TECHNICAL PAPER

Hitachi Unified Compute Platform 1000 Supports High Availability and System Replication for SAP HANA with Automated Failover in Scale-up Environments

Reference Architecture Guide

By Milind Pathak and Abhishek Dhanuka

December 2016

Hitachi Data Systems

Feedback

Hitachi Data Systems welcomes your feedback. Please share your thoughts by sending an email message to SolutionLab@hds.com. To assist the routing of this message, use the paper number in the subject and the title of this white paper in the text.

Contents

Hitachi Unified Compute Platform 1000 Supports High Availability and System Replication for SAP HANA with Automated Failover in Scale-up Environments

Reference Architecture Guide

Use this reference architecture guide to implement high availability for scale-up converged solutions with SAP HANA on Hitachi Unified Compute Platform 1000 (UCP). This environment uses SAP HANA system replication and automated failover. Implement this solution using Red Hat High Availability Add-On or SUSE Linux High Availability Extension.

Note — Testing of this configuration was in a lab environment. Many things affect production environments beyond prediction or duplication in a lab environment. Follow the recommended practice of conducting proof-of-concept testing for acceptable results in a non-production, isolated test environment that otherwise matches your production environment before your production implementation of this solution.

SAP HANA High Availability Features

SAP HANA database runs run mission-critical SAP applications. Solutions for SAP applications must remain highly available to users at all times. This requires that these systems make a faster recovery after system component failure (high availability) or after a disaster (disaster recovery). This should happen without any data loss (zero RPO) and in very short recovery time (low RTO).

This converged solution for SAP HANA high availability uses the high availability features of Hitachi Unified Compute Platform 1000, SAP HANA software, and either the Red Hat Linux operating system or SUSE Linux Enterprise Server as follows:

- **Hardware Failures**

  Hitachi Unified Compute Platform 1000 offers redundant hardware components to provide fault tolerance. This includes redundant power supplies, fans, and multiple Ethernet interfaces.

- **Software Failures**

  SAP HANA software includes a watchdog function that automatically restarts configured services in case of a failure to provide fault recovery. This includes the index server, name server and other SAP HANA database services.

In addition, this solution offers the following high availability mechanisms for SAP HANA. These solutions are based on completely redundant servers and/or storage.

- **Host Auto-Failover**

  Add one or more standby nodes to a SAP HANA system that are configured to work in standby mode. In case of failure, a standby node takes over the data and log volumes of a failed worker node. The standby node becomes a worker node, taking over user load. This type of solution does not need additional storage, only additional servers.

- **Storage Replication**

  Achieve data replication by storage mirroring that is independent from the database software. Mirror disks without a control process from SAP HANA. SAP HANA hardware partners offer this solution. This type of solution needs additional servers and storage.

- **SAP HANA System Replication**

  Configure the primary SAP HANA system to constantly replicate all data to a secondary SAP HANA system. Preload data constantly in the memory of the secondary system to minimize the recovery time objective (RTO). This type of solution needs additional servers and storage.

This reference architecture guide focuses on SAP HANA system replication.

See SAP HANA – High Availability FAQ to read more about SAP HANA high availability. This reference architecture uses SAP HANA System Replication.

SAP HANA System Replication Features

Implement SAP HANA system replication between two different SAP HANA systems with same number of active nodes. After setting up system replication between the two SAP HANA systems, all of the data replicates from the primary HANA system to the secondary HANA system (initial copy). After the initial copy, any logged changes in the primary system are sent also to the secondary system, but log entries are not replayed.

SAP HANA sends data snapshots from the primary system to the secondary system at regular intervals. This means whenever the secondary has to takeover, only log entries received after the last data snapshot need to be replayed.

With the data snapshot, the primary system also sends information about the tables loaded in memory, if the preload_column_tables parameter is set to **true**. If you set this parameter to **true** on the secondary system, these tables are preloaded in the memory of the secondary database.

These reduce RTO, making SAP HANA system replication a faster high availability solution, in terms of recovery. The entire process of data replication occurs on the software level and is fully controlled by the SAP HANA database kernel.

The following replication modes are available for system replication:

- **Synchronous on disk (mode=sync)**

  The transaction is committed after writing log entries on the primary system and the secondary systems.

  The synchronous replication mode (sync) can run with “full sync” enabled. In full sync operation, transaction processing on the primary site blocks, when the secondary is not currently connected and newly created redo log buffers cannot be shipped to the secondary site.

  This behavior ensures that no transaction can be committed locally without shipping the redo log buffers to the secondary site.

- **Synchronous in memory (mode=syncmem)**

  The Transaction is committed after the secondary system receives the logs, but before they are written to disks.

- **Asynchronous (mode=async)**

  The Transaction is committed after log entries are sent without any response from the secondary system.

You can run SAP HANA system replication in two different operation modes:

- **delta_datashipping**

  In addition to the continuous redo log shipping taking place, the secondary system requests a delta data shipping from time to time. The default time is every 10 minutes. During takeover the redo log needs to be replayed up to the last arrived delta data shipment.

  This is the default operation mode of SAP HANA system replication that is used for validation of this reference architecture.

- **logreplay**

  In this operation mode, pure redo log shipping is done after the system replication was initially set up with one full data shipping. The redo log is replayed on the secondary immediately after arrival, making this step superfluous during a takeover. This shortens RTO by factors.

  Additionally, this reduces the amount of data which needs to be transferred to the secondary site dramatically, because no delta data shipping is required anymore.

The procedure described in this reference architecture guide is applicable to all three replication modes. Set up system replication using SAP HANA Studio or the command line interface. Refer to SAP HANA Administration Guide to read more about system replication.

Figure 1 shows an overview of SAP HANA system replication.

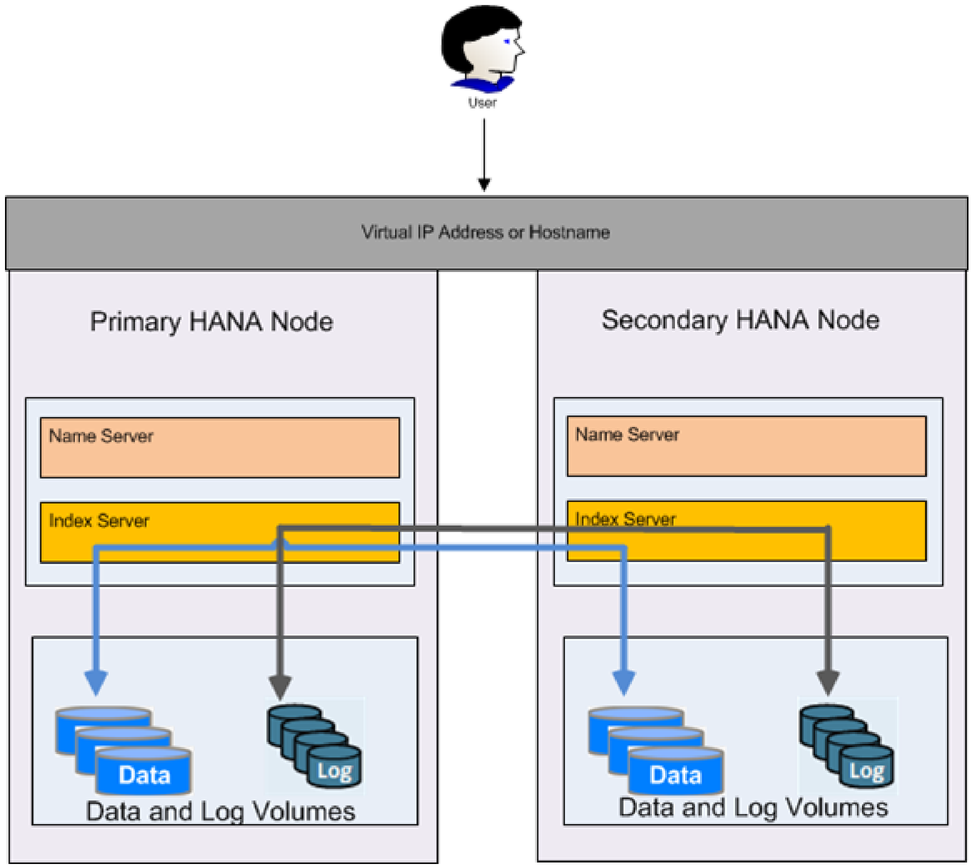


Figure 1

If the primary SAP HANA system fails, the system administrator must perform a manual takeover. Perform the takeover using SAP HANA Studio or the command line interface. Manual failover requires continuous monitoring, leading to higher recovery times.

To automate the failover process, Red Hat Linux High Availability Add-on or SUSE Linux Enterprise Server High Availability Extension can be used. Both of these meet those requirements by enhancing native SAP HANA replication and failover technology to automate the failover process. This helps you achieve service level agreements for SAP HANA downtime by enabling faster recovery without any manual intervention.

