Best Practices and Recommendations for Scale-Out Deployments of SAP HANA on VMware vSphere

August 2015

DEPLOYMENT AND TECHNICAL CONSIDERATIONS GUIDE

Table of Content

Introduction

Since the production support statement for VMware virtualized Scale-Up configurations in 2014, customers have been able to enjoy the flexibility, high availability and lower total cost of ownership by running SAP HANA virtualized on VMware vSphere®.

By extending platform support to SAP HANA Scale-Out configurations, running on VMware vSphere virtualization and cloud computing platforms, any database sizes that are supported by SAP on physical SAP HANA Scale-Out systems can now be deployed on top of a VMware virtualized infrastructure, allowing customers to deploy any SAP HANA type and workload virtualized.

Using the SAP HANA platform with VMware vSphere virtualization infrastructure for SAP HANA Scale-Out configurations provides benefits physical deployments of SAP HANA cannot provide, like:

- Increased security
- Built-in multi-tenancy support via system encapsulation in a virtual machine (VM)
- Abstraction of the hardware layer
- Higher hardware utilization rates
- Live migration of running SAP HANA instances to other vSphere host systems

These and other advanced features—found exclusively in virtualization—lower the total cost of ownership and provide better service-level agreements (SLAs) and flexibility.

This guide complements "*Best Practices and Recommendations for Scale-Up Deployments of SAP HANA on VMware vSphere*[i]," and describes the virtualized, sizing, best practices and recommendations for configuring, deploying, and optimizing SAP HANA Scale-Out deployments running on VMware virtualization infrastructure.

Many of the findings in this guide are a result of joint testing conducted by VMware, VCE, HDS and SAP to characterize the performance of SAP HANA Scale-Up and Scale-Out deployments powered by VMware vSphere.

Audience

This guide is intended for IT directors and managers who want to leverage SAP HANA's new in-memory database and are looking for guidance on how to implement SAP HANA cost efficiently and flexibly by meeting the organization's IT virtualization strategy. It was also written for architects, engineers, and administrators who are responsible for configuring and deploying the SAP HANA platform in a VMware virtualization environment.

It is assumed that the reader has a basic knowledge of VMware vSphere concepts and features, SAP HANA, and related SAP products and technologies.

Production Support

In November 2012, SAP announced initial support for SAP HANA on vSphere 5.1 for non-production environments. In April 2014, SAP extended this support to SAP HANA Scale-Up production environments on vSphere 5.5, including SAP Tailored Datacenter Integration (TDI) deployments and multiple SAP HANA VMs running on a single vSphere host in controlled availability (CA). In May 2015, SAP provided production CA support for the last remaining deployment option of SAP HANA - the so-called Scale-Out configurations.

In a scale-out scenario, an SAP HANA database system is distributed over multiple virtual machines running on several physical vSphere host systems, either as the sole VM on a host, or alongside other SAP HANA instances (or “any workload” VMs) on the same host system.[ii]

Leveraging SAP HANA Scale-Out configurations for SAP Business Warehouse (BW) workloads, customers can incrementally add new physical host server systems and VMs as needed—each with up to 3 TB[iii] of memory—to adapt easily and quickly to data growth.

VMware can support SAP HANA Scale-Out system RAM sizes up to a total of 1 TB per vSphere 5.5 VM, and once supported on vSphere 6, up to 4 TB RAM per VM. Dozens of SAP HANA virtual Scale-Out machines running on a single or multiple VMware vSphere cluster can combine to scale to the required SAP HANA database size, limited only by hardware resources (server, storage and network) and SAP support status.

SAP HANA System Types

Overview

The SAP HANA platform is data source agnostic and a flexible data source in-memory platform that allows customers to analyze large volumes of data in real time.[iv] It is deployable on-premise or in the cloud. It is a revolutionary platform best suited for performing real-time analytics and developing and deploying real-time applications. SAP HANA can be deployed in two different deployment options: Scale-Up or Scale-Out. Which deployment option an organization selects depends strongly on the planned use case, the application and the size and data growth. For BW deployments the following additional points may get considered:

- BW supports Scale-Up and Scale-Out deployments
  - Size of the source database
  - Expected data growth

Business Suite applications are not yet generally supported on Scale-Out deployments and should get deployed on Scale-Up configurations.

The SAP HANA installation guide refers to these deployment options as SAP HANA Single- or Multiple-Host configurations[v].

Scaling an SAP HANA system depends on the deployed system type and by either:

- Adding more compute resources
  - Like RAM and CPU power to a single host, and then
  - Extending the VM size to a maximum of 1 TB RAM and 64 vCPUs with vSphere 5.5, and up to 4 TB and 128 vCPUs with vSphere 6 once supported[vi]
- Or when installed in a Multiple-Host configuration by adding more VMs.

For a simplified view of these two deployment options on a VMware vSphere virtualized environment - and how to scale up or out to a larger SAP HANA system – see figure 1. Both solutions provide their unique benefits, like easier management for a Scale-Up configuration, or cost efficient expansion for greater storage requirements. For instance, it is more expensive to expand the memory of a single server than it is to buy small servers with smaller, less expensive memory module sizes. See table 1 for an example calculation with list prices of different sized memory modules.

The current RAM size limitations of a VM are 1 TB for vSphere 5.5 and 4 TB for vSphere 6, once vSphere 6 is supported for SAP HANA workloads.

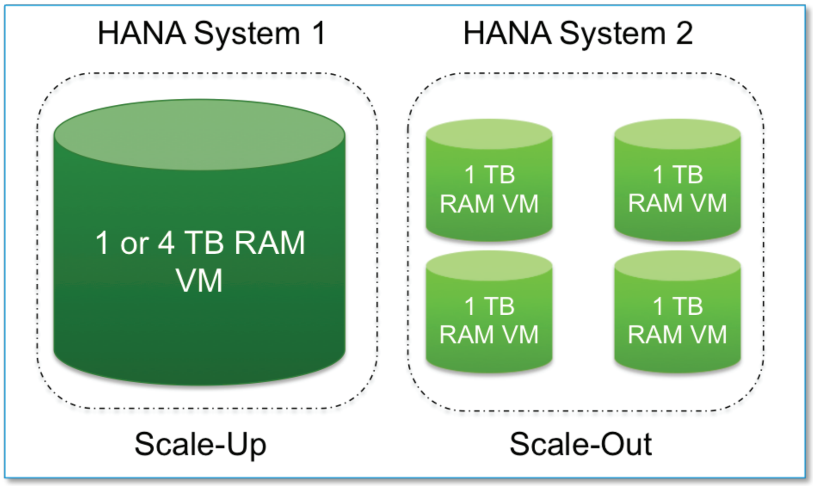


Figure 1. Scaling Options of an SAP HANA System on VMware vSphere

Table 1 summarizes the costs of 4 TB RAM when using different memory module sizes and when calculating with list prices as of April 2015. Using 64 GB DRAM modules are up to 6.7 times more expensive than 16 GB DRAM modules for the same amount of memory (and with existing memory module limitations of today's servers.

MODULE	16 GB DRAM	32 GB DRAM	64 GB DRAM
Target: 4 TB	256 modules	128 modules	64 modules
List price per module as of April 2015	$260 USD[vii]	$700 USD[viii]	$7,000 USD[ix]
Total list price	$66,560 USD	$89,600 USD	$448,000 USD
Price factor	1.0	1.346	6.731

Table 1. Cost Comparison of RAM Module List Prices as of April 2014

A typical modern 4-socket server can support up to 96 DIMM slots. This means that 64 GB DIMM modules must be used in that server, when 4 TB RAM is needed. When using an 8-socket server, 32 GB DIMM modules can be used, as the server provides more DIMM slots. This could dramatically reduce the cost of RAM, but the initial server costs are higher compared to a 4-socket server. Also, the SAP HANA socket-to-RAM ratios for Business Suite and Business Warehouse on HANA (BWoH) have to be followed, and may require—especially for very large SAP HANA systems—the usage of 8-CPU socket systems; this again will increase the costs compared to smaller server systems.

Depending on the organization's IT hardware strategy, there may also be additional dependencies on hardware (HW) standards and purchasing decisions; table 1—with the prices for RAM—presents only a simplified view of the costs of a large- vs. small-sized memory configuration. The price comparison does not include complex management tasks or operations which would compare the total cost of ownership of a Scale-Out vs. Scale-Up environment. When comparing the overall costs of a server landscape, all costs—not just the memory price—need to be considered, but the RAM prices provide an indication of the pure hardware costs of large memory footprint server systems and should be considered when deciding on scaling an SAP HANA system up or out.

Using SAP HANA allows scaling-up and -out as required, and is defined by the organization's IT strategy; in the case of pure memory costs, scaling SAP HANA out is an attractive option and is fully supported by VMware vSphere.

When scaling an SAP HANA instance, remember that sizing must meet SAP sizing requirements and follow the currently existing fixed-CPU socket-to-RAM ratios for BW and Business Suite workloads.

SAP HANA Scale-Up Option

The Scale-Up, or Single-Host SAP HANA configuration as named in the installation guide, describes in the context of virtualization, a single SAP HANA instance that runs on a single VM on a vSphere host.

With a Scale-Up configuration, a single SAP HANA VM can be configured to consume all available host resources. Once a VM reaches the host or vSphere limits, if cannot grow larger. Due to the CPU socket-to-RAM ratio limitation, the biggest server systems and RAM sizes available for BWoH workload are 8-socket server systems, with 2 TB RAM and up to 6 TB RAM for specific SAP Business Suite workloads. Larger systems like 16- or 32-socket server systems will become available, and will allow the consolidation/co-deployment of several maximum-sized SAP HANA VMs on a single server system.

Figure 2 shows such a single SAP HANA instance running on a single VM, that consumes all available server/VM resources. An SAP HANA system that already consumes the maximum server or VM limit won't be able to grow beyond these limitations. Adding resources or changing the VM resource configuration is required to support scale-up growth.

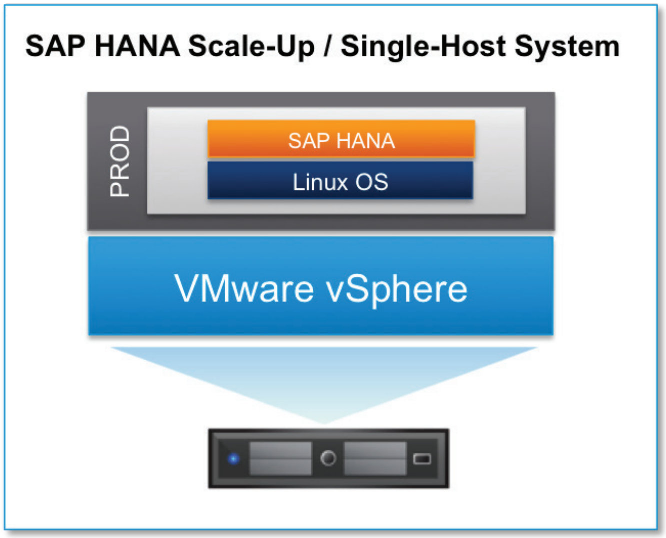


Figure 2. SAP HANA Single-Host System (Scale-Up) Configuration

Figure 3 shows the same configuration (but in a so-called **Multi-VM configuration** mode) of single, independent SAP HANA systems running co-deployed on a single host system. SAP HANA multi-VM configurations that run on a single host are supported in production in controlled availability and non-production workloads. It is required that the host system has enough resources (CPU, RAM, network and storage) available to support all co-deployed VMs, and that CPU and memory resources of production VMs get reserved. Non-production, non-performance critical workload VMs can share resources to better utilize available hardware resources. If several production-level SAP HANA VMs are co-deployed on a single host, required resources must be reserved for all VMs.

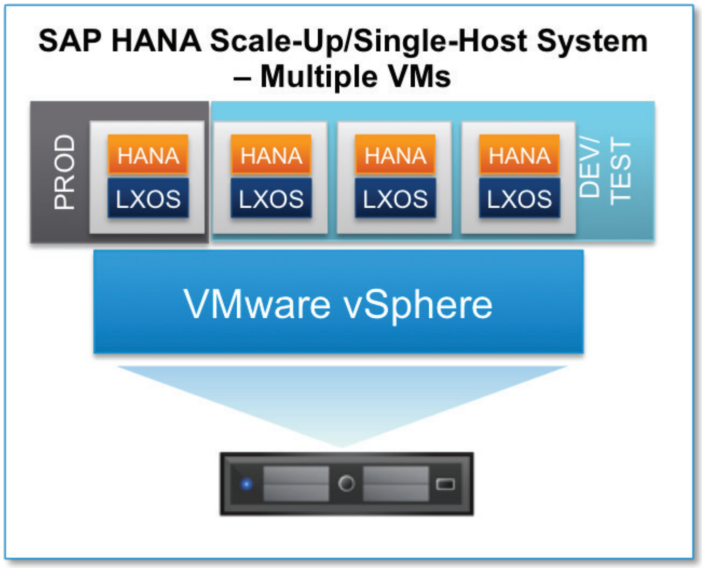


Figure 3. SAP HANA Multi-VM Single-Host System (Scale-Up) Configuration Example

SAP HANA Scale-Out Option

In an SAP HANA Scale-Out configuration running on vSphere, multiple SAP HANA "worker VMs" are distributed over multiple vSphere hosts. These SAP HANA worker VMs comprise a single, distributed SAP HANA system. In other words, an SAP HANA Scale-Out configuration connects several SAP HANA instances together into one large distributed SAP HANA database system.

The benefit of this configuration is that it is possible to grow the SAP HANA database size over time, and it is easy to add more vSphere hosts and SAP HANA worker VMs as needed.

