successful
A SAP HANA standby fails. It could be either a master name server candidate or a plain standby.

SAP HA does typically not repair this situation.

The running SAP HANA database is not directly influenced, but the HA capacity of the site gets degraded.
**SAP HANA Scale-Out – Standby Failure**

Recover

**SAPHanaSR**
- detects the outage of the SAP HANA standby node or instance
- restarts the failed SAP HANA standby instance, if the node is still part of the pacemaker cluster or rejoining the cluster
- takes care of the SAP HA fail-over “capacity” and increases the build-in SAP high availability
- checks, if the situation allows the restart of the standby or not.
**SAP HANA Scale-Up: Storage Replication**

<table>
<thead>
<tr>
<th><strong>Node 2 Usage:</strong></th>
<th>Dedicated</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Data pre-load on Secondary:</strong></td>
<td>No</td>
</tr>
<tr>
<td><strong>Take-over decision:</strong></td>
<td>Depends on storage vendor</td>
</tr>
<tr>
<td><strong>Take-over process:</strong></td>
<td>Depends on storage vendor</td>
</tr>
<tr>
<td><strong>Take-over reaction time:</strong></td>
<td>Depends on storage vendor</td>
</tr>
<tr>
<td><strong>Take-over speed:</strong></td>
<td>Slower since secondary copy must be completely loaded into memory</td>
</tr>
</tbody>
</table>

(See SAP Note 1755396 for solutions)
## Resource Agents (RA) and Monitoring SAP HANA

<table>
<thead>
<tr>
<th>RA</th>
<th>Scale</th>
<th>Decission</th>
<th>Monitoring-Interface</th>
</tr>
</thead>
<tbody>
<tr>
<td>SAPHana</td>
<td>Scale-Up</td>
<td>Landscape</td>
<td>landscapeHostConfiguration.py</td>
</tr>
<tr>
<td>SAPHanaController</td>
<td>Scale-Out</td>
<td>Landscape</td>
<td>landscapeHostConfiguration.py</td>
</tr>
<tr>
<td>SAPDatabase</td>
<td>Scale-Up</td>
<td>SAP ANY DB</td>
<td>SAPHOSTAGENT (saphostctrl/saphostexc)</td>
</tr>
<tr>
<td>SAPInstance</td>
<td>Scale-Up</td>
<td>SAP Instance</td>
<td>SAPSTARTSRV (sapstartsrv/sapctrl)</td>
</tr>
</tbody>
</table>
Unpublished Work of SUSE LLC. All Rights Reserved.
This work is an unpublished work and contains confidential, proprietary and trade secret information of SUSE LLC. 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.