Guest Operating System Options

The following options are available as guest operating system:

- **Red Hat Enterprise Linux (RHEL)**

  Using the stability and flexibility of Red Hat Enterprise Linux, reallocate your resources towards meeting the next challenges instead of maintaining the status quo. Deliver meaningful business results by providing exceptional reliability on military-grade security. Use Enterprise Linux to tailor your infrastructure as markets shift and technologies evolve

- **SUSE Linux Enterprise Server (SLES) for SAP Applications**

  Compete more effectively though improved uptime, better efficiency, and accelerated innovation using SUSE Linux Enterprise Server. This is a versatile server operating system for efficiently, deploying highly available enterprise-class IT services in mixed IT environments with performance and reduced risk.

  SUSE Linux Enterprise Server was the first Linux operating system to be certified for use with SAP HANA. It remains the operating system of choice for the vast majority of SAP HANA customers.

Changing the configuration settings is only supported along the guidelines of SAP and the operating system distributor and may otherwise cause significant performance problems. The following SAP Notes for SLES and RHEL are a good starting point for information on this topic:

- 1944799 - SAP HANA Guidelines for SLES Operating System Installation
- 2009879 - SAP HANA Guidelines for Red Hat Enterprise Linux (RHEL) Operating System

For more details, see section 2.1.4.1, “Updating and Patching the Operating System” in the SAP HANA Technical Operations Manual.

Red Hat High-Availability Add-on

The Red Hat High Availability Add-On is an integrated set of software components that can be deployed in a variety of configurations to suit your needs for performance, high availability, load balancing, scalability, file sharing, and economy.

The add-on used in this reference architecture is based on ClusterLabs Pacemaker. A cluster configured with Pacemaker comprises the following:

- Separate component daemons that monitor cluster membership
- Scripts that manage the services
- Resource management subsystems that monitor the disparate resources

See High Availability Add-On Overview on the Red Hat website for the components of the Pacemaker architecture.

High availability setup for SAP HANA also uses **resource-agents-sap-hana** provided by Red Hat Linux for SAP HANA. It provides the following components:

- **SAPHanaTopology:** Gathers information about the current status of SAP HANA System Replication.
- **SAPHana:** Manages pre-configured SAP HANA System Replication environment.

SKU Red Hat Enterprise Linux for SAP HANA with High Availability and Smart Management is required. Red Hat Enterprise Linux for SAP HANA with High Availability and Smart Management is a variant of Red Hat Enterprise Linux for SAP HANA which enhances native SAP HANA replication and fail-over technology to automate the takeover process. It consists of these components:

- Red Hat Enterprise Linux for SAP HANA
- Red Hat Enterprise Linux High Availability Add-on

Refer to article Overview of Red Hat Enterprise Linux for SAP HANA subscription for details about Red Hat Linux SKUs.

SUSE Linux Enterprise High Availability Extension for SAP HANA High Availability

SUSE Linux Enterprise High Availability Extension is an integrated suite of open source clustering technologies that enable you to implement highly available physical and virtual Linux clusters. SUSE and SAP have developed a high availability solution for SAP HANA using SAP HANA system replication and SUSE Linux High Availability Extension.

A part of SUSE Enterprise Linux Server, SUSE Linux High Availability Extension requires no additional license. When installed and configured, SUSE Linux High Availability Extension and SAP HANA system replication provide the automated takeover mechanism for SAP HANA system replication.

SUSE Linux Enterprise Server for SAP Applications supports SAP HANA system replication using components of SUSE Linux Enterprise High-Availability Extension and the following resource agents:

- **SAPHanaTopology Resource Agent**

  SUSE has developed the SAPHanaTopology resource agent. This agent runs on all nodes of a SUSE Linux Enterprise High-Availability Extension cluster. It gathers information about the status and configurations of SAP HANA system replications. (Information from the SUSE website.)

- **SAPHana Resource Agent**

  This resource agent from SUSE supports scale-up scenarios by checking the SAP HANA database instances for whether a takeover needs to happen. Unlike with the pure SAP solution, takeovers can be automated.

  This resource agent is configured as a master/subordinate resource: The master assumes responsibility for the SAP HANA databases running in primary mode. The subordinate, sometimes called the slave, is responsible for instances that are operated in synchronous (secondary) status.

  In case of a takeover, the secondary (subordinate resource instance) can automatically be promoted to become the new primary (master resource instance). (Information from the SUSE website.)

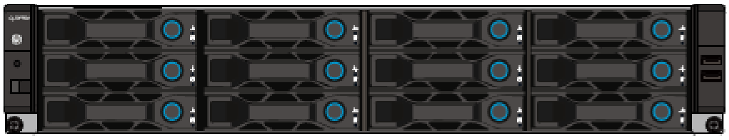
Rack Optimized Server for Solutions, 2U Single Node

The rack optimized server for solutions, 2U single node, is a rack mounted server designed for optimal performance and power efficiency. It supports up to 1.5 TB highly scalable memory capacity. It is powered by the Intel Xeon E5-2600 v3 processor product family for complex and demanding workloads. It supports flexible OCP and PCIe I/O expansion card options.

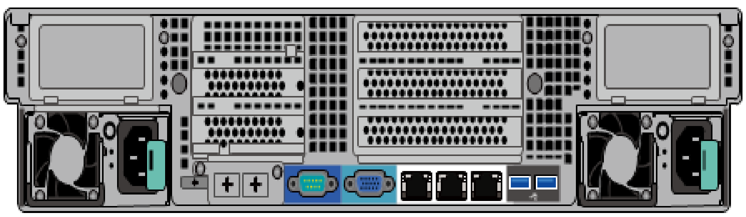
This reference architecture uses two identical servers for the primary and the secondary system in a SAP HANA database replication scenario. Use internal drives for storage.

On each of these two servers, install a scale-up SAP HANA system with same SID and instance number, as explained further Solution Design.

Refer to Hitachi Unified Compute Platform 1000 for the SAP HANA Platform in a Scale-up Configuration Using a Rack Optimized Server for Solutions, 2U Single Node Reference Architecture Guide (PDF) for details about architecture, network design, storage design, and the supported sizes of SAP HANA database.



rack optimized server for solutions,
2U single node (Front)



rack optimized server for solutions,
2U single node (Back)

Figure 2

Solution Overview

This reference architecture guide provides an example configuration of for SAP HANA high availability using SAP HANA system replication and automated takeover using one of these operating systems:

- Red Hat Enterprise Linux with the Red Hat High Availability Add-on
- SUSE Linux Enterprise Linux with the SUSE Linux High Availability Extension

Scale-up SAP HANA systems follow the architecture defined in Hitachi Unified Compute Platform 1000 for the SAP HANA Platform in a Scale-up Configuration using a Rack Optimized Server for Solutions, 2U Single Node Reference Architecture Guide (PDF). Connect the two SAP HANA systems through Brocade VDX 6740 Ethernet switches. Use internal drives for storage.

Validation of this environment was on two SAP HANA database nodes with 768 GB memory. Since SAP HANA system replication is a solution offered by SAP with automated takeover for SAP HANA system replication is a solution offered by Red Hat High-Availability Add-on or SUSE Linux High Availability Extension, this solution may be applied to all other scale-up systems supported by Hitachi Data Systems and SAP with appropriate hardware and network configuration changes. See “Cluster Configuration During Solution Validation,” starting on page 25.

Figure 3 on page 8 gives an overview of this solution.

- Set up SAP HANA system replication from node 1 to node 2.
- Red Hat High-Availability Add-on or SUSE Linux High Availability Extension with its SAPHana resource, performs the actual checks on the HANA database. Configure this resource as a master/subordinate resource. The master is responsible for primary SAP HANA system. The subordinate is responsible for the secondary system.
- The SAPHanaTopology resource monitors the status and configuration of SAP HANA system replication.
- Resource agents use the script **landscapeHostConfiguration.py** to monitor the status of the database.
- In case of failure of the primary node, the following happens:
  - The secondary node is promoted to take over the primary node role.
  - The virtual IP is moved to the secondary server.
  - When you initiate a connection to the SAP HANA database, you are automatically redirected to the active server. This requires no manual monitoring and takeover.
  - If **data preload** option is used, takeover time is significantly less than other SAP HANA high availability options (host auto-failover or storage replication).