SAP HANA Scale-Out certified appliances are available from 4- to 56-nodes with 2- to 8-CPU sockets, and between 128 GB and 3 TB of RAM. Currently, Scale-Out configurations are limited to SAP BW workloads and are generally supported by SAP for up to 16 SAP HANA worker systems. Business Suite workloads are not generally supported; for some scenarios (native and virtualized), however, pilot phases are in-process. For up-to-date details on this please refer to SAP Note 1825774. *Also support for more than 16 SAP HANA worker nodes are available upon request from SAP.*

In a physical Scale-Out configuration, a Standby host (working as a "hot-spare") is required to provide high availability. In a virtualized environment it is required to have only the compute resources available to support a failover and restart in case an SAP HANA VM fails. Unlike with a physical setup, spare resources can be used for other workloads, so they are not wasted.

Figure 4 shows such a configuration, where every ESXi™ host supports a single SAP HANA VM. The storage configuration is not shown here, and requires a shared TDI storage subsystem. To ensure enough resources are available in the event of a failure, hosts must be configured to provide spare resources, or resources used by non-critical VMs (shown in the figure as "ANY Workload") must be available to support the failover and restart of any failed SAP HANA VMs.

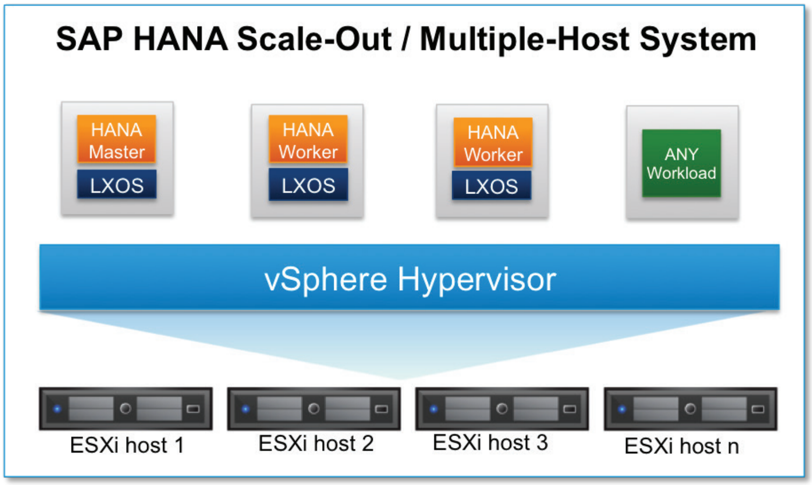


Figure 4. SAP HANA Multiple-Host System (Scale-Out) Configuration

When the configured ESXi hosts are able to provide enough resources, it is also possible to co-deploy another SAP HANA Scale-Out system on the same physical vSphere ESXi host cluster, or any other workload that an organization may want to run in virtualized mode. Figure 5 shows such a configuration, where two independent SAP HANA Systems (1 and 2) are running on several vSphere ESXi hosts that work together as two independent SAP HANA systems (SAP HANA System 1 and 2). When selecting SAP HANA system—or any other virtualized workload that has overlapping resource peaks and valleys—it is possible to optimally utilize all available HW resources of the vSphere cluster.

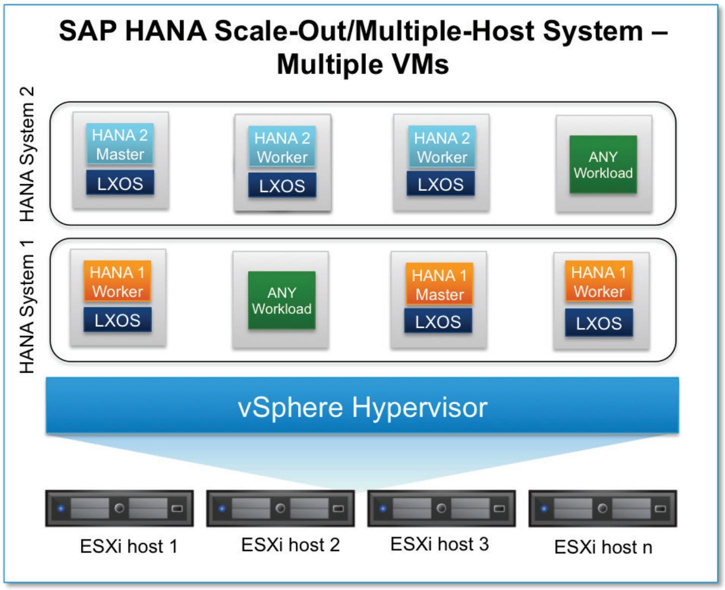


Figure 5. SAP HANA Multiple-Host System (Scale-Out) – Multi-VM Configuration

For detailed SAP HANA deployment information please visit the main SAP HANA webpage at: https://help.sap.com/hana_platform or review the "SAP HANA Master Guide and SAP HANA Server Installation and Upgrade Guide."

SAP HANA Appliance Model

The SAP HANA appliance is a flexible, multi-purpose, data source agnostic in-memory appliance. It combines SAP software components to run on selected and pre-configured Intel-based hardware configurations delivered by various SAP hardware partners such as Hitachi (HDS), VCE, Dell, Cisco, Lenovo, HP, NEC, Fujitsu and others. For a list of certified SAP HANA appliances and supported hardware components please visit the "SAP HANA Hardware Directory.[x]"

A VMware virtualized SAP HANA system can - but doesn't have to - come pre-configured as an SAP HANA appliance, such as those offered by Hitachi[xi] and other vendors. The minimum requirement is that:

- The OEM hardware be supported by SAP for SAP HANA workloads
- The server hardware corresponds to the bill of material (BOM) of a certified SAP HANA appliance or is listed as a supported SAP HANA server system
- The selected server is VMware-certified

In addition to the core server system, Scale-Out SAP HANA appliances must also include the required network components to interconnect multiple SAP HANA hosts; unlike Scale-Up configurations, a shared storage area network (SAN) or network-attached storage (NAS) storage system is used instead of local attached disks.

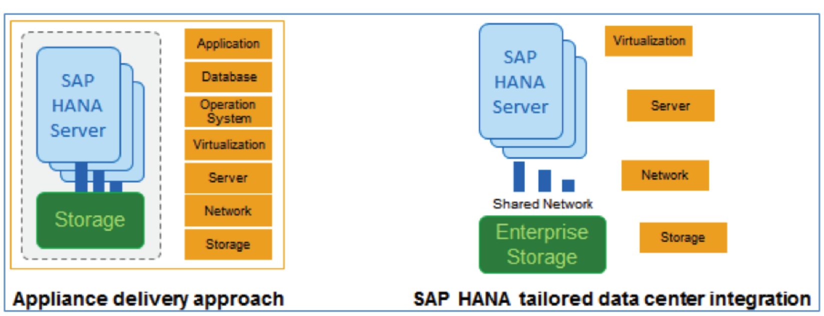
SAP HANA Tailored Datacenter Integration

The SAP HANA Tailored Datacenter Integration deployment option allows customized SAP HANA solutions, where a customer can choose any supported SAP HANA server from an approved vendor, along with any supported SAP enterprise storage. TDI allows also the usage of any hardware and infrastructure components that may already exist in an organization's data center that complies with SAP HANA hardware requirements.

The customer benefits in many ways when using VMware virtualization combined with TDI:

- Faster SAP HANA deployment
- More flexibility in deploying SAP HANA
- Right-sizing instead of appliance T-shirt sizing
- No vendor lock-in
- Easier operation
- Higher SLAs due to virtualization policy-driven server management and operations

The figures below describe the appliance delivery model and the TDI model for SAP HANA systems. The graphic for the appliance model shows all components in a box, and shows how these components are all pre-configured, tested and certified as an appliance; the TDI model shows the different components loosely "coupled" like server and storage to highlight that a customer can choose from any supported vendor, but is responsible for the installation and implementation of the overall system.



Picture © SAP SE

Figure 6. Appliance Delivery Model vs. SAP Tailored Datacenter Integration[xii]

In Scale-Out configurations it is required to use either a NAS or SAN shared storage for data persistence.

Currently, SAP has limited the number of active SAP HANA Scale-Out worker hosts to a maximum of 16 per SAP HANA TDI environment; this limit is also valid for virtualized systems. Customers that need more than 16 active worker hosts must contact the SAP HANA TDI Back Office to get support approval for larger SAP HANA Scale-Out systems. In some cases, "standby systems" do not count toward this limitation; for updated information please refer to the *SAP HANA TDI FAQ document*.[xiii]

Co-Deployment of Mixed Workloads in VMware vSphere Virtualized SAP HANA Scale-Out TDI Multi-VM Environments

The VMware vSphere TDI deployment option allows an organization to deploy several SAP HANA systems in production and non-production use cases depending on available compute resources; this can include non-SAP HANA workloads running on the same physical server. Co-deployment and consolidation of independent workloads running on different VMs is one of the primary use cases of virtualization, but must follow strict isolation and encapsulation of the memory and instructions a VM is using. The VMware document "*Security of the VMware vSphere Hypervisor*[xiv]" describes it this way:

> **"Isolation of CPU, memory, and I/O now is done at a hardware level, with the vSphere ESXi hypervisor managing how much of the hardware resources a virtual machine can use, similar to a choreographer or traffic officer. This part of the hypervisor is called the virtual machine monitor (VMM). With the ability to leverage these CPU extensions, the attack surface of the hypervisor shrinks considerably.**
>
> **Memory pages that are identical in two or more virtual SAP HANA machines are stored once in the host system's RAM, and each of the virtual machines has read-only access. Such shared pages are common, for example, if many virtual machines on the same host run the same OS. As soon as any one virtual machine attempts to modify a shared page, it gets its own private copy. Because shared memory pages are marked copy-on-write, it is impossible for one virtual machine to leak private information to another through this mechanism. Transparent page sharing is controlled by the VMkernel and VMM and cannot be compromised by virtual machines. Transparent page sharing can get disabled on a per-host or per-virtual machine basis."**

Figure 7 shows this concept, as discussed in the VMware document "Security of the VMware vSphere Hypervisor.[xiv]"

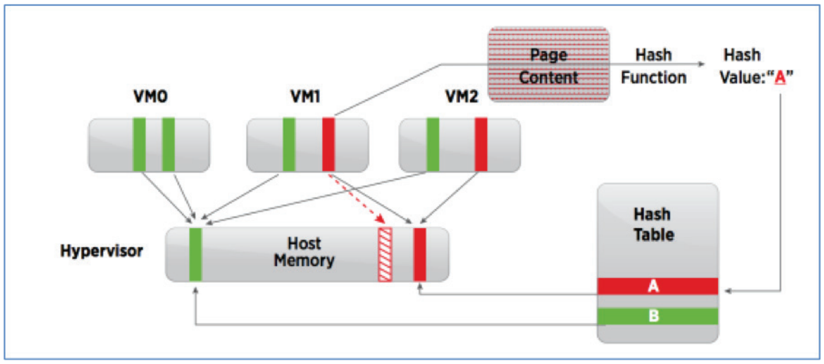


Figure 7. Transparent Page Sharing – Page-Content Hashing[xiv]

Using Transparent Page Sharing will reduce the amount of RAM of co-deployed similar VMs. The sharing of RAM pages used, for instance, by the OS and SAP kernel, can help reduce the overall consumed RAM on a vSphere host, and would allow the deployment of larger systems on a single vSphere host.

Note: When using this option please be aware of technical information in the document VMware KB 2080735,[xv] "*Security considerations and disallowing inter-Virtual Machine Transparent Page Sharing*."

In order for SAP HANA VMs to run in a secure and protected mode, instructions and memory must be isolated; this is fully provided and supported by VMware vSphere. Therefore, you can safely run as many concurrent SAP HANA VMs as your underlying hardware is able to support.

Besides CPU instruction and memory isolation, virtualized device drivers are also limited to instruction and memory elements of their VM, and are unable to access information from other VMs.

In a virtualized environment, compute resources are shared among all virtual machines, but because system resources can be managed, you can enable limits on virtual machines. SAP HANA systems that run in production must be configured to avoid resource over-commitments, or there is the risk of being impacted by resource constraints caused by co-deployment of other VMs. Using resource reservations, or—alternatively—high resource shares for SAP HANA VMs is a strict requirement for production SAP HANA VMs.

Co-deployments of SAP HANA VMs, and the operation of transactions within the VMs, is secure and reliable and protected by default by vSphere. Protecting and managing the network communication of an SAP HANA VM needs to get implemented as well to ensure that outside communication is protected against dangerous threads and misuse. vSphere provides the tools and solutions to secure any network communication of single or multiple SAP HANA VMs running on the same host, or several hosts. Comparatively, protecting physical running SAP HANA systems at the same level requires specific routers, switches, and firewalls and is very costly, complex, and inefficient.

Virtual network communication and security works the same as it does with any other physical OS or application that is natively installed on a server and connected to a network. The primary difference is that it is much easier to isolate and protect network traffic of SAP HANA VMs running on vSphere, by leveraging the VMware-provided software for virtual network and solutions.

Figure 8 shows how network communications can become limited and secured by leveraging the built-in hypervisor network capabilities, in a way that is similar to virtual switches.

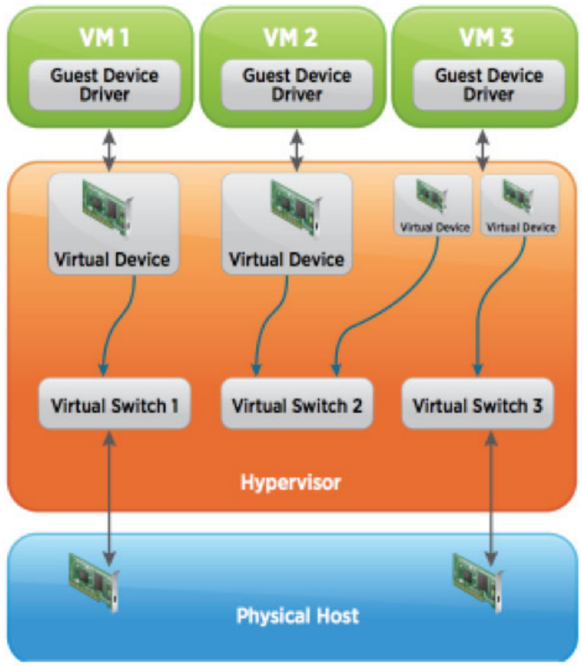


Figure 8. Network isolation of a Virtualized Server Environment[xiv]

Virtual machine network isolation, as outlined in figure 8, works like this:[xiv]

- Virtual Machine 1 does not share a virtual switch with any other virtual machines within the host. VM1 is therefore completely isolated from other virtual networks within the host, but would be able to communicate via Virtual Switch 1 to any network devices connected to the same network.
- If no physical network adapter is configured for Virtual Machine 2, then the virtual machine is completely isolated from any physical networks. In this example, the only access to a physical network is if Virtual Machine 3 acts as a router between Virtual Switch 2 and Virtual Switch 3.
- A virtual machine can span two or more virtual switches only if configured by the administrator. Virtual Machine 3 shows such a configuration.

Through the use of a virtualized network interface card (vNIC)–level firewall, a virtual machine can be isolated from other virtual machines, even on the same switch (using “Layer 2 isolation”).

Using NSX™—VMware’s network virtualization solution—in combination with vSphere, it is possible to attach network security services and policies to a VM with a policy-driven approach. This will allow instant micro-network segmentation and increased network and firewall security policies per VM. For more information on VMware NSX, please visit the NSX product webpage at: http://www.vmware.com/products/nsx/.

The consolidation, co-deployment or multi-tenant SAP HANA database installations on physical servers lack security and resource control; but a VMware-virtualized SAP HANA environment provides these attributes, and also allows secure, reliable and controllable co-deployments of multiple, independent SAP HANA systems or other workloads.

SAP HANA Tailored Datacenter Integration Support Process

The SAP HANA appliance delivery model is different from the SAP and OEM HW partner model where the hardware vendor validates and provides ongoing maintenance for the solution; in the SAP HANA TDI approach, it is the responsibility of the customer to negotiate with VMware, the OS, and hardware partner, to arrange individual support agreements for solution maintenance and ongoing support.[xvi]

Nevertheless, the actual support process in gaining support from SAP for SAP HANA issues is unchanged, and existing support contracts with the OS and hardware vendors, or VMware, can still be used.

Besides taking care of proper support contracts, the customer is also responsible for the end-solution validation and go-live testing. SAP and VMware provides several tools and services for SAP HANA TDI customers to help them with the end-solution validation, go-live testing, and ongoing support[xvi]:

- "The HW Configuration Check Tool for SAP HANA[xvii]" provides tests for customers to determine if the hardware configuration meets the minimum performance criteria required to run SAP HANA in production. This tool also allows ongoing checks and measurements to ensure required performance levels when additional SAP HANA VMs get deployed on the same hardware.
- SAP AGS provides go-live-checks and other support services for SAP HANA customers as part of their SAP support license. For details please review the SAP AGS IT Planning webpage for more details.
- VMware provides virtualized SAP HANA health and architecture review checks. For details please visit the VMware pages on SAP SCN at http://scn.sap.com/docs/DOC-60470
- Freedom of choice of support contracts allows organizations to harmonize SAP HANA support to other support processes and contracts an organization may already have in place for other business-critical applications, like existing SAP ERP or database applications.

Benefits of Leveraging the TDI Deployment Option for SAP HANA Scale-Out Configurations with VMware vSphere

By leveraging the TDI deployment option, the customer gains (regarding SAP[xii] above) the following benefits compared to the appliance delivery model.

- Reduced hardware and operation costs by reusing existing hardware components and operation processes
- Mitigate risk and optimize time-to-value by enabling existing IT management processes for SAP HANA implementation
- Increased flexibility in hardware vendor selection by leveraging the existing ecosystem

By combining the power of the SAP HANA database Scale-Out configurations—running on VMware vSphere—organizations can achieve the following additional benefits that go beyond what TDI already provides.[xviii]:

- Standardization of support contracts, IT processes, management and operations for organizations by leveraging existing virtualization best-practice strategies.
- Faster time-to-value, by using the deployment and operation options a virtualized SAP data center provides.
- Better service levels and more flexibility by using the advanced features vSphere provides, such as VMware vSphere vMotion®, VMware vSphere Distributed Resource Scheduler™ (DRS), VMware vSphere High Availability (HA), and Disaster Recovery.
- Lower risks by implementing a virtualized SAP HANA system that can get deployed on standard SAP HANA supported IT components. These can be changed, modified or quickly extended to react to changing business needs.
- Up to 99.9 percent out-of-the-box high availability, with the option to use SAP HANA Standby systems or third-party HA agents to increase this even further.
- Lower total cost of ownership (TCO), by reducing the capital expenditures (CAPEX) by up to 70 percent and operating expenditures (OPEX), by up to 56 percent, for SAP HANA production environments with full SAP and VMware support.
- Fully isolated and secured co-deployment of multi-tenant and mixed workloads in VMware vSphere virtualized SAP HANA TDI environments.

Virtualized SAP HANA Scale-Out TDI Reference Architecture

An SAP HANA Scale-Out configuration depends on external infrastructure components for storage and network, which interconnects the server systems hosting the SAP HANA instances that build an SAP HANA Scale-Out system.

These infrastructure components must meet the minimum requirements SAP has defined for SAP HANA Scale-Out TDI configurations, and must pass the tests as defined in the HW Configuration Check Tool for SAP HANA.[xvii]

This chapter describes in detail a VMware vSphere virtualized SAP HANA Scale-Out architecture and covers the current system maximums, the network configuration, and storage options like Fibre Channel-SAN or NFS attached shared storage systems.

The following documents describe SAP HANA TDI prerequisites and recommendations and will be referenced later in this guide.

- "SAP HANA Tailored Data Center Integration Frequently Asked Questions"
- "SAP HANA Tailored Data Center Integration Overview Presentation"
- "SAP HANA Storage Requirements"
- "SAP HANA Server Installation and Update Guide"
- "Best Practices and Recommendations for Scale-up Deployments of SAP HANA on VMware vSphere"

General Concepts and Overview

Figure 9 shows a conceptual overview of a virtualized SAP HANA Scale-Out system. As with any other SAP HANA Scale-Out system, it can be configured with several active and idle Standby hosts. The SAP HANA software allows a distributed installation to support even the largest SAP HANA databases, and can balance the load between different hosts.

Unlike native, physical server systems, an installed Scale-Out SAP HANA system is installed *inside* virtualized machines running on top of a VMware vSphere hypervisor; the hypervisor is installed on the physical server host systems instead, and server resources needed for the Standby host are not wasted and idle, as they are used for other workloads running in the VM.

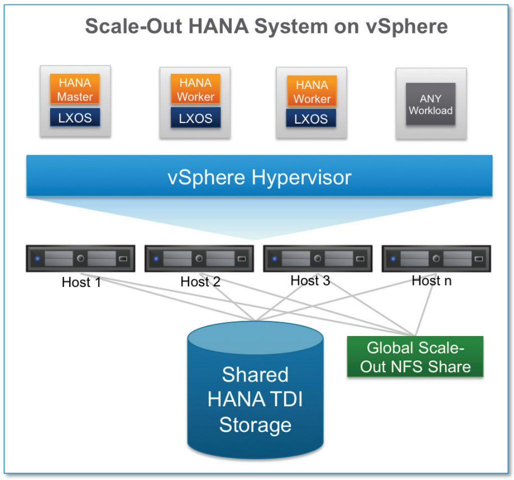


Figure 9. Virtualized SAP HANA Scale-Out Overview

Note: Certain required infrastructure components, like network switches, are not shown in figure.

Another distinction from SAP HANA installed on native physical servers is the ability to use advanced virtualization features like VMware vMotion, which allows the live migration of SAP HANA instances to other physical servers for outage-free server maintenance. More about the benefits of running SAP HANA on VMware vSphere can get found in the "Best Practices and Recommendations for Scale-Up Deployments of SAP HANA on VMware vSphere.[xix]"

The benefits described in the Scale-Up document are also applicable to virtualized Scale-Out configurations, as the system architectures for both deployment options—when based on a TDI configuration—are the same (except for the need of a shared file system with the Scale-Out deployment option). So it is possible to migrate a Scale-Up system very easily to a Scale-Out configuration to add more SAP HANA worker VMs to an existing Scale-Out SAP HANA system when additional database space is required.

Figure 10 shows how adding additional SAP HANA worker VMs would change an SAP HANA Scale-Up system to a Scale-Out configuration. The "plus" (+) sign in the figure indicates that this is only possible when hardware resources like storage, network and servers are added at the same time to support the expected additional higher workload.

In case an SAP HANA worker VM gets removed due to reduced resource needs, the now free HW resources can be used by any other workload ensuring resources are not wasted.

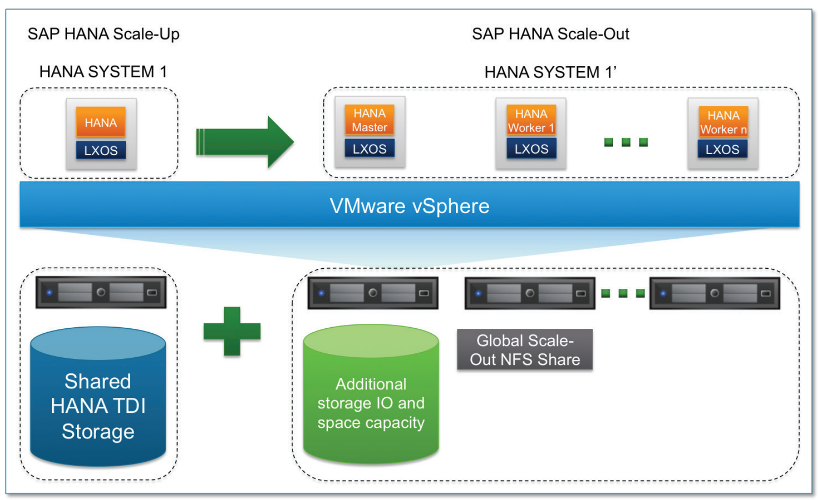


Figure 10. SAP HANA Scale-Up to Scale-Out Migration and Easy Extension

SAP Scale-Out on vSphere Storage Considerations

Before we discuss which storage options are supported for SAP HANA databases running on VMware vSphere, here are some important details on how virtual machines access different types of storage.

- When a virtual machine communicates with its virtual disk stored on a datastore, it issues SCSI commands. Because datastores can exist on various types of physical storage, the SCSI commands are encapsulated into other forms, which are determined by the protocol the ESXi host uses to connect to a storage device.
- ESXi supports Fibre Channel (FC), Internet SCSI (iSCSI), Fibre Channel over Ethernet (FCoE), and NFS protocols.
- Regardless of the type of storage device a host uses, the virtual disk always appears to the virtual machine as a mounted SCSI device.
- The virtual disk hides the physical storage layer from the virtual machine's operating system. This allows running operating systems in a very standardized way without any specific storage requirements, such as SAN host bus adapter (HBA) drivers or other settings.

Figure 11 shows the storage types vSphere uses, and how the storage is connected to the vSphere host –and how a VM accesses it. More details about the different storage types can be found in the *"vSphere storage guide.*[xx]*"* For SAP HANA systems, currently only NAS over NFS and fiber-based storage arrays formatted with Virtual Machine File System (VMFS) are supported.

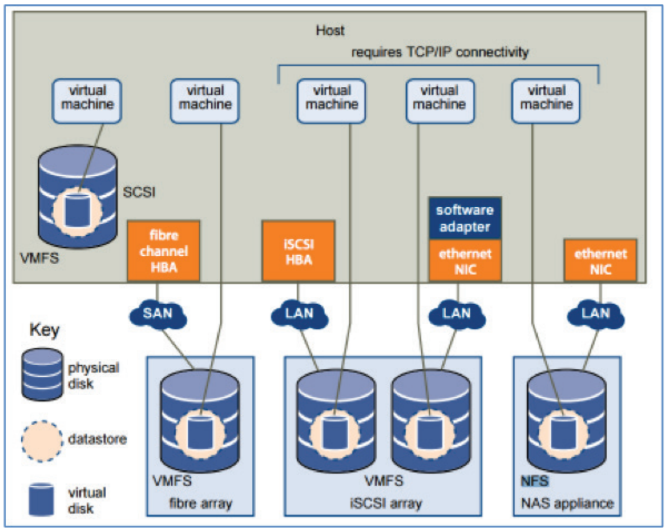


Figure 11. Virtual Machines Accessing Different Types of Storage[xx]

The figure shows how the physical disks/logical unit numbers (LUNs) get connected via a network card (NFS) or a Fibre Channel HBA (fiber-based SAN storage) on the vSphere host, and how the actual VMDK file of a VM is placed inside a datastore. Datastores on fiber arrays are formatted with VMware's own cluster file system VMFS; datastores on NAS storage use NFS as the cluster file system.

A VM with an NFS client installed can directly access any NFS-based NAS in the network. This feature is required when installing the SAP HANA Scale-Out deployment option on vSphere, because:

- SAP HANA server software in a Scale-Out configuration must be installed partly on a shared NFS based storage system, and
- The file system has to be mounted by all VMs in the SAP HANA system.

Figure 9, on page 17, shows this file system share as "Global Scale-Out NFS Share." The individual SAP HANA host/VM files (like SAP HANA database, or log files) can either be installed on NFS or SAN storage.