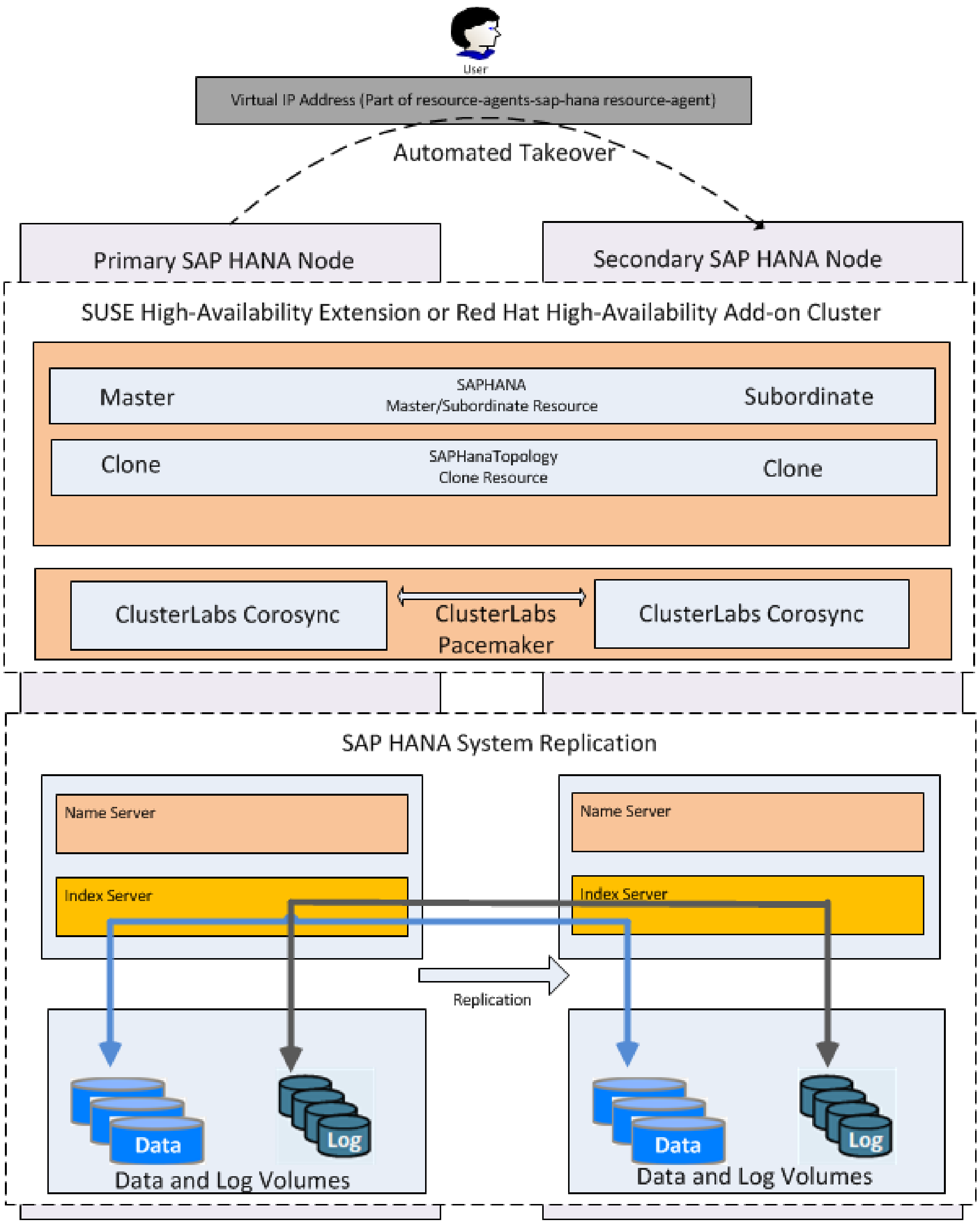

User
Virtual IP Address (Part of resource-agents-sap-hana resource-agent)
Automated Takeover
Primary SAP HANA Node
Secondary SAP HANA Node
SUSE High-Availability Extension or Red Hat High-Availability Add-on Cluster
Master
SAPHANA
Master/Subordinate Resource
Subordinate
Clone
SAPHanaTopology
Clone Resource
Clone
ClusterLabs Corosync
ClusterLabs
Pacemaker
ClusterLabs Corosync
SAP HANA System Replication
Name Server
Index Server
Name Server
Index Server
Replication
Data
Log
Data and Log Volumes
Data
Log
Data and Log Volumes

Figure 3

Key Solution Elements

These are the key hardware and software elements used in this reference architecture.

Hardware Elements

Table 1 describes the hardware for rack optimized server for solutions, 2U single node. In this solution, two such nodes are used for SAP HANA system replication.

Table 1. Hardware Elements

Hardware	Quantity	Configuration	Role
Rack Optimized Server for Solutions, 2U Single Node	1	■ Use only one of these processor types in any installation: ■ 2 Intel Xeon E5-2680 v3 processors (12 C, 2.5 GHz, 120 W) or ■ 2 Intel Xeon E5-2680 v4 processors (14 C, 2.3 GHz, 120 W) ■ 2 Heat sinks CPU 0/1 ■ RAM per SAP HANA node: ■ 64 GB (4 × 16 GB DIMMS): for non-production use only ■ 128 GB (8 × 16 GB DIMMS) ■ 256 GB (16 × 16 GB DIMMS) ■ 384 GB (24 × 16 GB DIMMS) ■ 512 GB (16 × 32 GB DIMMS) ■ 768 GB (24 × 32 GB DIMMS)	CPU for the SAP HANA server node
Internal storage drives on the Rack Optimized Server for Solutions, 2U Single Node (internal storage implementations only)	9	■ 1.2 TB, 10k RPM, 6 Gb/sec SAS internal hard drives	Block storage, when using internal storage
	2	■ 400 GB SATA solid state drives	
RAID controller card on the rack optimized server for solutions, 2U single node (internal storage implementations only)	1	■ LSI SAS 3108 12 Gb/sec 1 GB RAID controller mezzanine card	Required when using internal storage
	1	■ Activation key/card for LSI MegaRAID CacheCade Pro 2.0	

Table 1. Hardware Elements (Continued)

Hardware	Quantity	Configuration	Role
OCP mezzanine network card (all implementations)	1	■ Dual port 10 GbE Intel 82599ES SFP+ OCP mezzanine card	For SAP HANA 10 GbE client network
	2	■ 10 Gb/sec 850nm Multimode Datacom SFP+ transceiver	
Intel PCIe network cards (all implementations)	1	■ Dual port 10 GbE 82599ES SFP+ PCIe card	For replication and cluster communication using the 10 GbE network
	2	■ Ethernet SFP+ SR optics dual rate 10GBASE-SR/1000BASE-SX	
Brocade ICX 6430 24 port switch	1	■ 24 × 1 GbE ports	Optional switch for management network. Other 1 GbE network switch can be used.
Brocade VDX 6740 48 port switches	2	■ 48 × 10 GbE ports	Optional switches for things such as the client network or additional backup network. Other 10 GbE network switches can be used.
Minkel Global Solutions Rack (optional)	1	■ 1 standard rack	Optional rack for mounting server
PDUs (optional)	2	■ 12 outlet vertical PDUs	Optional PDUs for solution

Software Elements

Table 2 describes the software products used to deploy the two high availability SAP HANA nodes.

Table 2. Software Elements

Software Component		Software Version
SAP HANA		SPS12 or later
Operating System (use only one)	Red Hat Linux	Red Hat Enterprise Linux for SAP HANA with High Availability and Smart Management 7.2
	SUSE Linux Enterprise Server	SUSE Linux Enterprise Server 12 for SAP Applications
High availability add-on	For Red Hat Enterprise Linux: **resource-agents-sap-hana**	3.9.5-54.el7_2.17.x86_64
	For SUSE Linux Enterprise Server: **SAPHANASR**	0.151-0.11.1

Note — The resource-agents-sap-hana 3.9.5-54.el7_2.17.x86_64 (RHEL) and SAPHANASR 0.151-0.11.1 (SLES) was used to validate this reference architecture. However, Hitachi Data Systems strongly recommended that you use the latest version in your implementation. Also, use latest version of Red Hat High Availability Add-on or SUSE Linux High Availability Extension packages with their dependencies.

Solution Design

This is the detailed solution design of this reference architecture. It includes the following sections:

Planning and Prerequisites

To implement this solution, this reference architecture guide requires that the following:

- The hardware is already setup using the network architecture, as explained in "Network Design" on page 13
- The installation of the SAP HANA nodes on the primary and secondary servers has been performed using the design from Hitachi Unified Compute Platform 1000 for the SAP HANA Platform in a Scale-up Configuration Using a Rack Optimized Server for Solutions, 2U Single Node Reference Architecture Guide (PDF)

No license key is required for the secondary SAP HANA node, as this information will be replicated from the primary node to the secondary node.

Choose a replication mode from "SAP HANA System Replication Features" on page 2, based on your requirements.

When using data preload is used, set the parameter 'preload_column_tables' on both SAP HANA nodes. This preloads the main parts of the column tables into memory of the secondary node when they were loaded in the primary node.

With preload, the first access to a table that was previously used in the primary node only loads the changes. In most cases, the changes are typically much smaller than the main part of the column tables and can be loaded within seconds.

Other system parameters may be required for system replication to work efficiently. Their values depend on individual customer environments. Refer to System Replication Configuration Parameters for the list of parameters and their description.

Refer to Red Hat article Automated SAP HANA System Replication with Pacemaker on RHEL Setup Guide for supported scenarios and pre-requisites. Similarly refer to SUSE documentation on SAP HANA System Replication on SLES for SAP Applications for supported scenarios and prerequisites.

Also, check the prerequisites described by SAP from SAP HANA System Replication in the SAP HANA Administration Guide.

Following items are required for the setup of this high availability solution:

- SAP HANA system replication network IP addresses for the primary and the secondary SAP HANA servers
- Cluster Communication IP addresses and host names for the primary and the secondary SAP HANA servers
- Logical names for primary and secondary SAP HANA system replication sites
- SAP HANA SID and instance numbers, which must be identical for primary node and secondary node
- IPMI IP addresses for the primary and the secondary SAP HANA servers

Table 3 lists the information used to setup high availability in this reference architecture. Your configuration may use different values.

Table 3. Information Used for High Availability Setup

Configuration Detail	SAP HANA Node 1	SAP HANA Node 2
Public host name	saphanan1	saphanan2
IP address for client network	192.168.150.91	192.168.150.92
Cluster host name	node001	node002
IP Address for cluster Communication network	192.168.125.91	192.168.125.92
IP Address for replication network	192.168.175.91	192.168.175.92
IP Address for IPMI network	192.168.50.91	192.168.50.92
SAP HANA SID	HIT	HIT
SAP HANA instance number	10	10
Logical name for system replication	SITEA	SITEB

STONITH Device (Fencing)

STONITH is an acronym for “shoot the other node in the head.”

A STONITH device protects your data from being corrupted by rogue nodes or concurrent access. Just because a node is unresponsive, this does not mean the node is not accessing your data.

The only way to be 100% sure that your data is safe is to fence the node using STONITH so you can be certain that the node is truly offline before allowing the data to be accessed from another node.

STONITH also has a role to play in the event that a clustered service cannot be stopped. In this case, the cluster uses STONITH to force the whole node offline. This makes it safe to start the service elsewhere.

A Red Hat or SUSE cluster without the STONITH mechanism is not supported. There are multiple ways to implement STONITH. This reference architecture used the IPMI method. To implement IPMI based fencing, both the cluster nodes must be able to communicate with each other's BMC IP address. This is achieved by following the design explained in “Network Design” on page 13.

Network Design

The network architecture used in this reference architecture can be classified in two categories:

- “Management and IPMI Network” on page 13
- “Compute Network” on page 14

Management and IPMI Network

The management and IPMI network serves two purposes:

- Provide the management network for the following:
  - The rack optimized server for solutions, 2U single node
  - Brocade VDX switches
  - Brocade ICX switches

  Using this network, perform administrative tasks such as configuration of switches or accessing the BMC console of the server.
- Provide the network for implementing STONITH (fencing) through IPMI. In case one node needs to fence the other node, STONITH sends the reboot or shutdown command using this network.

Connect the 1 GbE management port on the rack optimized server for solutions, 2U single nodes and management server to a Brocade ICX 6430 24 port switch or to any other external 1 GbE switch for management and IPMI network connectivity as shown in Figure 4. Configure Ethernet interfaces of both rack optimized server for solutions, 2U single nodes, and assign IP addresses so that both the nodes can access each other's BMC IP addresses.

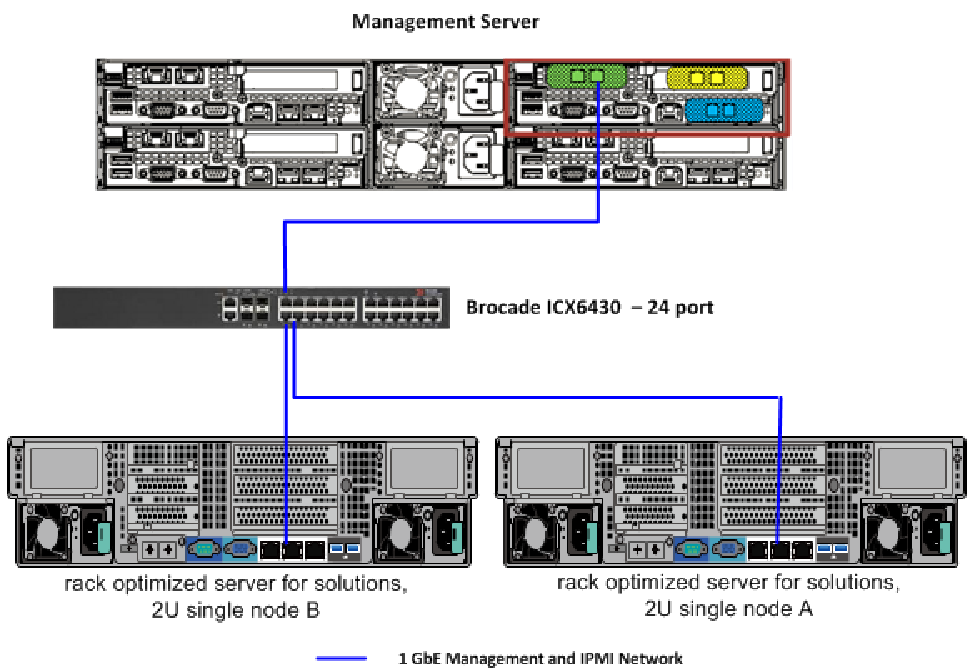


Figure 4

Compute Network

This solution requires three separate compute networks:

- **SAP HANA Client Network** — This provides communication between the SAP HANA production system and SAP production application servers.
- **SAP HANA System Replication Network** — This provides system replication between the SAP HANA production primary node and secondary node.
- **Cluster Communication Network** — This provides the cluster communication network for Red Hat Enterprise Linux High-Availability Add-on or SUSE Linux High Availability Extension.

Configure the SAP HANA system replication network and cluster communication network as a separate VLAN on the same network bonding interface. The SAP HANA client network is a dedicated network bond. Make the following 10 GbE network connections for the network setup, as shown in Table 4, “Network Setup,” on page 15:

- **Client Network**
  - Connect the following to Brocade VDX 6740 48 port switches:
    - Port 0 of dual port 10GbE Intel 82599ES SFP+ OCP mezzanine card
    - Port 1 of dual port 10GbE Intel 82599ES SFP+ OCP mezzanine card
  - Bond these corresponding two ports as bond0: eth9901 and eth9902. Bond these corresponding two ports as bond1: eth9911 eth9912. Do this at the operating system level using an **active-active** network bond mode with the options on both servers:

    mode= 802.3ad miimon=100 xmit_hash_policy=layer3+4 updelay=5000 lacp_rate=fast
  - Assign client network IP address to bond0.

    Bond0 acts as the client network for the SAP HANA nodes.
- **Replication Network and Cluster Communication Network**
  - Connect the following to Brocade VDX 6740 48 port switches:
    - Port 0 of dual port 10 GbE Intel 82599ES SFP+ PCI card
    - Port 1 of dual port 10 GbE Intel 82599ES SFP+ PCI card
  - Bond these corresponding two ports as bond1: eth9911 and eth9912. Do this at the operating system level using **active-active** network bond mode with the options on both servers:

    mode= 802.3ad miimon=100 xmit_hash_policy=layer3+4 updelay=5000 lacp_rate=fast
  - Create separate VLANs for the replication network and the cluster communication network.
  - Assign the replication network IP address and the cluster communication network IP addresses to the VLANs.
- Configure the external switch ports with a short LACP timeout value to speed up detecting corrupted connections.

Table 4. Network Setup

Network Card	Port	Network Description	VLAN
Slot 1	0	Client network for the SAP HANA node	150
	1	Client network for the SAP HANA node	
Slot 2	0	Replication and Cluster Communication network for the SAP HANA node	125 and 175
	1	Replication and Cluster Communication network for the SAP HANA node	

Figure 5 shows the connectivity of servers and switches.

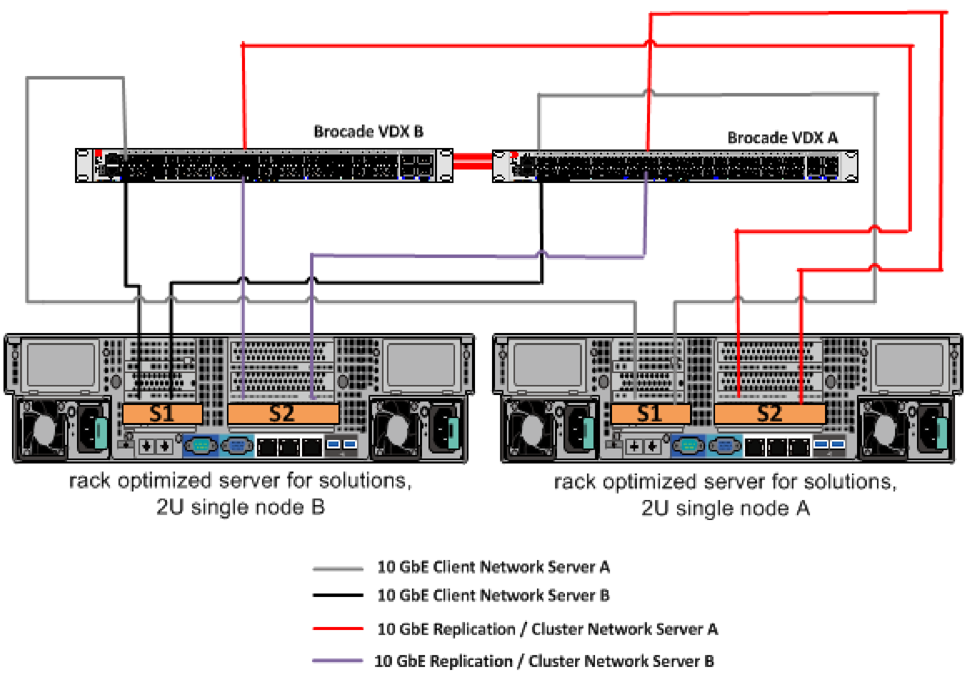


Figure 5

Configure Name Resolution for Replication Network

This solution uses a dedicated VLAN for the SAP HANA system replication network, as described in “Network Design” on page 13. SAP HANA database nodes must be configured to identify the replication network. Do this before configuring system replication.