Note: The VMware-supported Scale-Out solution requires the installation of the SAP HANA shared file system on an NFS share. Other solutions—like Oracle Cluster File System, or IBM General Parallel File System—are not supported by VMware.

STORAGE TYPE	BOOT VM	VMOTION	DATASTORE	RDM	VM CLUSTER	VMWARE HA AND DRS
Local storage	Yes	No	VMFS	No	Yes*	Yes
Fibre Channel	Yes	Yes	VMFS	Yes	Yes*	Yes
iSCSI	Yes	Yes	VMFS	Yes	No*	Yes
NAS over NFS	Yes	Yes	NFS	No	No*	Yes

Table 2. vSphere Features Supported by Storage Type[xxi]

*In guest cluster software support, e.g., SUSE Linux Enterprise High Availability Extension[xxii]

Table 2 summarizes the vSphere features supported by the different storage types.

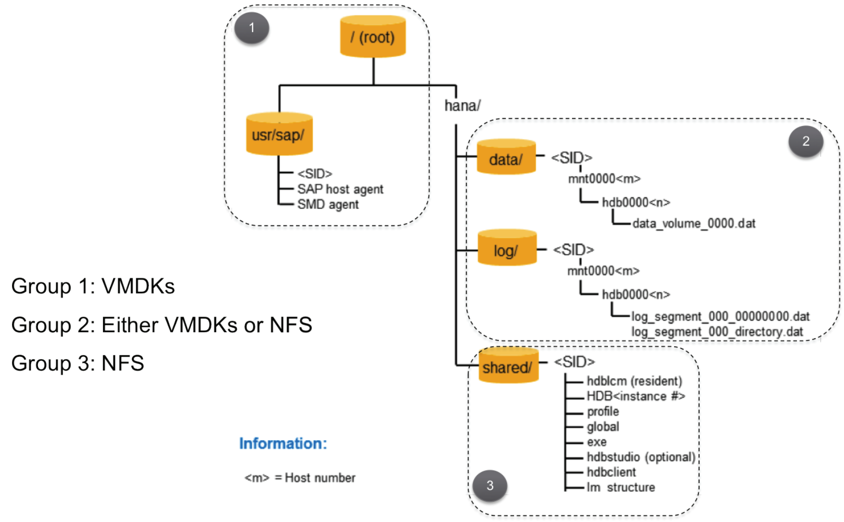
As previously stated, as of today only Fibre Channel and NAS over NFS-based storage are supported by VMware for SAP HANA Scale-Out configurations.

Figure 12 shows the SAP-recommended SAP HANA file system layout, which is the suggested layout when running SAP HANA virtualized. Grouping the file system layout in three groups helps you decide whether to use VMDK files (Groups 1 or 2), or an NFS mount-point as Group 3 to store the files and directories.

Group 1: For the first VMDK file that contains the OS and usr/sap files. This VMDK could be split into two independent VMDK files: one VMDK for the OS (vmdk-hx-os-1), and one for the SAP HANA binaries (vmdk-hx-sap-1). This provides additional flexibility by allowing an increase to the volume the SAP HANA /usr/sap is stored on. In most cases using only one VMDK for both OS and /usr/sap may be sufficient.

Group 2: This contains the non-shared DATA and LOG files, which can either be stored directly on NFS-mounted NAS storage, or inside VMDKs, that are stored in separate LUNs, when Fibre Channel-based storage is used.

Group 3: This contains the SAP HANA Scale-Out shared files and must be stored on NFS storage and is mounted on all connected SAP HANA VMs of a single SAP HANA Scale-Out system.



Picture © SAP SE[xxiii]

Figure 12. SAP-recommended SAP HANA File System Layout

The SAP HANA installation guide[xxiii] defines the following storage capabilities to be met in order to support the failover of a defective VM worker:

1. The Standby host has file access
2. The failed VM worker host no longer has access to write-to-file; this is called "fencing."

These requirements are met when using NFS- or SAN-based storage for VMware datastores.

The "SAP HANA Server Installation and Update Guide[xxiv]" describes the supported storage options for SAP HANA. Please review this guide for further storage information.

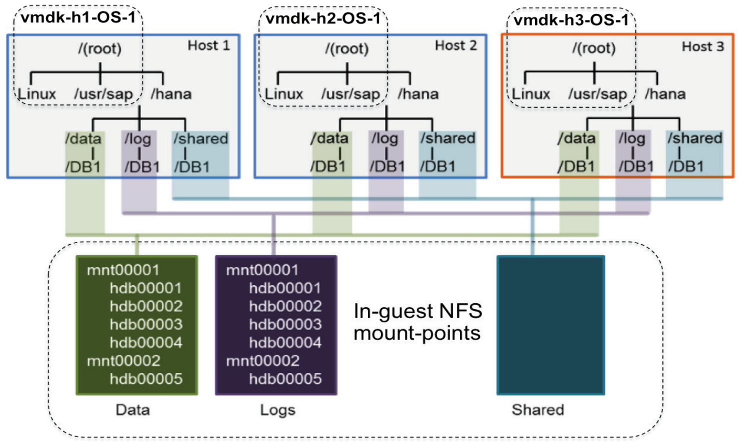
In a shared storage solution, with NFS, the storage subsystem provides dynamic mount-points for SAP HANA VMs. As the SAP installation guide describes, shared storage subsystems vary in their handling of fencing; therefore, it is the responsibility of the storage vendor to develop a corruption-safe failover solution when the SAP HANA Auto-Host failover option is used. In a vSphere environment, "fencing off" a failed SAP HANA worker VM is relatively easy, as the fencing can be done via standard vSphere commands as issued by the SAP HANA cluster manager/SAP HANA storage connector. Existing storage vendor scripts will need to be modified accordingly.

Fencing is required to ensure that after the Standby VM takes over the data and log files, the old worker VM no longer has access to these files. This is also required when a Standby host gets used with FC-SAN storage. In Linux cluster environments this solution is also called STONITH, which, interestingly, is short for "Shoot The Other Node In The Head." STONITH is common in Linux cluster solutions for maintaining node integrity.

Note: If only VMware HA is used, vSphere ensures fencing and consistency, and no SAP HANA Standby host is needed. The Scale-Out auto-host failover solution is discussed later in the "*SAP HANA Scale-Out on vSphere High Availability*" section of this guide.

NAS Over NFS Storage Option

Figure 13 shows the storage architecture used when running Scale-Out SAP HANA on vSphere with NFS storage.



Picture © SAP SE, source[xxiv]

Figure 13. Scale-Out SAP HANA on vSphere and NFS Storage

Besides the installation of the OS and SAP HANA binaries in one or two VMDK-based disks for OS and /usr/sap (shown in the figure as "vmdk-hx-OS-1") the installation of an SAP HANA Scale-Out configuration on vSphere with NFS based storage is exactly the same as when done with physical server systems (folder structure below / root//hana).

Please refer to the NFS storage vendor-specific SAP HANA Scale-Out configuration guidelines to configure NFS storage, and to determine how many dedicated 10 GbE network cards you will need to configure the ESXi host systems to support the I/O bandwidth requirements the vendor has sized for the configuration, and how the NFS vendor implements STONITH/fencing functionality.

The *"SAP HANA on NetApp FAS Systems with NFS[xxv]"* configuration guideline document is a good start for designing a NAS over NFS storage.

Note: A virtualized SAP HANA Scale-Out configuration running on NFS storage provides both HA solutions either VMware HA or the SAP HANA Standby host solution. The choice is up to the customer and if the NFS storage vendor supports an SAP HANA Scale-Out Standby solution with vSphere.

Fibre Channel-based SAN Storage Option

Also, when fiber-based storage is used there are two possible options for installing SAP HANA: either with a Standby host, or without a Standby host, leveraging only VMware HA. To support a Scale-Out configuration, both options require a cluster-wide NFS mount-point for the shared SAP HANA Scale-Out files.

The SAN installation with and without the SAP HANA Storage Connector application programming interface (API) will be discussed later in this section.

When installing a virtualized SAP HANA Scale-Out solution with fiber-based storage, the major difference -when compared to a physical installation - is that the SAP HANA system is not aware of the underlying storage hardware, and will only see SCSI devices. The FC-HBAs are installed and configured in the vSphere ESXi host system; the vSphere hypervisor ensures the VMs can use the storage.

As with NFS storage, it is recommended that best-practice guidelines of the storage vendor be followed. Storage sizing - and how many disks a LUN should support - is beyond the scope of this guide. As a default, a LUN-per-data and -log file VMDK should be used. OS and or /usr/sap should always be installed in an independent VMDK file.

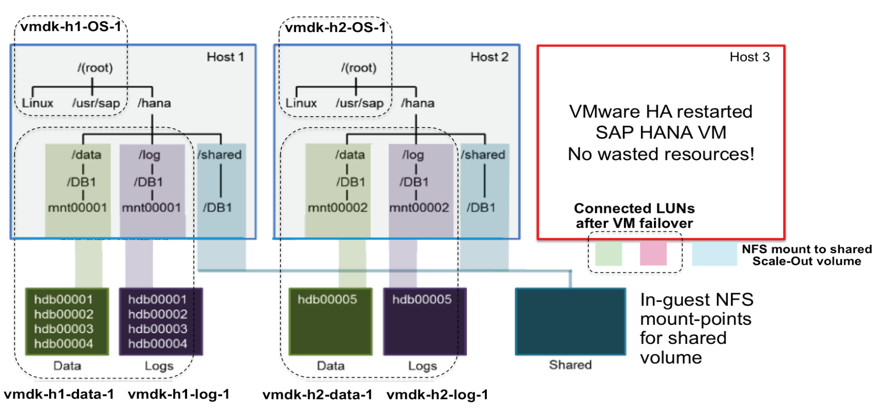
Please refer to vendor-specific documentation, like EMC's "*Storage configuration best practices for SAP HANA Tailored Data Center Integration on EMC VNX Series Unified Storage Systems,*[xxvi]" or any other appropriate vendor-specific SAP HANA TDI storage documentation based on your selected HW partner. The storage configuration—especially the sizing—of a physical and virtual SAP HANA installation is the same, and will require the same amount of HW resources in order to meet the requirements of SAP HANA KPIs for TDI storage.

Figure 12 above describes the SAP-recommended SAP HANA file system layout and shows three groups of volumes: Groups 1 and 2 require dedicated FC-based storage, and Group 3 contains the shared Scale-Out SAP HANA files. Figure 14 provides a more detailed view on the storage configuration, and which files and volumes get stored in a VMDK, or get mounted via NFS. This is shown in the figure in a block diagram connecting the required storage parts with the SAP HANA VMs.

In this configuration, VMware HA will protect against host, VM, or OS failures. In the case of such failures, the failed VM will get restarted on the same host, or moved and restarted on another host as shown in figure 14. The VM with a red border represents a restarted SAP HANA VM. No specific storage re-connecting or fencing is needed, as the failed SAP HANA VM simply gets used and restarted. This is the simplest configuration and does not require special testing or integration of the SAP HANA Storage Connector API or STONITH scripts. It may also be sufficient for most use cases, as 99.9 percent[xxvii] availability is only achievable by leveraging VMware HA.

By leveraging the SAP HANA Auto-Restart Feature with VMware HA, the overall SAP HANA system availability can increase to over 99.9 percent, as the SAP HANA watchdog will detect and solve SAP HANA application issues.

Combining SAP HANA Auto-Restart with VMware HA provides a very solid HA solution that meets most HA requirements.



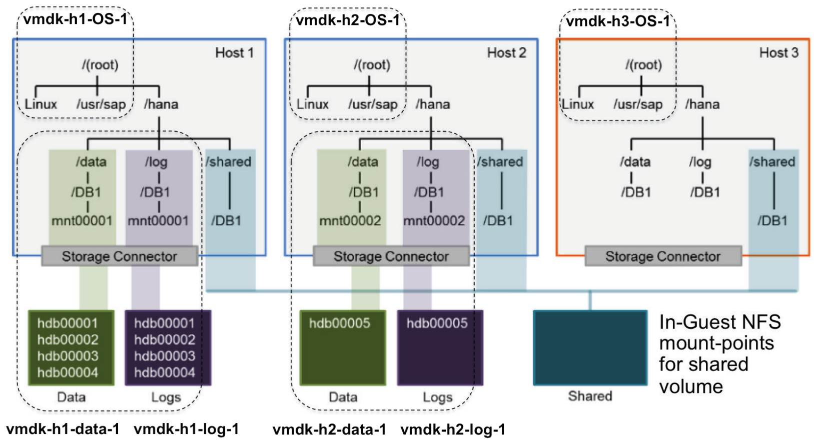
Picture © SAP SE, source[xxiv]

Figure 14. Fiber-based Storage Option Without SAP HANA Standby Node

When an organization decides that SAP HANA needs to achieve protection at the application level higher than 99.9 percent availability, using the SAP HANA Standby host/host auto-failover option—or another third-party SAP HANA HA solution—is required.

Figure 15 shows the storage block diagram when an SAP HANA Standby node/VM should get leveraged. The SAP HANA cluster manager will detect SAP worker node failures and will initiate the failover of the failed worker VM to the Standby VM.

The drawback is that an idle VM must run as an SAP HANA Standby VM, and will consume resources; this configuration is limited to specific storage vendors that provide support for this particular vSphere solution.



Picture © SAP SE, source[xxiv]

Figure 15. Fiber-based Storage Option with SAP HANA Standby Node

All disks of the SAP HANA Scale-Out System are VMDK disks; only SAP HANA shared volumes are NFS-mounted and accessible from all VMs. A Standby SAP HANA system is installed on a dedicated VM just as it would be on a physical system.