Refer to [Network Configuration for SAP HANA System Replication]() to configure name resolution. Configure this in the *system_replication_hostname_resolution* section in the *global.ini* file.

Figure 6 on page 16 shows how to configure the IP addresses on the primary node and the secondary node.

```
hitadm@saphanan1:/hana/shared/HIT/global/hdb/custom/config> more global.ini
[persistence]
basepath_datavolumes = /hana/data/HIT
basepath_logvolumes = /hana/log/HIT

[system_replication]
operation_mode = delta_datashipping

[system_replication_hostname_resolution]
192.168.175.92 = saphanan2
hitadm@saphanan1:/hana/shared/HIT/global/hdb/custom/config>
```

```
hitadm@saphanan2:/hana/shared/HIT/global/hdb/custom/config> more global.ini
[persistence]
basepath_datavolumes = /hana/data/HIT
basepath_logvolumes = /hana/log/HIT

[system_replication]
operation_mode = delta_datashipping

[system_replication_hostname_resolution]
192.168.175.91 = saphanan1
hitadm@saphanan2:/hana/shared/HIT/global/hdb/custom/config>
```

Figure 6

Create a User and Configure the Userstore for Resource Agents

Create a monitoring user with sufficient privileges in SAP HANA for the resource agents to be able to run queries on the system replication status. Add a user key to hdbuserstore.

Note — Perform this procedure for the root user on the primary node and the secondary node.

For the procedure to create a user and configure the userstore for resource agents, refer to the following:

- Automated SAP HANA System Replication with Pacemaker on RHEL Setup Guide for Red Hat.
- SAP HANA System Replication on SLES for SAP Applications for SLES.

Verify that the key has been set correctly by running following command as the root user:

```
/usr/sap/HIT/HDB10/exe/hdbsql -U <USER KEY> -i 10 "SELECT CURRENT_USER FROM DUMMY"
```

You should be able to connect to the database using HDBSQL to see the output.

Setup SAP HANA System Replication

Configure system replication setup using the SAP HANA System Administration Guide. The configuration can be further tuned by applying parameters described in the System Replication Configuration Parameters and SAP HANA System Administration Guide, based on your individual requirements.

Validate that the system replication is running by running this command on the primary SAP HANA server as user <sid>adm, as shown in Figure 7.

```
hdbnsutil -sr_state
```

```
hitadm@saphanan1:/usr/sap/HIT/HDB10> hdbnsutil -sr_state
checking for active or inactive nameserver ...

System Replication State
~~~~~~~~~~~~~~~~~~~~~~~~
online: true

mode: primary
site id: 1
site name: SITEA

Host Mappings:
~~~~~~~~~~~~~~

saphanan1 -> [SITEA] saphanan1
saphanan1 -> [SITEB] saphanan2


done.
hitadm@saphanan1:/usr/sap/HIT/HDB10>
```

Figure 7

Configure the Cluster

This explains the installation and configuration of the Red Hat High-Availability Add-on or the SUSE Linux High Availability Extension cluster.

This reference architecture was validated using IPMI-based fencing. However, you can use shared storage based fencing, as well. For shared storage fencing, both SAP HANA nodes must be able to access the same device (LUN). To implement a shared storage fencing solution, refer to one of the following:

- Configuring Storage-based Fence Devices with Unfencing from Red Hat
- Section 5.2.2, “Adapting the Corosync and sbd Configuration” in SAP HANA SR Cost Optimized Scenario

Configure Red Hat High-Availability Add-on

Use this when using the Red Hat Enterprise Linux operating system.

Install the Add-on

Install the Red Hat High-Availability Add-on packages. The latest versions `of these packages` and their dependencies are provided with RHEL 7.2. These were used in validating this reference architecture. Perform the installation on the primary and the secondary SAP HANA database servers.

Configure the Add-on

See the following from Red Hat for the steps to configure a two node cluster and to configure required properties:

- Creating a Red Hat High-availability Cluster with Pacemaker
- Automated SAP HANA System Replication with Pacemaker on RHEL Setup Guide

Configure SUSE High Availability Extension

Use this when using the SUSE Linux Enterprise Server operating system.

Install the Extension

Install the SUSE Linux High Availability Extension packages. Latest versions `of these packages` (and their dependencies) provided with SLES12 are used in validation of this reference architecture. The installation must be performed on both primary and secondary SAP HANA database servers.

Configure the Extension

Refer to documentation from SUSE on SAP HANA SR Cost Optimized Scenario for steps to configure a two node cluster and configure required properties.

Configure the Resource Agent

This is how to configure the `resource agents`.

- For Red Hat High-Availability Add-on, configure the **resource-agents-sap-hana** package using the Red Hat cluster configuration command line interface. Follow the procedure described in Automated SAP HANA System Replication with Pacemaker on RHEL Setup Guide for the configuration steps.
- For SUSE Linux High Availability Extension, configure the SAPHANASR resource agents using **crm shell**. Follow the procedure described in SAP HANA SR Cost Optimized Scenario for the configuration steps.

Configure the following parameters:

- **SAP SID (SAP System Identifier)**. The SAP SID is always a 3-character alphanumeric string. **HIT** is the SAP SID used in this reference architecture guide.
- **SAP Instance Number**. The instance number must be a two-digit number, normally including a leading zero. However, instance number **10** is used in this reference architecture guide.
- **Virtual IP Address**. Configure the virtual IP address on the host where the primary database is running. Use this for application connectivity. Virtual IP address **192.168.150.199** is used in this reference architecture guide.

- **PREFER_SITE_TAKEOVER**. This parameter is required during configuration of SAP HANA resources. It decides whether the resource agent prefers to switch over to the subordinate instance instead of restarting the master instance locally in case of failure.
  - **true** — Prefer takeover to remote site
  - **false** — Prefer to restart locally

  This reference architecture guide uses **true** as the default value.

- **DUPLICATE_PRIMARY_TIMEOUT**. SAP HANA resources require configuration of this parameter. It is the time difference (in seconds) needed between two primary time stamps, if a dual-primary situation occurs. If the time difference is less than the time gap, then the cluster holds one or both instances in a "WAITING" status. This gives an administrator a chance to react on a takeover.

  If AUTOMATED_REGISTER is set to “true,” a failed former primary will be registered after the time difference is passed. After this registration as a secondary to the new primary, all data will be overwritten by the system replication.

- **AUTOMATED_REGISTER**. SAP HANA resources require configuration of this parameter. If a takeover event has occurred and the DUPLICATE_PRIMARY_TIMEOUT has expired, this decides whether a former primary instance should be registered automatically as a secondary by the resource agent during cluster/resource start. The default value is “false,” which was used when validating this solution.

After the configuration is complete, check that cluster members and all resources are online. Figure 8 on page 19 shows the cluster resources for Red Hat Enterprise Linux. Figure 9 on page 20 shows the cluster resources for SUSE Linux Enterprise Server. See “Cluster Configuration During Solution Validation” on page 25 for the complete cluster configuration used when validating this solution.

```
[root@saphanan1 ~]# pcs status
Cluster name: hanaHSR
Stack: corosync
Current DC: node002 (version 1.1.15-9.el7-e174ec8) - partition with quorum
Last updated: Mon Oct 17 08:24:27 2016          Last change: Mon Oct 17 08:24:04

2 nodes and 7 resources configured

Online: [ node001 node002 ]

Full list of resources:

 hana_ipmi_node1        (stonith:fence_ipmilan):        Started node001
 hana_ipmi_node2        (stonith:fence_ipmilan):        Started node002
 rsc_ip_SAPHana_HIT_HDB10       (ocf::heartbeat:IPaddr2):       Started node001
 Clone Set: rsc_SAPHanaTopology_HIT_HDB10-clone [rsc_SAPHanaTopology_HIT_HDB10]
     Started: [ node001 node002 ]
 Master/Slave Set: msl_rsc_SAPHana_HIT_HDB10 [rsc_SAPHana_HIT_HDB10]
     Masters: [ node001 ]
     Slaves: [ node002 ]

Daemon Status:
  corosync: active/disabled
  pacemaker: active/disabled
  pcsd: active/enabled
```

Figure 8

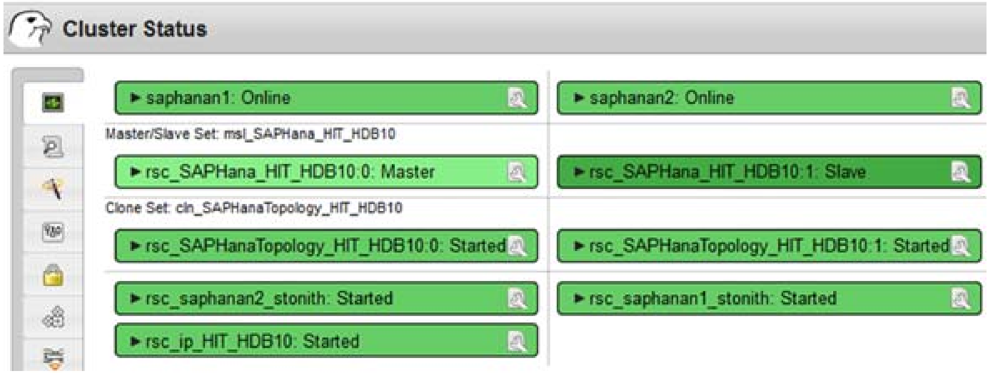


Figure 9

Note — These default timeout parameters configuration are a good starting point. Perform intensive testing and tune these parameters so they work better for your environment.

Engineering Validation

Table 5 lists some of the test cases performed in the Hitachi Data Systems lab to validate this solution. Before starting next test case, SAP HANA system replication and the cluster was brought to its original state.

Table 5. Test Cases for Engineering Validation

Test Case Number	Test Case Description	Result
1	Shutdown SAP HANA database on primary node *saphanan1* by command **HDB stop** to check automatic failover to secondary.	The SAP HANA database started on the secondary node *saphanan2* automatically within a few minutes and connectivity resumed in the SAP HANA Studio database through the virtual IP address.
2	Shut down the SAP HANA database on primary node *saphanan1* by stopping the **hdbdeamon** process by doing the following: 1. Run this command: `kill -9 <process ID of hdbdeamon>` 2. Check automatic failover to secondary.	The SAP HANA database started on the secondary node *saphanan2* automatically within a few minutes and then connectivity resumed in SAP HANA Studio through the virtual IP address.
3	Gracefully shut down the primary SAP HANA database server *saphanan1* by doing the following: 1. Run this command: `shutdown -h now` 2. Check the automatic failover to secondary.	The SAP HANA database started on secondary node *saphanan2* automatically within a few minutes and connectivity resumed in the SAP HANA Studio database through the virtual IP address.
4	Abruptly shut down the primary SAP HANA database server *saphanan1* by doing the following: 1. Run this command: `"echo 'b' > /proc/sysrq-trigger"` 2. Check the automatic failover to secondary.	The SAP HANA database started on secondary node *saphanan2* automatically within a few minutes. Connectivity resumed in the SAP HANA Studio database through the virtual IP address.

Table 5. Test Cases for Engineering Validation (Continued)

Test Case Number	Test Case Description	Result

5	Manually fence the SAP HANA database server *saphanan1* using the following command from *node002*:
	▪ **Red Hat Enterprise Linux:**
	`pcs stonith fence node001`
	▪ **SUSE Linux Enterprise High A**
	`crm node fence node001`
		The SAP HANA database started on the secondary node *saphanan2* automatically within a few minutes. Connectivity resumed in SAP HANA Studio database through the virtual IP address.

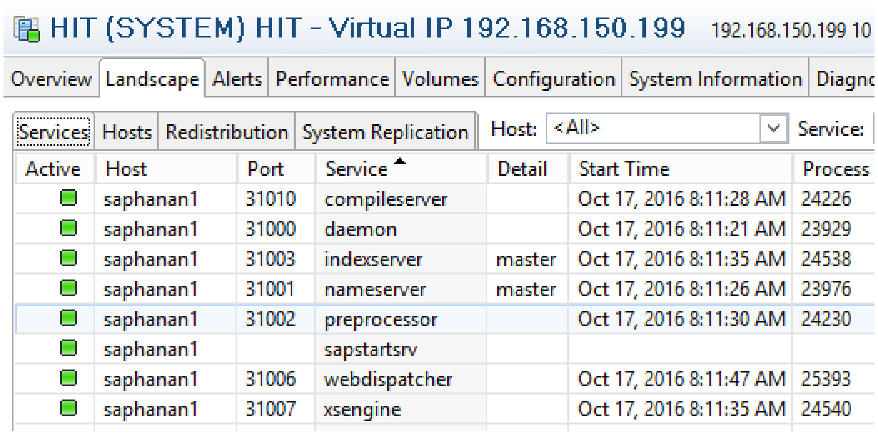
6	Shut down the cluster communication network on the primary SAP HANA database server *saphanan1* by running this command:
	`ifdown ifcfg-bond1.125`
	`Node001 is immediately fenced and rebooted.`
		The SAP HANA database started on the secondary node *saphanan2* automatically within a few minutes. Connectivity resumed in SAP HANA Studio database through the virtual IP address.

"Test Automated Failover" on page 23 describes the procedure to test automated failover using test case 1 in Table 5

Test Automated Failover

Automated failover is explained using test case 1 from Table 5, “Test Cases for Engineering Validation,” on page 21. Perform all other test cases to validate all the cluster and SAP HANA failover scenarios.

1. In SAP HANA Studio, verify that the primary node is running on the virtual IP address and on server *saphanan1*, as shown in Figure 10. All cluster resources must be up and running and SAP HANA system replication must be active.



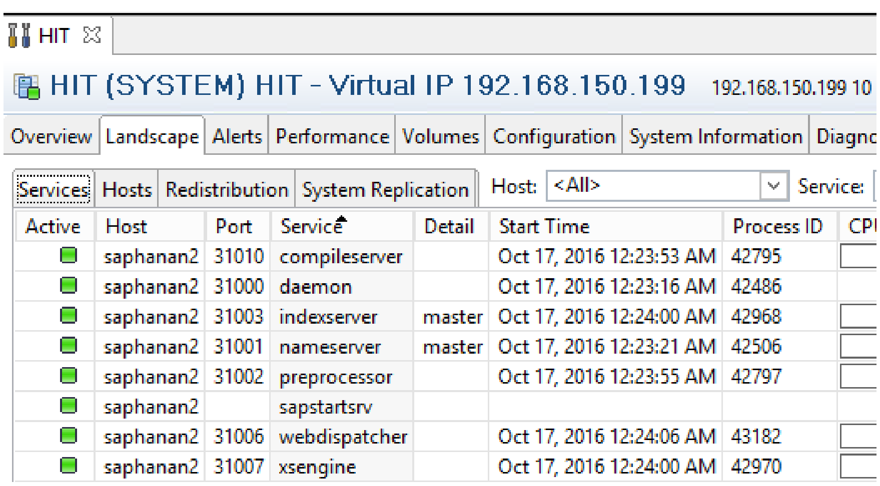
Active	Host	Port	Service	Detail	Start Time	Process
	saphanan1	31010	compileserver		Oct 17, 2016 8:11:28 AM	24226
	saphanan1	31000	daemon		Oct 17, 2016 8:11:21 AM	23929
	saphanan1	31003	indexserver	master	Oct 17, 2016 8:11:35 AM	24538
	saphanan1	31001	nameserver	master	Oct 17, 2016 8:11:26 AM	23976
	saphanan1	31002	preprocessor		Oct 17, 2016 8:11:30 AM	24230
	saphanan1		sapstartsrv			
	saphanan1	31006	webdispatcher		Oct 17, 2016 8:11:47 AM	25393
	saphanan1	31007	xsengine		Oct 17, 2016 8:11:35 AM	24540

Figure 10

2. Stop the HANA database on the primary server. Log on as user <sid>adm to execute this command: `HDB stop`

   Within a few seconds, the cluster detects that the HANA database is down on the primary server and automatic failover to the secondary server starts. The subordinate on the secondary server is promoted as master and the virtual IP address is moved to the secondary server. These events are written in cluster log file.

   After failover is complete, the HANA database on virtual IP address is now running on the secondary server *saphanan2*, as shown in Figure 11.



Active	Host	Port	Service	Detail	Start Time	Process ID	CPU
	saphanan2	31010	compileserver		Oct 17, 2016 12:23:53 AM	42795	
	saphanan2	31000	daemon		Oct 17, 2016 12:23:16 AM	42486	
	saphanan2	31003	indexserver	master	Oct 17, 2016 12:24:00 AM	42968	
	saphanan2	31001	nameserver	master	Oct 17, 2016 12:23:21 AM	42506	
	saphanan2	31002	preprocessor		Oct 17, 2016 12:23:55 AM	42797	
	saphanan2		sapstartsrv				
	saphanan2	31006	webdispatcher		Oct 17, 2016 12:24:06 AM	43182	
	saphanan2	31007	xsengine		Oct 17, 2016 12:24:00 AM	42970	

Figure 11

3. After the primary SAP HANA node '*saphanan1*' is up and running again, perform the steps described in How to Perform System Replication for SAP HANA to register it as a secondary node in system replication.

4. Verify from the cluster monitoring tool that the secondary node '*saphanan2*' is now the master server and the primary node '*saphanan1*' is the subordinate server.

This concludes the test.

Cluster Configuration During Solution Validation

These are the cluster configurations used to validate this solution in the Hitachi Data Systems lab. Further tuning of certain parameters may be required to make these configurations work in your environment. This includes updating these configurations to use the IP addresses and hostnames in your environment.