The SAP HANA cluster manager ensures—via the SAP HANA Storage Connector API and a STONITH script—that the failing VM no longer has write access to its SAP HANA log and data files. This ensures files cannot be corrupted. The STONITH script that needs to be implemented is basically the same for all vendors and storage solutions. The STONITH script is used to shut down a failed SAP HANA worker/master VM to ensure that the old, failed SAP HANA VM can no longer access the storage it previously owned. The script will send the shutdown VM command via public APIs to vCenter. Different SDKs, like the VMware vSphere Perl SDKs, are available to implement this.

Although the SAP HANA Standby host HA solution provides benefits like rolling OS or SAP HANA binary upgrades (or the possibility exists to react to SAP HANA application failures) it also creates some drawbacks when it comes down to VMware management and supported features, like cloning or storage vMotion of VMs that have Multi-Writer enabled VMDKs. More details on supported and unsupported features and actions can be found in the referenced "*VMware KB Article 1034165*."[xlii]

Note: Support of a Standby solution is currently only available from limited vendors, which have tested their SAP HANA Storage Connector API or STONITH setup with vSphere. A list of validated storage vendors that provide support in vSphere environments can be found at http://scn.sap.com/docs/DOC-60470

Table 3 provides a summary of an SAP HANA Scale-Out on vSphere storage configuration with required VMDK disks.

	OS*	/USR/SAP*	DATA	LOG	SHARED
FC-SAN	.VMDK	.VMDK	.VMDK	.VMDK	NFS mount point
NAS-NFS	.VMDK	.VMDK	NFS mount point	NFS mount point	NFS mount point

Table 3. SAP HANA Scale-Out on vSphere Storage VMDK File Configuration

*Depending on the preference, one or two .vmdk files for OS and /usr/sap

SAP Scale-Out on vSphere Network Considerations

vSphere offers so-called "standard and distributed" switch configurations. Both switches can be used when configuring an SAP HANA in a vSphere environment. A vSphere standard switch is very similar to a physical Ethernet switch: it can bridge traffic internally between virtual machines in the same VLAN, and link to external networks by connecting the physical network interface cards (NICs) of the hosts to uplink ports on the standard switch. Virtual machines have network adapters you connect to port groups on the standard switch. Every port group can use one or more physical NICs to handle their network traffic. If a port group does not have a physical NIC connected to it, virtual machines on the same port group can only communicate with each other, but not with the external network. Detailed information can be found in the vSphere networking guide.[xxxi]

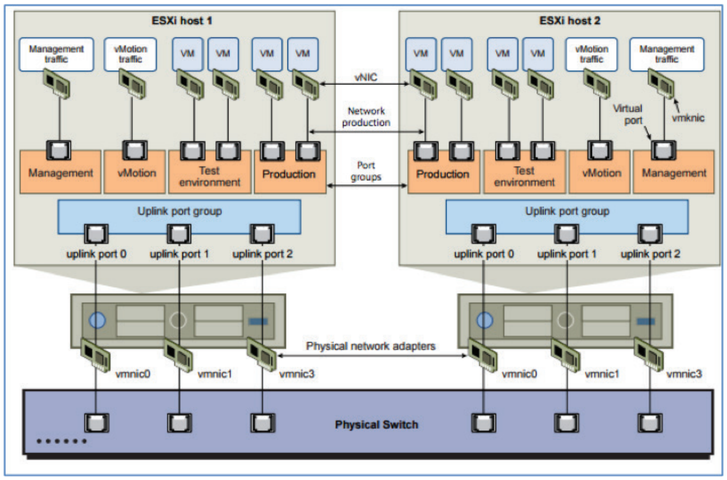


Figure 16. vSphere Standard Switch Architecture[xxxi]

Using a vSphere distributed switch in SAP HANA environments is recommended as it provides centralized management and monitoring of the networking configuration of all SAP HANA hosts that are associated with the switch. You can set up a distributed switch on a vCenter Server system, and its settings are propagated to all hosts associated with the switch. This ensures network consistency and error-free configuration. It also provides enhanced security and monitoring for virtual machines migrated via VMware vMotion through maintenance and migration of port runtime state and enhanced provisioning and traffic management capabilities through private VLAN support and bi-directional virtual machine rate-limiting.

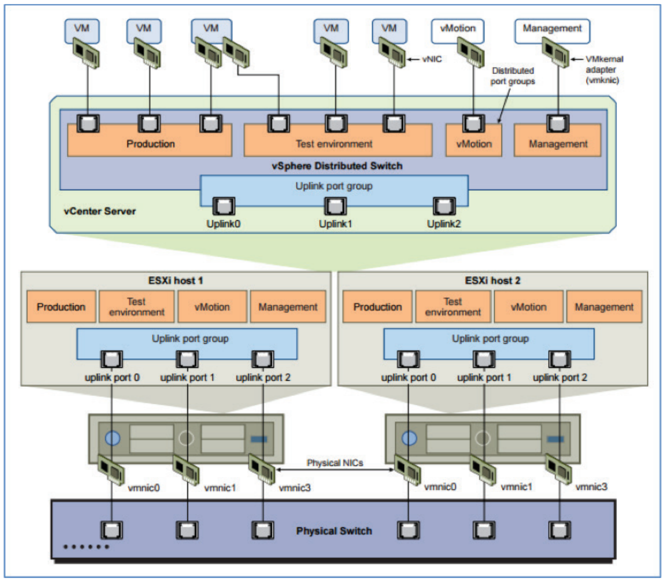


Figure 17. vSphere Distributed Switch Architecture

The virtual switch an organization decides to use for SAP HANA depends on the VMware vSphere license the organization has purchased. The minimum requirements for setting up SAP HANA on vSphere are fulfilled with a standard switch configuration.

Since SAP HANA Scale-Out configuration multiple-host systems work together as a single system, they must be connected, as recommended in the "SAP HANA Server Installation and Update Guide,[xxxii]" via dedicated 10 GbE network connection(s).

In addition to the SAP HANA internode communication network, management, HA, and a client network - with at least a 1 GbE on bandwidth - needs to be configured, and if advanced vSphere features like vMotion are required, a dedicated network is also required. If NAS instead of Fibre Channel-based storage is used to store the data and log files, an additional 10 GbE network must be installed.

Table 4 summarizes the network cards and configuration for SAP HANA on vSphere environments. vmNIC is the name for the physical NIC in the hosts, and is connected to the physical network switches that build the organization's network.

When a distributed or standard switch gets created, the suggested network labels should be used to allow easy detection of the networks the cards are connected to.

Depending on storage, additional 10 GbE NICs and switches may be required to satisfy the network bandwidth an SAP Scale-Out cluster requires. In table 4, the vmNIC (pNIC) row shows the physical network cards. In this example, four physical NICs are used; a 1 GbE NIC is used once, and 10 GbE NICs are used two or n-times, when NFS storage for data and log is used. The row with vNIC shows the number of virtual network cards needed to configure for the VM which will be visible inside the OS of the VM.

	MGT, HA AND EXTERNAL NETWORK	SAP HANA SHARED (NFS MOUNT) NETWORK	VMOTION NETWORK	SAP HANA INTER-NODE NETWORK	SAP HANA NFS STORAGE-N*
Network label	External	SAP HANA-NFS-shared	vMotion	SAP HANA-internal	SAP HANA-NFS-storage-n
Bandwidth	1 GbE	1 GbE	10 GbE	10 GbE	10 GbE
MTU size**	default	default	9,000	9,000	9,000
vmNIC (pNIC)	0		1	2	n
VLAN***	0	1			
vNIC	0	1		2	3
NAS storage	✓				
FC storage	✓				†

Table 4. SAP HANA Scale-Out vSphere Network Configuration of a Physical vSphere Host

* One dedicated 10 GbE NIC per SAP HANA VM that runs on a vSphere host. For example: One HANA VM per host, 1 x 10 GbE; two HANA VMs per host, 2 x 10 GbE NICs.

** Default MTU size as defined by IT provider

*** Depending on the customer needs, more VLANs are possible and may require separate application servers from server management network traffic.

If NFS is used to connect each SAP HANA VM running on a vSphere host to the storage subsystem, a dedicated 10 GbE link is required to fulfill the SAP HANA storage bandwidth requirements. Close network monitoring and planning is required before adding new SAP HANA worker nodes.

For more information on the network requirement of NAS-NFS storage based SAP HANA systems please review "SAP HANA on NetApp FAS Systems with NFS Configuration Guide.[xxxiii]"

Figure 18 shows the network configuration of an SAP HANA on vSphere and FC-SAN storage system using distributed switches. In this case, four dedicated physical networks are configured.

Note: It is recommended that each vSphere host have redundant network connections. It is also recommended to have Jumbo frames (with an MTU size of 9,000) for all networks but the client network. The client network MTU size is up to the IT standards of your organization.

Figure 18 shows the network diagram for an SAP HANA Scale-Out system running on vSphere by leveraging a virtual distributed switch. The SAP HANA-NFS-Storage-x network is optional and only needed when NAS-NFS storage is used for SAP HANA data and log files. If FC-SAN storage is used then this network is not required.

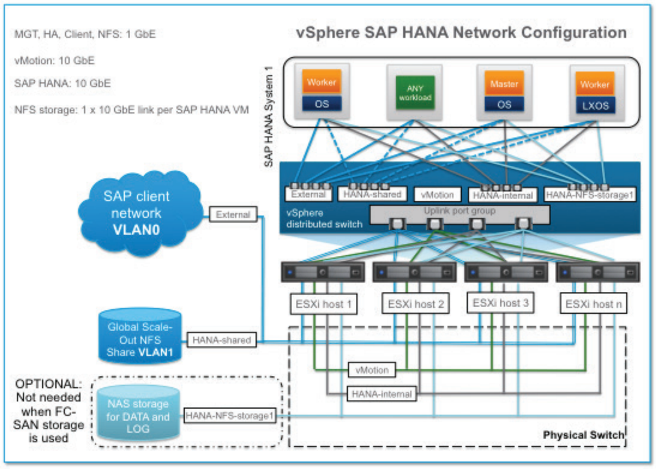


Figure 18. vSphere SAP HANA Scale-Out Network Configuration of a Virtualized Three-node SAP HANA Scale-Out System (three VMs) with a Dedicated Failover Server

SAP HANA Scale-Out on vSphere High Availability

SAP HANA offers several methods for HA and disaster recovery. There are auto failover, service restart options, backups, system replication and Standby host systems. In VMware virtualized environments additional VMware HA can be invoked to minimize unplanned downtime due to faults and disasters. Different Recovery Point-(RPO) and Recovery Time-Objectives (RTO) can be assigned to different faults and HA/DR solutions. SAP describes[xxxiv] the phases of high availability in figure 19.

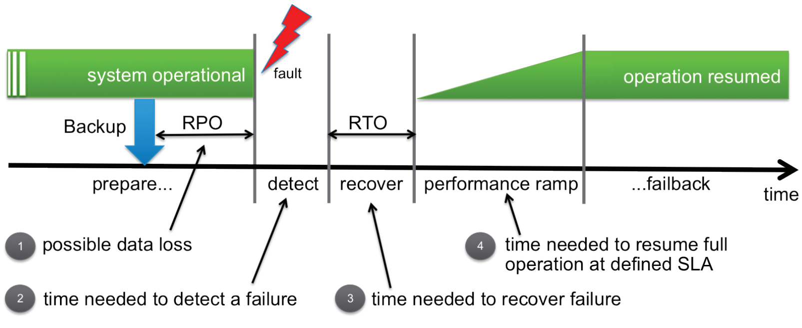


Figure 19. SAP HANA System High Availability Phases

1. Point 1, RPO, specifies the amount of possible data that can be lost due to a failure. It is the time between the last valid backup and/or last available SAP HANA savepoint and last saved transaction log file that is available for recovery. All changes made within this time may be lost and are not recoverable.
2. Point 2 shows the time needed to detect a failure and to start the recovery steps. This can be seconds, but also minutes/hours, without automation like with VMware HA, which minimizes this detection time.
3. Point 3 is the time needed to recover from a fault. Depending on the failure this may require restoring of a backup or a simple restart of the SAP HANA processes.
4. Point 4 shows the performance ramp, which describes the time needed for a system to run at the same SLA level as before the fault (data consistency and performance).

SAP HANA systems tend to be business-critical systems. Because of this, it is required to minimize downtime (RTO) and to avoid any possible data loss (RPO). This chapter focuses on SAP HANA data persistence, VMware HA and SAP HANA Standby server to protect against faults.

Backup and system replication is not covered in this guide, but they are important, and required in combination with the described HA solutions. For more information on SAP HANA high availability—especially for SAP HANA system backup and replication—please see the "SAP HANA High Availability Guide.[xxxiv]"

SAP HANA Data Persistence

SAP HANA is an in-memory database, storing all its data in volatile memory, which maintains data only for as long as the server is powered. To secure the memory content and all possible changed memory blocks, SAP HANA tracks all changes in the transaction log. When a transaction is committed, or when the transaction log becomes full, the log is written—using sequential writes in 1 MB blocks—to persistent storage on a storage device. Point 2 in figure 20 shows log writes. During normal operation, SAP HANA also automatically saves in full five minute increments (default value) the complete memory image representing the date-to-disk by creating a savepoint (see point 1). In the event of a failure, like a power outage or hardware failure (point 3 in the figure), SAP HANA restarts and returns to its last consistent state by replaying the log since the last savepoint (crash recovery).