Red Hat High-Availability Add-on

Cluster Name: hanaHSR

Corosync Nodes:

node001 node002

Pacemaker Nodes:

node001 node002

Resources:

Resource: rsc_ip_SAPHana_HIT_HDB10 (class=ocf provider=heartbeat type=IPaddr2)

Attributes: ip=192.168.150.199

Operations: start interval=0s timeout=20s (rsc_ip_SAPHana_HIT_HDB10-start-interval-0s)

stop interval=0s timeout=20s (rsc_ip_SAPHana_HIT_HDB10-stop-interval-0s)

monitor interval=10s timeout=20s (rsc_ip_SAPHana_HIT_HDB10-monitor-interval-10s)

Clone: rsc_SAPHanaTopology_HIT_HDB10-clone

Meta Attrs: clone-max=2 clonenode-max=1 interleave=true is-managed=true

Resource: rsc_SAPHanaTopology_HIT_HDB10 (class=ocf provider=heartbeat type=SAPHanaTopology)

Attributes: SID=HIT InstanceNumber=10

Operations: start interval=0s timeout=600 (rsc_SAPHanaTopology_HIT_HDB10-start-interval-0s)

stop interval=0s timeout=300 (rsc_SAPHanaTopology_HIT_HDB10-stop-interval-0s)

monitor interval=10 timeout=600 (rsc_SAPHanaTopology_HIT_HDB10-monitor-interval-10)

Master: msl_rsc_SAPHana_HIT_HDB10

Meta Attrs: interleave=true clone-max=2 notify=true is-managed=true clone-node-max=1

Resource: rsc_SAPHana_HIT_HDB10 (class=ocf provider=heartbeat type=SAPHana)

Attributes: SID=HIT InstanceNumber=10 PREFER_SITE_TAKEOVER=true DUPLICATE_PRIMARY_TIMEOUT=7200 AUTOMATED_REGISTER=false

```
  Operations: start interval=0s timeout=3600 (rsc_SAPHana_HIT_HDB10-start-interval-0s)
              stop interval=0s timeout=3600 (rsc_SAPHana_HIT_HDB10-stop-interval-0s)
              promote interval=0s timeout=3600 (rsc_SAPHana_HIT_HDB10-promote-interval-0s)
              demote interval=0s timeout=3600 (rsc_SAPHana_HIT_HDB10-demote-interval-0s)
              monitor interval=59 role=Master timeout=700 (rsc_SAPHana_HIT_HDB10-monitor-interval-59)
              monitor interval=61 role=Slave timeout=700 (rsc_SAPHana_HIT_HDB10-monitor-interval-61)

Stonith Devices:
 Resource: hana_ipmi_node1 (class=stonith type=fence_ipmilan)
  Attributes: ipaddr=192.168.50.221 login=admin passwd=admin pcmk_host_list=node001
  Operations: monitor interval=60s (hana_ipmi_node1-monitor-interval-60s)
 Resource: hana_ipmi_node2 (class=stonith type=fence_ipmilan)
  Attributes: ipaddr=192.168.50.222 login=admin passwd=admin pcmk_host_list=node002
  Operations: monitor interval=60s (hana_ipmi_node2-monitor-interval-60s)
Fencing Levels:

Location Constraints:
Ordering Constraints:
  start rsc_SAPHanaTopology_HIT_HDB10-clone then start msl_rsc_SAPHana_HIT_HDB10 (kind:Mandatory)
(non-symmetrical) (id:order-rsc_SAPHanaTopology_HIT_HDB10-clone-msl_rsc_SAPHana_HIT_HDB10-mandatory)
Colocation Constraints:
  rsc_ip_SAPHana_HIT_HDB10 with msl_rsc_SAPHana_HIT_HDB10 (score:2000) (rsc-role:Started) (with-rsc-role:Master)
(id:colocation-rsc_ip_SAPHana_HIT_HDB10-msl_rsc_SAPHana_HIT_HDB10-2000)
Ticket Constraints:

Alerts:
 No alerts defined

Resources Defaults:
 default-resource-stickness: 1000
 default-migration-threshold: 5000
```

Operations Defaults:

timeout: 600s

Cluster Properties:

cluster-infrastructure: corosync

cluster-name: hanaHSR

dc-version: 1.1.15-9.el7-e174ec8

have-watchdog: false

last-lrm-refresh: 1476668661

no-quorum-policy: ignore

Node Attributes:

node001: hana_hit_remoteHost=saphanan2 hana_hit_site=SITEA hana_hit_srmode=syncmem hana_hit_vhost=saphanan1 lpa_hit_lpt=30

node002: hana_hit_remoteHost=saphanan1 hana_hit_site=SITEB hana_hit_srmode=syncmem hana_hit_vhost=saphanan2 lpa_hit_lpt=1476718740

SUSE Linux High Availability Extension

node 1084777563: saphanan1 \

attributes lpa_hit_lpt=10 hana_hit_vhost=saphanan1 hana_hit_site=SITEA hana_hit_srmode=syncmem hana_hit_remoteHost=saphanan2

node 1084777564: saphanan2 \

attributes hana_hit_vhost=saphanan2 hana_hit_remoteHost=saphanan1 hana_hit_site=SITEB hana_hit_srmode=syncmem lpa_hit_lpt=1477431126

primitive rsc_SAPHanaTopology_HIT_HDB10 ocf:suse:SAPHanaTopology \

params SID=HIT InstanceNumber=10 \

op monitor interval=10 timeout=600 \

op start interval=0 timeout=600 \

op stop interval=0 timeout=300

```
primitive rsc_SAPHana_HIT_HDB10 ocf:suse:SAPHana \
        params SID=HIT InstanceNumber=10 PREFER_SITE_TAKEOVER=yes AUTOMATED_REGISTER=no
DUPLICATE_PRIMARY_TIMEOUT=7200 \
        op start interval=0 timeout=3600 \
        op stop interval=0 timeout=3600 \
        op promote interval=0 timeout=3600 \
        op monitor interval=60 role=Master timeout=700 \
        op monitor interval=61 role=Slave timeout=700
primitive rsc_ip_HIT_HDB10 IPaddr2 \
        params ip=192.168.150.199 \
        op start timeout=20 interval=0 \
        op stop timeout=20 interval=0 \
        op monitor interval=10 timeout=20
primitive rsc_saphanan1_stonith stonith:external/ipmi \
        params hostname=saphanan1 ipaddr=192.168.50.221 userid=admin passwd=admin interface=lanplus \
        op stop interval=0 timeout=60 \
        op start interval=0 timeout=60 \
        op monitor interval=0 timeout=60
primitive rsc_saphanan2_stonith stonith:external/ipmi \
        params hostname=saphanan2 ipaddr=192.168.50.222 userid=admin passwd=admin interface=lanplus \
        op stop interval=0 timeout=60 \
        op start interval=0 timeout=60 \
        op monitor interval=0 timeout=60 \
        meta target-role=Started
ms msl_SAPHana_HIT_HDB10 rsc_SAPHana_HIT_HDB10 \
        meta clone-max=2 clone-node-max=1 interleave=true
clone cln_SAPHanaTopology_HIT_HDB10 rsc_SAPHanaTopology_HIT_HDB10 \
        meta is-managed=true clone-node-max=1 interleave=true
location loc_saphanan1_stonith rsc_saphanan1_stonith -inf: saphanan1
location loc_saphanan2_stonith rsc_saphanan2_stonith -inf: saphanan2
colocation col_saphana_ip_HIT_HDB10 2000: rsc_ip_HIT_HDB10:Started msl_SAPHana_HIT_HDB10:Master
```

```
order ord_SAPHana_HIT_HDB10 Optional: cln_SAPHanaTopology_HIT_HDB10 msl_SAPHana_HIT_HDB10
property cib-bootstrap-options: \
        stonith-enabled=true \
        stonith-action=reboot \
        stonith-timeout=300s \
        startup-fencing=false \
        no-quorum-policy=ignore \
        placement-strategy=balanced \
        have-watchdog=false \
        dc-version=1.1.13-14.7-6f22ad7 \
        cluster-infrastructure=corosync \
        cluster-name=hacluster \
        last-lrm-refresh=1477429129
rsc_defaults rsc-options: \
        resource-stickiness=1000 \
        migration-threshold=3000
op_defaults op-options: \
        timeout=600 \
        record-pending=true
```

Network Design for 520X B3 Server Blades

This solution was validated on the rack optimized server for solutions, 2U single node. However, since SAP HANA System Replication is a solution offered by SAP, and automated takeover for SAP HANA System Replication is a functionality offered by Red Hat High-Availability Add-on or SUSE Linux High Availability Extension, the procedure described in this reference architecture guide may be applied to other solutions from Hitachi Data Systems and SAP-supported SAP HANA platform servers. If implemented on other server blades from Hitachi Data Systems, the network design must be changed as required.

This describes the network design for the 520X B3 server blades (used in 4-socket scale-up systems) in a Hitachi Compute Blade 2500 chassis. Refer to Hitachi Unified Compute Platform 6000 for the SAP HANA Platform in a Scale-up Configuration with Intel Xeon E7-88xx v4 Processors Reference Architecture Guide (AS-503-01 or later, PDF) for details about the design of a scale-up solution.