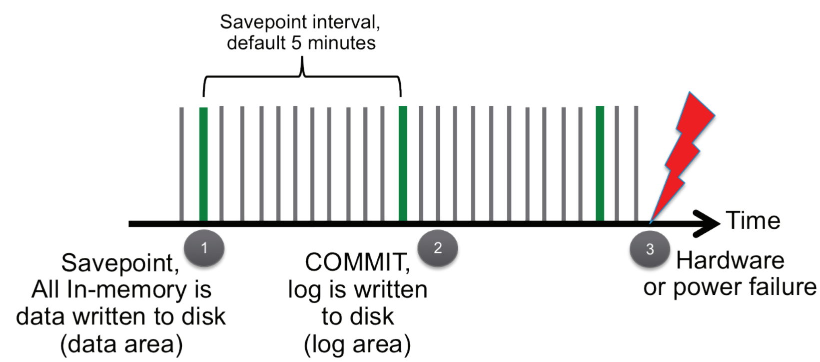


Figure 20. SAP HANA Data Persistency

Crash recovery steps are performed automatically when SAP HANA starts after an unexpected crash or failure. For the actual restart of an SAP HANA instance/VM a HA solution like the SAP Standby server or VMware HA solutions are required to restart the failed SAP HANA instance/VM on the same or another host. SAP HANA Standby server or VMware HA solutions restart either the complete SAP HANA VM or by mounting the log and data volumes of the failed instance to the Standby VM, and start loading the data into memory from the data volumes or last safe point, along with corresponding log files.

To minimize downtime it is recommended you enable the automated service restart feature of SAP HANA services. When *Autostart* is enabled, a VM - which got protected with VMware HA - will automatically start not only the OS, but also the SAP HANA database. Use the SAP HANA instance profile parameter "*Autostart = 1*" to enable autostart. For more details please refer to section *SAP HANA Auto-Restart Feature with VMware HA*.

As noted earlier, SAP HANA Standby or VMware HA solutions are able to restart the SAP HANA services or the complete SAP HANA VM, but these solutions are not able to protect against file corruption, logical or site failures, or failures affecting the SAP HANA storage subsystem. To protect an SAP HANA system against such failures and situations, backup and system replication solutions should be implemented to protect the data.

For available options and more details please refer to:

- "*The SAP HANA High Availability Guide*[xxxv]"
- "*The SAP HANA Backup/Recovery Overview presentation*[xxxvi]"
- "*Specific SAP HANA business continuity solutions available from different vendors like EMC*[xxxvii] *or HDS*"

VMware HA-protected SAP HANA Systems

VMware provides vSphere products with built-in and optional high availability and disaster recovery solutions to protect a virtualized SAP HANA system at all levels.

The power behind VMware's HA/DR solutions are how they are layered to protect against failures at every level of the data center, from individual components, like NIC or HBA card teaming, all the way up to the entire site. With vSphere replication, the solutions provide protection against both planned and unplanned downtime. Figure 21 shows the different solutions to protect against component-level up to complete site failures.

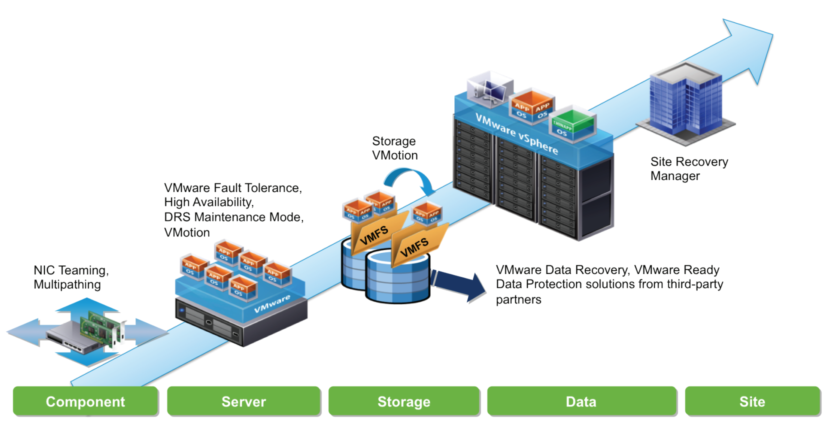


Figure 21. VMware HA/DR Solutions - Protection at Every Level

Many of the key features of virtualization, such as encapsulation and hardware independence, already offer inherent protections. From there, additional protections throughout the vSphere platform are provided to ensure organizations can meet their availability requirements and provide the following features:

- Protection against hardware failures
- Planned maintenance with zero downtime
- Protection against unplanned downtime and disasters

This chapter focuses on VMware HA, which is the easiest way to protect an SAP HANA system against unplanned downtime caused by OS or HW crashes.

VMware vSphere High Availability[xxxviii] delivers the availability required by most applications running in virtual machines, independent of the operating system or applications running on them. VMware HA provides uniform, cost-effective failover protection against hardware and operating system outages within your virtualized IT environment. HA can:

- Monitor VMware vSphere hosts and virtual machines to detect hardware and guest operating system failures, and restarting the VM on the same or another host in the vSphere cluster.
- Restart virtual machines, without any dependencies on the applications running on the VM, or on other vSphere hosts in the cluster without manual intervention when a server outage is detected.
- Reduce SAP HANA downtime by automatically restarting virtual machines upon detection of an operating system failure.

Using VMware HA to protect an SAP HANA system is the easiest and most cost efficient way to protect a virtualized SAP HANA system against OS and HW failures, without the dependency on any external components like DNS servers or features like the SAP HANA Storage Connector API provides. In the event of a failover the whole SAP HANA VM gets restarted either on the same or another host in the vSphere cluster, when the HW of the initial host is defective. Since all virtualized disks, like OS, data and log VMDKs and in-guest mounted NFS volumes get failed over to a new ESXi host as well, no specific storage tasks or cluster solutions are required.

By leveraging VM restart polices and affinity and anti-affinity rules it is possible to provide HA protection for independent SAP and non-SAP HANA VMs on a single cluster. The only requirement is that the failover host has enough resources left to start the failed VM, and that SAP HANA services are configured to start automatically after an OS reboot.

In an SAP HANA production environment, resource commitments are required and must be enforced to ensure optimal performance. In the case of an SAP HANA VM HW event the VM will get restarted on a host with enough resources left. If a VM is already running on the failover/Standby ESXi host, then beside failing over the VM, the non-critical VMs that run on the Standby ESXi host will get shutdown or moved to another host first to free up the needed resources for the failed SAP HANA VM. Another method is to configure the non-critical VM without resource commitments to purposely allow overcommitting of resources. In the event of a failover the SAP HANA VM will request all needed resources from the ESXi host, and the non-critical VM will only get un-claimed resources remaining on the ESXi host.

Figure 22 shows two failure situations where VMware HA can react:

- Event 1 shows a HW defect, affecting two VMs running on a failed host.
- Event 2 affects only the VM itself.

While event 1 has occurred, two arrows show that “SAP HANA 1 worker” VM gets moved to the host that runs beside SAP HANA Master 2 “ANY workload” VM. “SAP HANA 2 worker” gets moved to host number 4 and will restart the SAP HANA worker 2 VM on this host.

Event 2 shows a VM/OS failure. This VM would get moved to host number 4, as on this host an ”ANY workload” VM runs. By using affinity and anti-affinity rules—and the possibility to shutdown/reconfigure a no-resource commitment for the “ANY workload” VM—it is possible to provide a HA environment that does not waste any resources.

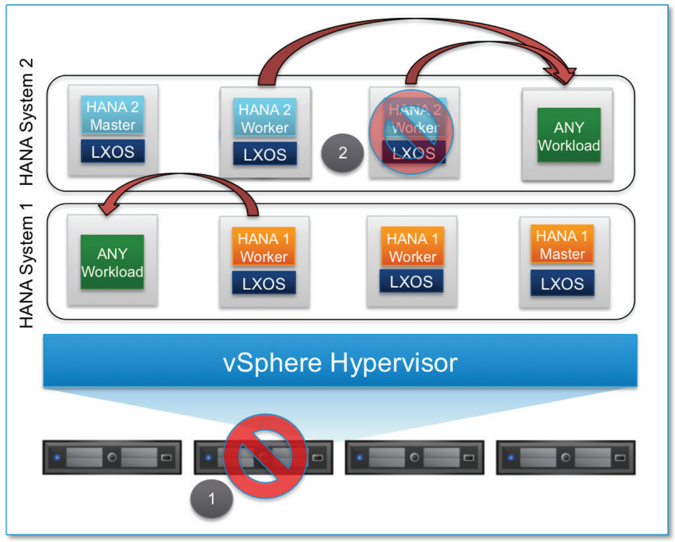


Figure 22. VMware HA-protected SAP HANA Systems[xxxix]

During the failover it is not possible to access the tables maintained by the affected SAP HANA worker/master VM. Depending on the last written savepoint and log, SAP HANA will initiate a crash recovery after the SAP HANA processes have started. After the crash recovery the SAP HANA system is accessible and usable as before the fault (RPO=0).

Note: Logical or file corruption failures are not recoverable.

SAP HANA Auto-Restart Feature with VMware HA

SAP HANA provides a Service Auto-Restart/watchdog function to automatically detect and restart stopped HANA processes.

All configured SAP HANA services (e.g., Index Server, Name Server, etc.) will be restarted by the SAP HANA Service Auto-Restart function, which automatically detects failures, whether a software failure or an intentionally stopped SAP HANA processes by an administrator. Once detected these stopped SAP HANA processes will restart automatically. The in-memory loaded SAP HANA data not affected by a process failure will not get reloaded. After the processes has restarted, only the affected SAP HANA service process data gets reloaded into memory, and SAP HANA resumes its function.

VMware HA combined with the SAP HANA Auto-Restart Feature provides a very solid HA solution to minimize failures due to power, HW, OS or SAP HANA application faults. Combining these two solutions will provide a high availability solution higher than 99.9 percent!

Figure 23 shows in a view of the protected components of virtualized SAP HANA system when the VMware HA and SAP HANA Auto-Restart features are invoked.

Additionally, SAP HANA application failures—like process crashes or OS failures—may also get protected via third-party solutions, such as that offered by Symantec[xxix]. A benefit of such a third-party solution is that it would be also able to detect OS failures, and would be able to leverage the vSphere HA API to initiate a VM restart or failover, whereas the SAP HANA watchdog is only monitoring and restarting the SAP HANA processes. Please check out the third-party vendor product pages of support availability for SAP HANA Scale-Out configurations.

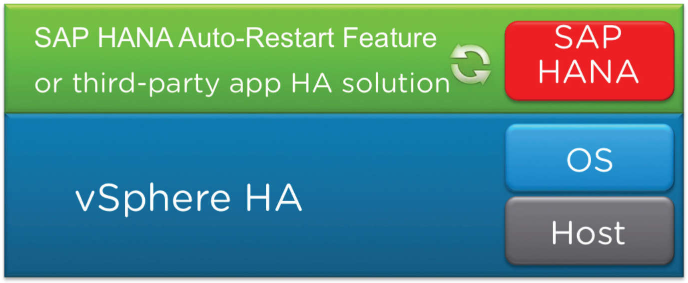


Figure 23. VMware HA-protected Components

Using VMware HA and SAP HANA Auto-Restart Feature provides:

- The ability of vSphere HA to protect against OS and host failures without specific configuration requirements!
- The SAP HANA Auto-Restart Feature can protect against SAP HANA application process failures to improve application uptime to more than 99.9 percent.[xl]
- Third-party VMware virtualized SAP HANA solutions can extend the monitoring beyond SAP HANA processes.

SAP HANA Systems Protected with Host Auto-Failover to a Standby VM

Protecting a VMware virtualized SAP HANA system with VMware HA already provides very high system availability. Customers who want to use the Host Auto-Failover solution to failover to a Standby VM can do this with selected storage vendors that offer SAP HANA Storage Connector API or STONITH implementations validated for the vSphere virtualized environment.

A benefit of using a Standby VM to protect an SAP HANA system is having an operational VM with a pre-installed OS and SAP HANA configuration that is available and online. This VM can be used for specific maintenance tasks only available with a Standby VM like OS or SAP HANA application binary patching, where the active SAP HANA system configuration can be moved over to the Standby node that already has a patched OS and SAP HANA binaries available. In a VMware HA environment this is not possible, as only one VM per SAP HANA worker exists at any time. Nevertheless, before performing such maintenance tasks it is possible to take a snapshot of the VM to undo/do, in the case of a binary problem, and any patches of configuration changes are restored within seconds. After the maintenance, such snapshots must be deleted, otherwise the I/O performance of this VM will be negatively impacted. An organization must decide which solution is best; virtualizing SAP HANA with VMware vSphere *will not* limit the customer's freedom of choice.

Host Auto-Failover

When an SAP HANA has an host auto-failover—also called SAP HANA Standby host solution—the active (worker) VMs will get monitored:

1. The SAP HANA name server, which also acts as the cluster manager monitoring all SAP HANA worker VMs regularly.
2. If a failing SAP HANA instance gets detected, the cluster manager decides which Standby system gets used for the failover and ensures that the Standby VM takes over the role of the failing VM and starts its database instance using the persisted data and log files of the failed instance. The SAP HANA storage connector implementation of the storage vendor will get executed which ensures that the failed SAP HANA instance no longer has write access to the log and data files stored in VMDK disks, to avoid data corruption. This is done via a STONITH script and the Storage Connector API.
3. Once the SAP HANA storage connector returns successfully, the disks will be accessible only by the Standby VM.
4. The data and log files will get loaded into memory while SAP HANA automatically starts the required crash recovery steps to ensure data consistency for this instance.

For details about the storage configuration and requirements, please check out the storage chapter on page 19 and figure 25.

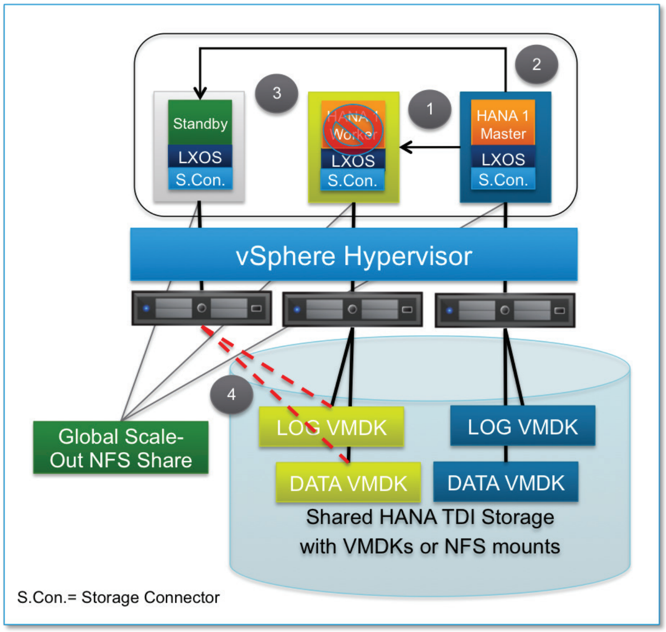


Figure 24. SAP HANA Host Auto-Failover Steps

Host Grouping[xli]

Before deploying an SAP HANA Scale-Out system it is necessary to consider the number of required SAP HANA worker VMs, and how these systems should be protected. If SAP HANA Standby VMs are used, host grouping will be required for anything greater than eight host cluster systems. This is also required when more than one Standby VM needs to be available to protect the SAP HANA system to provide a higher redundancy level.

Host grouping is a logical grouping of SAP HANA worker VMs that get protected by a Standby VM. It does not impact performance or load distribution among the VMs. As mentioned, host grouping is required when more than eight vSphere ESXi hosts are required.

"*VMware KB Article 1034165*[xlii]" provides more detail on how to enable the multi-writer option to allow VMs to share VMDKs.

If nothing gets configured, then all SAP HANA worker VMs belong to the "default" group. Figure 25 shows the default host group of a four-host ESXi vSphere cluster with one Standby VM. With the default group a cluster with up to eight physical ESXi host systems with FC-based storage is supported.

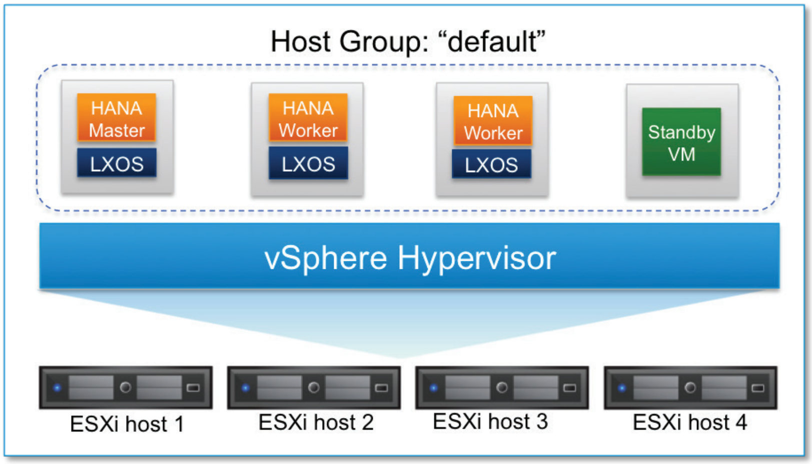


Figure 25. SAP HANA Host Group "Default"

For larger SAP HANA systems that exceed eight vSphere ESXi hosts, a host group for each vSphere HA cluster (running on FC-based storage) is required as outlined in figure 26. Again, SAP supports generally only Scale-Out cluster sizes of 16 hosts/VMs. Support for larger systems can get requested from SAP.

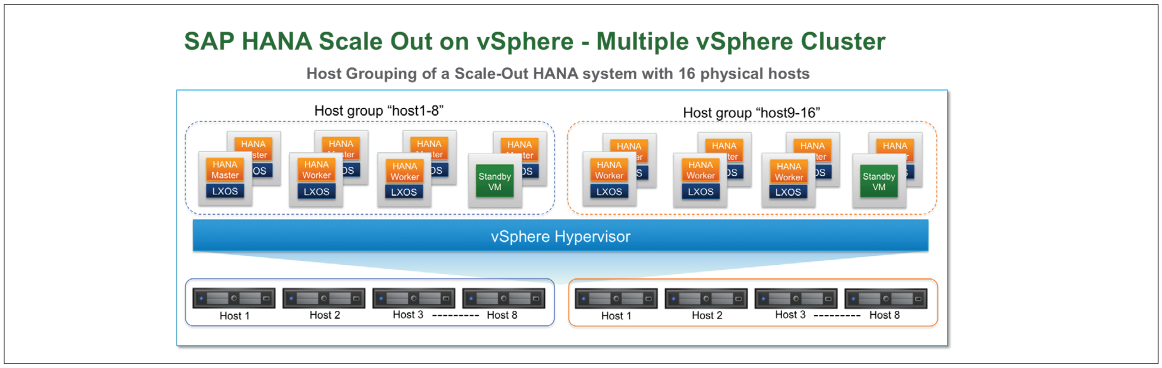


Figure 26. Host Grouping of a Scale-Out HANA System With 16 Physical Hosts

Each host group contains a dedicated Standby VM, which ensures that the VMs of this group can only fail only over to this VM and not to a VM hosted on another vSphere HA cluster. This ensures that the Standby VM will have access to the log and data files of a failed SAP HANA VM.

With host groups it is also possible to define Standby VMs with different sizes to support an SAP HANA system with mixed machine resources.[xliii]

Note: The support of the SAP HANA Standby host solution depends on the storage vendor-specific SAP HANA Storage Connector implementation, and customization to adapt the actual storage configuration at the customer site is often required. NFS v3 protocol-based NAS storage requires the STONITH script to ensure exclusive write access of the Standby VM to the failed over disks. NFS v4 protocol-based NAS storage can leverage NFS v4 file locking. A STONITH script with this storage solution may not be required.

Since the SAP HANA Storage Connector implementation varies from vendor to vendor, it is required to reference the storage vendor’s specific documentation to ensure proper installation in a vSphere environment. A list of validated storage vendors that provide support in vSphere environments is listed on the VMware SAP SCN webpage: http://scn.sap.com/docs/DOC-60470.

SAP HANA Disaster Recovery Solutions with VMware vSphere

After describing the HA solutions that recover from local failures—like component or OS failures—the SAP HANA disaster recovery solutions that replicate the data from the primary data center to VMs in a secondary data center is supported with vSphere. SAP HANA System Replication provides a very robust solution to replicate the SAP HANA database content to a secondary disaster site; this storage-based system replication can be used as well.

When using SAP HANA System Replication the same number of SAP HANA VMs must exist at the disaster recovery site.[lxxxiii] These VMs must be configured and installed similar to a natively-running SAP HANA system with System Replication enabled.

SAP HANA System Replication provides different modes for system replication:

- Synchronous
- Synchronous in-memory
- Asynchronous

Depending on requirements, the disaster recovery VMs can consume more or less resources on the disaster recovery VMware vSphere cluster. For instance, selecting the synchronous in-memory mode will consume the same amount of RAM as the primary systems. This mode is only required if the customer requests the shortest recovery time. In most customer scenarios, using synchronous data replication should be sufficient. SAP states that by only replicating the data, around 10 percent of the system resources are required,[lxxxiii] allowing up to 90 percent of the resources to continue being used by other systems, like test or QA systems.

Figure 27 shows the scenario of a co-deployed configuration where all SAP HANA Scale-Out VMs get replicated to a disaster recovery site per VM level.

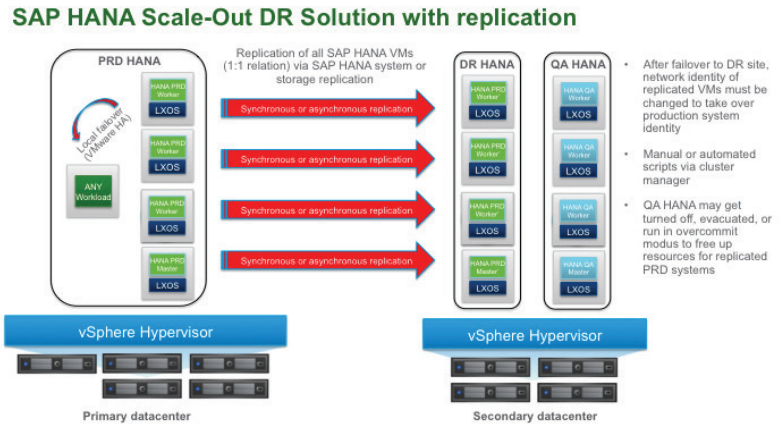


Figure 27. 1:1 VM Replication via SAP HANA System or Storage Replication

In this scenario resource over commitments get used to allow the co-deployment of such an environment. By using resource pools and resource shares it is possible to provide the needed resources to the disaster recovery SAP HANA Scale-Out VMs. The co-deployed system, with fewer resource shares, will experience performance degradation after the disaster recovery systems get used after a site failover. Evacuate these VMs to other available vSphere systems to free up all resources for the disaster recovery SAP HANA VMs; this is another option rather than running both systems in parallel—with resource limitations—on the same platform.

System Replication via storage—or the SAP HANA replication solution—needs additional steps after a site failover has taken place to switch the network identity (IP redirect) of the replicated systems from the disaster recovery configuration to the production configuration. This can be done manually or via automated tools like HP Service Guard, SUSE Cluster, SAP Landscape Virtualization Manager (SAP LVM) or other cluster managers. The configuration of such a solution in a virtualized environment is similar to the configuration of natively running systems. Please get in contact with your storage vendor to discuss a cluster manager solution supported by their storage solution.

Summary SAP HANA HA and Disaster Recovery Solutions with VMware vSphere

Table 5 provides an overview of the different available SAP HANA HA and Disaster Recovery Solutions when running on VMware vSphere.

The complexity and cost level can increase with the certain solutions, starting with VMware HA combined with the SAP HANA Auto-Restart Feature, up to a completely replicated SAP HANA system to a second DR site. Tools like the VMware vCenter™ Site Recovery Manager™ (SRM) can help reduce the complexity of a DR solution.

HIGH AVAILABILITY SOLUTION	VMWARE HIGH AVAILABILITY	VMWARE HA + SAP HANA AUTO-RESTART FEATURE	SAP HANA HOST AUTO-FAILOVER (STANDBY VM)	SAP HANA SYSTEM REPLICATION	STORAGE SYSTEM REPLICATION	STORAGE SYSTEM REPLICATION + VMWARE SRM
Scenario description	VMware standard HA solution. VMware HA to restart or failover of VM to another host in case of a detected OS/ HW failure.	Standard VMware HA combined with SAP HANA Auto-Restart watchdog running inside a VM to monitor SAP HANA application status and triggers an SAP HANA process restart. OS and HW failures will get handled by VMware HA.	SAP HANA Standby VM to take automatically over the role of another SAP HANA VM, in case of a detected failure.	Data replication between primary and secondary sites by leveraging SAP HANA System Replication functionality; there is no automated failover process. Third-party cluster solutions can be used to automate site failover.	Data replication between primary and secondary site by leveraging storage system replication functionality. No automated failover process. Third-party cluster solution can get used to automate site failover.	VMware SRM to automate the storage replicated SAP HANA system VMs and all related servers to another site.
Operating system failures	Yes	Yes	Yes	Yes	Yes	Yes
Hardware failures	Yes	Yes	Yes	Yes	Yes	Yes
Application failures	No	Yes	Yes	Yes	Yes	Yes
IP redirect/DNS update	Not necessary	Not necessary	Not necessary	Yes, manual or third-party software, such as cluster manager solution, is needed	Not necessary as complete VMs with all data disks get replicated	Not necessary as complete VMs with all data disks get replicated
RTO	Medium (crash recovery of DB)	Medium (crash recovery of DB)	Short to Medium (only if in-memory data loading)	Shortest to medium RTO, depending on IP redirect solution	Long and manual restart (crash recovery of DB)	Medium (crash recovery of DB) Automated restart
RPO	0	0	0	0 (synchrony) > 0 (asynchrony)	0 (synchrony) > 0 (asynchrony)	0 (synchrony) > 0 (asynchrony
Performance ramp	Minutes to hours (bootstrap+loading)	Minutes to hours (bootstrap+loading)	Minutes to hours (bootstrap+loading)	Seconds to minutes if synchronous replication into memory is selected, depending on the selected modus and IP redirect solution.	Hours to days (bootstrap+loading)	Hours (bootstrap+loading)
Cost	**Included**	**Included**	**Included**	**High**	**High**	**High**
Complexity	**Low**	**Low**	**Medium**	**High**	**High**	**High**

Table 5. SAP HANA Scale-Out HA and Disaster Recovery Solutions with VMware vSphere

VMware vSphere Site Recovery Manager

VMware site recovery manager can help to reduce the complexity of a system replication DR solution, by automating the complex disaster recovery steps on any level.

Site Recovery Manager is designed for disaster recovery of a complete site or data center failure. Site Recovery Manager supports both uni- and bi-directional failover. In addition, it also supports "shared recovery site," allowing organizations to failover multiple protected sites into a single, shared recovery site. This site could, for instance, also be a VMware vCloud® Air™[xliv] provided cloud data center, VMware's own cloud service offering.