The 520X B3 server blades offer two dedicated 10 GbE ports used for system replication. There are two 10GBASE-SR 2-port LAN adapters installed on the PCIe slots of the I/O board module of blade 1 of the Hitachi Compute Blade 2500 chassis. This solution uses two 10 GbE ports on the 10GBASE-SR 2-port LAN adapters for connectivity with the 10 GbE external switches. This solution requires three separate compute networks:

- **SAP HANA Client Network** — This provides communication between the SAP HANA production system and SAP production application servers.
- **SAP HANA System Replication Network** — This is used for SAP HANA system replication between the primary production node and the secondary node.
- **Cluster Communication Network** — This is used for the cluster communication network.

Configure the SAP HANA System Replication network and cluster communication network as separate VLANs on the same network bonding interface. The SAP HANA client network is a dedicated network bond.

Connect the management module on the Hitachi Compute Blade 2500 chassis to an external switch for management connectivity. Make the following network connections for the client, replication, and cluster communication networks, as shown in Figure 12 on page 31.

- Port 0 of the I/O board module on PCIe slot IOBD 01B to port 1 of Brocade VDX 6740-48B
- Port 0 of the I/O board module on PCIe slot IOBD 02B to port 1 of Brocade VDX 6740-48A
- Port 1 of the I/O board module on PCIe slot IOBD 01B to port 3 of Brocade VDX 6740-48B
- Port 1 of the I/O board module on PCIe slot IOBD 02B to port 3 of Brocade VDX 6740-48A
- Port 0 of the I/O board module on PCIe slot IOBD 05B to port 2 of Brocade VDX 6740-48B
- Port 0 of the I/O board module on PCIe slot IOBD 06B to port 2 of Brocade VDX 6740-48A
- Port 1 of the I/O board module on PCIe slot IOBD 05B to port 4 of Brocade VDX 6740-48B
- Port 1 of the I/O board module on PCIe slot IOBD 06B to port 4 of Brocade VDX 6740-48A

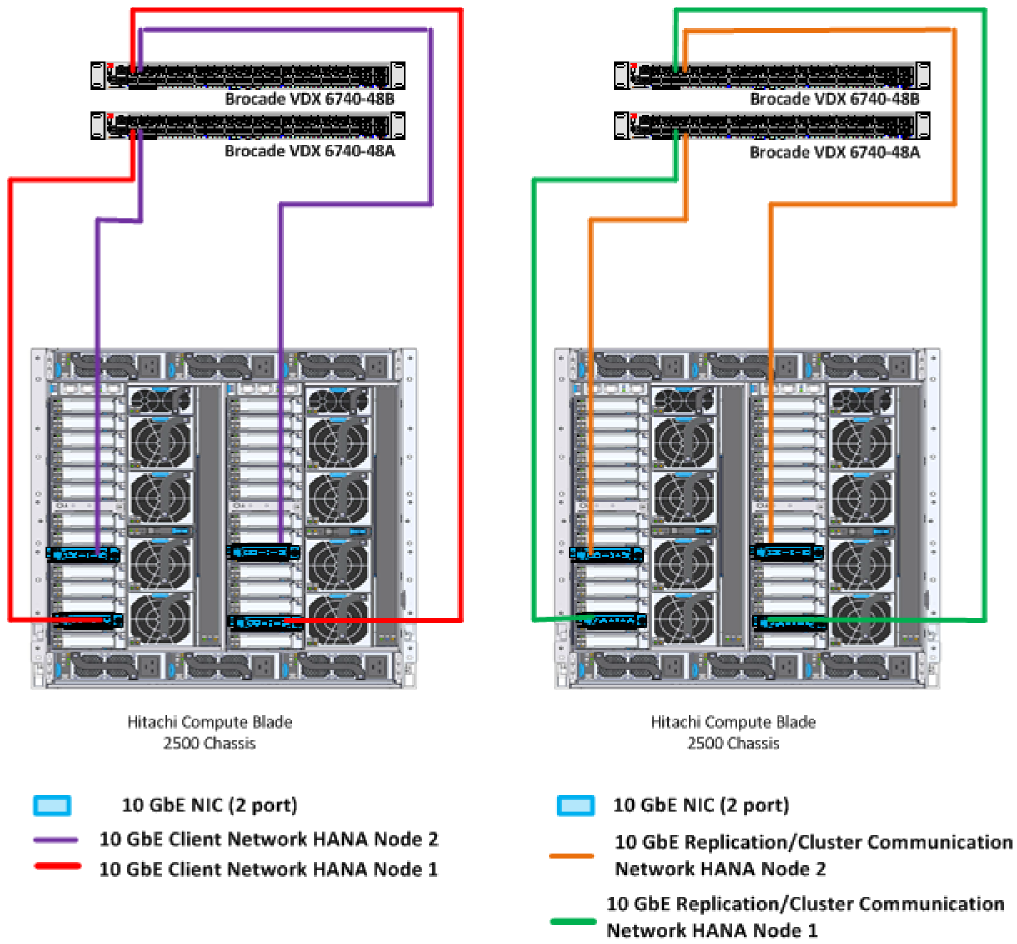


Figure 12

This solution connects two **Brocade VDX 6740** switches together using ISL. It enables both switches to act together as one single logical switch with the characteristics that, if one switch fails, there still is a path to the hosts. The client network is a dedicated two port bond.

- At the operating system level, create bond0 and bond1 using the ports as described in Table 6, "Network Setup Using 10GBASE-SR 2-Port LAN Adapter," on page 32. Use **active-active** network bond mode with these options:

  ```
  mode= 802.3ad miimon=100 xmit_hash_policy=layer3+4 updelay=5000 lacp_rate=fast
  ```

- Assign IP address to bond0 as described in Table 6.
- Create separate VLANs 100 and 200 for the replication network and the cluster communication network.
- Assign the replication and cluster communication network IP addresses to the VLANs, as described in Table 6.

Table 6. Network Setup Using 10GBASE-SR 2-Port LAN Adapter

HANA Node	PCIe Slot Number	Switch Module Port	Network Description	Bond	IP Address
Primary SAP HANA server	IOBD 01B	0	Client network for SAP HANA node 1	Bond0	192.168.150.111
	IOBD 02B	0	Client network for SAP HANA node 1		
	IOBD 01B	1	Replication network for SAP HANA node 1	Bond1	VLAN 100: 192.168.100.111 VLAN 200: 192.168.200.111
	IOBD 02B	1	Replication network for SAP HANA node 1		
Secondary SAP HANA server	IOBD 05B	0	Client network for SAP HANA node 2	Bond0	192.168.150.112
	IOBD 06B	0	Client network for SAP HANA node 2		
	IOBD 05B	1	Replication network for SAP HANA node 2	Bond1	VLAN 100: 192.168.100.112 VLAN 200: 192.168.200.112
	IOBD 06B	1	Replication network for SAP HANA node 2		

Shared Storage-based STONITH Device

This solution was validated using an IPMI-based fencing mechanism. However, if a shared LUN exists between the two cluster nodes, storage-based fencing can be used.

Create a small LUN (50 MB) on the storage array that is shared between the cluster members. Map this LUN to the primary and the secondary SAP HANA servers through the storage ports. Make note of the SCSI identifier of the block device (/dev/disk/by-id/scsi*) of this LUN. Use the same SCSI identifier on the primary and the secondary SAP HANA servers. It is possible to add more than one device in a cluster for redundancy.

For an environment using Red Hat Enterprise Linux, refer to Using SCSI Persistent Reservation Fencing (fence_scsi) with pacemaker in a RHEL 6 or 7 High Availability cluster for pre-requisites and further details.

For an environment using SUSE Linux Enterprise Linux, refer to Storage-based Fencing for pre-requisites and further detail.

For More Information

Hitachi Data Systems Global Services offers experienced storage consultants, proven methodologies and a comprehensive services portfolio to assist you in implementing Hitachi products and solutions in your environment. For more information, see the Hitachi Data Systems Global Services website.

Live and recorded product demonstrations are available for many Hitachi products. To schedule a live demonstration, contact a sales representative. To view a recorded demonstration, see the Hitachi Data Systems Corporate Resources website. Click the **Product Demos** tab for a list of available recorded demonstrations.

Hitachi Data Systems Academy provides best-in-class training on Hitachi products, technology, solutions and certifications. Hitachi Data Systems Academy delivers on-demand web-based training (WBT), classroom-based instructor-led training (ILT) and virtual instructor-led training (vILT) courses. For more information, see the Hitachi Data Systems Services Training and Certification website.

For more information about Hitachi products and services, contact your sales representative or channel partner or visit the Hitachi Data Systems website.

Hitachi Data Systems

Corporate Headquarters
2845 Lafayette Street
Santa Clara, CA 96050-2639 USA
www.HDS.com community.HDS.com

Regional Contact Information
Americas: +1 408 970 1000 or info@hds.com
Europe, Middle East and Africa: +44 (0) 1753 618000 or info.emea@hds.com
Asia Pacific: +852 3189 7900 or hds.marketing.apac@hds.com

AS-564-00. December 2016.