The key elements that make up a Site Recovery Manager deployment for SAP are:

- **Site Recovery Manager** – Designed for virtual-to-virtual disaster recovery. Site Recovery Manager requires a vCenter Server management server at each site; these two vCenter Servers are independent, each managing their own site; Site Recovery Manager makes them aware of the virtual machines they will need to recover if a disaster occurs.
- The Site Recovery Manager service takes care of managing, updating, and executing disaster recovery plans. Site Recovery Manager is managed via a vCenter Server plug-in.
- Site Recovery Manager relies on storage vendors' array-based replication Fibre Channel, or NFS storage, that supports replication at the block level to replicate SAP HANA data and log files to the DR site. Site Recovery Manager communicates with the replication via storage replication adapters that the storage vendor offers and has been certified for Site Recovery Manager.
- vSphere Replication has no such restrictions on use of storage-type or adapters and can be used for non-performance-critical or static VMs, like infrastructure services or SAP application servers with RPO time objectives of 15 minutes or longer.

Figure 28 shows an example SRM- and storage-protected SAP landscape. The VMs running on the primary site contain all needed infrastructure and SAP components like LDAP, SAP HANA database and SAP application servers, as in an SAP Business Suite implementation. The VMs can be replicated, depending on the RPO needs via vSphere, SAP HANA or storage replication. vSphere replication can be used with VMs that tolerate an RPO time of 15 minutes or longer.

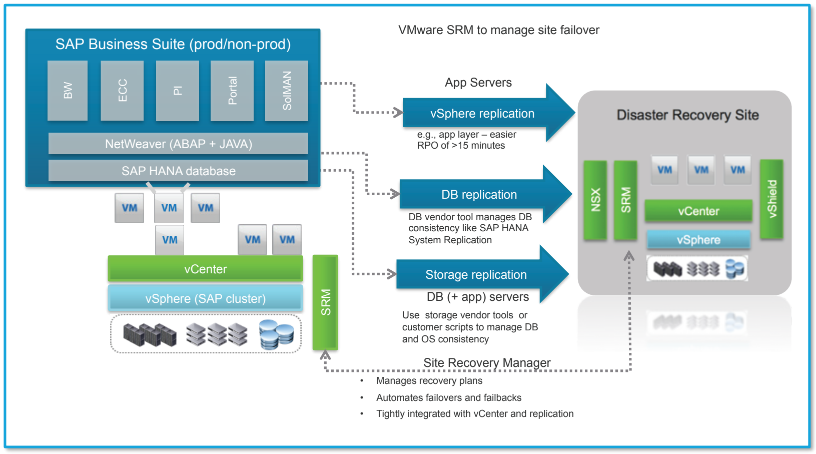


Figure 28. VMware SRM-protected SAP Landscape

The VMware vCenter Site Recovery Manager datasheet details the benefits of using VMware SRM for managing the DR process for SAP landscapes. Here is a summary[xlv]:

- Lower the cost of disaster recovery by up to 50[xlv] percent
- Application-agnostic protection eliminates the need for app-specific point solutions
- Support for vSphere Replication and array-based replication offers choices and options for synchronous replication with zero data loss
- Centralized management of recovery plans directly from the vSphere Web Client replaces manual runbooks
- Self-service, policy based provisioning via VMware vRealize™ Automation™ automates protection
- Frequent, non-disruptive testing of recovery plans ensures highly predictable recovery objectives
- Automated orchestration of site failover and failback with a single-click reduces RTOs with reliability
- Planned migration workflows enable disaster avoidance and data center mobility

Virtualized SAP HANA Business Warehouse Scale-Out Sizing

Sizing SAP HANA for a virtual environment is the same as sizing SAP HANA for a natively running physical environment, except for a few differences. These differences are due to the size limitations of virtual machines, virtual CPU to physical CPU mappings, and the slight memory and CPU overhead cost of virtualization. The basic process is to size the SAP HANA database as for a natively running physical one, and then make any needed adjustments to account for the virtualization overhead.

This chapter focuses on these differences and the sizing specifics of virtualized SAP HANA Scale-Out systems running BW workload. SAP Business Suite is not yet generally released[xlvi] on SAP HANA Scale-Out configurations. While these scenarios are rolled out for ramp-up, customers have to contact SAP directly for sizing guidelines and recommendations.

For SAP HANA basic sizing guidelines please refer to the specific SAP webpages on SCN for the most current information[xlvii], and also see SAP note 1514966 – "SAP HANA 1.0: Sizing SAP In-Memory Database[xlviii]" and SAP note 1736976 – "Sizing Report for BW on HANA[lxxx]"

General Considerations

Right-sizing of the server systems running SAP HANA BW workload is critical, and is based on accurate calculation of the required amount of memory, CPU processing power, storage and network IO.

SAP HANA BW workload sizing is driven by memory sizing, and consists of sizing for static data and dynamic objects created during runtime (data load and query execution).

SAP provides different methods to determine the required RAM of an SAP HANA database.

1. SAP HANA sizing using the Quick Sizer tool[lxxxi]
2. SAP HANA sizing using database-specific scripts
3. SAP HANA sizing using the ABAP report

Methods 2 and 3 will provide the most precise results, but require a running system and the rights to execute these scripts. The SAP HANA Quick Sizer tool, or using an assumed compression ratio for the database conversion to SAP HANA, can also be used to determine the memory needs of an SAP HANA system. If no SAP database system exists, then the Quick Sizer tool and using a compression ratio are the only methods used to determine the required RAM for a new SAP HANA system. The SAP HANA source database compression ratios vary depending on the source data and its distribution in the database tables.

As described in SAP note 1514966 - SAP HANA 1.0: Sizing SAP In-Memory Database, memory sizing of SAP HANA is determined by the amount of static and dynamic data that must get stored in memory. The usable amount of server RAM for the in-memory SAP HANA database is therefore only around 50 percent.

The static memory need of SAP HANA is determined by the amount of data tables stored in memory. Additional to this, dynamic RAM is required to store objects that are created dynamically when new data is loaded, or if queries are executed. Because of this, SAP is generally recommending to reserve as much memory for dynamic as for static objects; for sizing reasons, the determined static RAM needed is multiplied by two, or, the usable RAM provided by an SAP HANA server system is divided by two.

Focusing on SAP HANA Scale-Out configurations, the first question to ask is when to use Scale-Out instead of Scale-Up configurations.

SAP note 1855041 – "Sizing Recommendation for Master Node in BW-on-HANA[xlix]" provides general guidelines for systems larger than approximately 20 TB (the size of the original uncompressed BW system). An SAP HANA Scale-Out configuration based on 1 TB RAM per SAP HANA nodes is recommended to ensure stable operation. SAP also states that these are very general guidelines and they must be carefully assessed individually based on your system usage pattern and other operational aspects.

SAP is currently limiting the number of active SAP HANA Scale-Out worker hosts to a maximum of 16 per SAP HANA TDI environment. Customers with more than 16 active worker hosts must contact the SAP HANA TDI Back Office to get support approval for larger SAP HANA Scale-Out systems. SAP HANA Standby systems do not count against this limit. For more and current information please refer to the "*SAP HANA TDI FAQ document.*[l]"

If a customer wants to use a Standby server/VM to protect their SAP HANA system, a Scale-Out Standby host configuration may be considered for Scale-Up and –Out configurations. If a Standby server is used for a Scale-Up system, then a "two host Scale-Out" configuration can be implemented. If a customer wants to do this, the CA process as described in SAP note 1781986[xlvi] must be followed.

Currently, the largest SAP HANA Scale-Out system can be as large as eight CPU sockets with 1 or 2 TB of RAM. Since virtual SAP HANA Scale-Out sizing follows this SAP requirement, even the largest virtualized SAP HANA system can span eight CPU sockets with a maximum of 2 TB of RAM per VM.

Limiting the supported SAP configuration to these server systems for BW workloads also maximizes limits of the amount of RAM per CPU socket.

- An Intel Xeon processor E7-4870 CPU (Westemere) can have up to 128 GB RAM per CPU socket
- An Intel Xeon processor E7-4890 v2 CPU (Ivy-Bridge) can have up to 256 GB per CPU socket
- An Intel Xeon processor E7-8880 v3 CPU (Haswell) can have up to 384 GB per CPU socket

These sizing limitations may change over time. For an updated certified SAP HANA Scale-Out server list please visit the SAP HANA hardware directory.[li]

Figure 29 shows a screen shot from certified systems sorted by Scale-Out appliance time and RAM size.

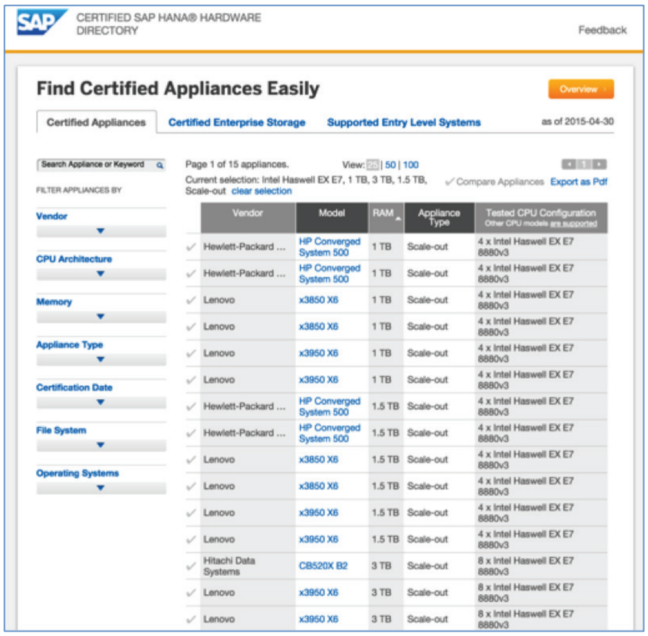


Figure 29. Certified SAP HANA Hardware Directory (screen shot as of May 2015)[lviii]

Maximum Virtual Machine Size

When configuring a virtual machine for the usage with SAP HANA, the primary sizing factors are CPU and memory. In SAP HANA Scale-Out configurations, network and storage is equally important as CPU and memory, as an SAP HANA Scale-Out system depends very strongly on network communication and a shared storage subsystem.

The vSphere 5.5[lii] and 6.0[liii] configuration maximums are documented in detail in the referenced documents. Table 6 provides a summary of some SAP HANA relevant figures and compares the vSphere 5.5 and 6.0 versions.

When starting an SAP HANA sizing process these maximums need to be taken into account, as they define the upper limits of a VM. The lower limits are defined by how hardware resources like CPU resources Non-Uniform Memory Access (NUMA) node can get shared in productive scenarios. In productive, performance critical scenarios, the sharing of a NUMA node between different SAP HANA systems are not supported.

	VSPHERE 5.5	VSPHERE 6.0*
Hosts per cluster	32	64
VMs per cluster	4,000	8,000
NUMA nodes per host	16	
Logical CPUs per host	320	480
RAM per host	4 TB	6 TB, 12 TB with specific OEMs
VMs per host	512	2,048
Combination of 10 GB and 1 GB Ethernet Ports	8 x 10 GB and 4 x 1 GB ports (12 network ports in total)	16 x 10 GB and 4 x 1 GB ports (20 network ports in total)
Physical HBAs ports	16 ports	
Virtual CPUs per VM	64	128
Virtual RAM per VM	1 TB	4 TB
Virtual NICs per VM	10	
Virtual SCSI adapters per VM	4	
Virtual disk size (VMDK)	62 TB	
Virtual disks per VM	60	

Table 6. vSphere 5.5 and 6.0 SAP HANA Relevant Configuration Maximums

*vSphere 6 is in SAP HANA validation

The CPU socket-to-RAM limitations for BWoH workloads show that the maximum RAM sizes of natively running systems are the same as for VM-based systems. As outlined in the table, there are no real limitations for Scale-Out configurations when a 4-socket server is used (amount of CPU threads and RAM). The largest VM that spans four CPU sockets (NUMA nodes) can have as many as 64 vCPUs and 1 TB of RAM with vSphere 5.5, and 128 vCPUs and 1.5 TB RAM with vSphere 6.

Memory Overhead and Configuration

Virtual machines have full access and full use of the RAM assigned to the virtual machine when it is created. However, some memory is needed to run the ESXi hypervisor and to manage the virtual memory assigned to the virtual machines. Many vSphere customers never notice this memory overhead due to the many memory optimization techniques that have been implemented.

Production SAP HANA virtual machines have to be configured with full memory reservation. This feature does not allow enablement of memory optimization techniques. For this reason, it is necessary to leave a small amount of unassigned RAM on the host to account for the memory needed by hypervisor and virtual machine memory overhead.

A very conservative estimate for the amount of RAM that needs to be unassigned to the SAP HANA virtual machines for overhead is **three to four percent**. For example, on a vSphere ESXi host system with 1 TB of physical RAM, approximately 30 to 40 GB would need to be left unassigned to the virtual machines. On this host, a single large virtual machine with around 980 GB—or two virtual machines with 490 GB each—could be created to leave enough unassigned memory for the hypervisor and virtual machine memory overhead. This memory overhead only needs to be considered when a single—or several—VMs will use all the memory resources of a server. If the server has more memory installed than the sum of all the VMs are able to consume, then the three percent overhead does not need to be considered.

When configuring memory reservations, watch the vSphere GUI carefully regarding the suggested memory settings. The maximum available memory that can be selected is the complete RAM of a server minus the RAM needed for other VMs and the hypervisor. When configuring an SAP HANA VM that spans all NUMA nodes, this is not an issue and the maximum usable value is available; but when configuring smaller VMs that might span only two CPU sockets (NUMA nodes), then remote memory could be used, and this can negatively impact memory performance. In this case, the amount of vCPUs of such a VM should be configured to consume all available CPU resources of the CPU sockets/NUMA nodes connected to the needed RAM. Production-level environments do not support consumption of remote memory. The amount of RAM of an SAP HANA VM that is supported for production workloads is a multiple of the available RAM per NUMA node.

Figure 30 shows local and remote memory access and the negative effect on memory performance (memory latency and bandwidth) of a VM that has an unbalanced memory configuration.

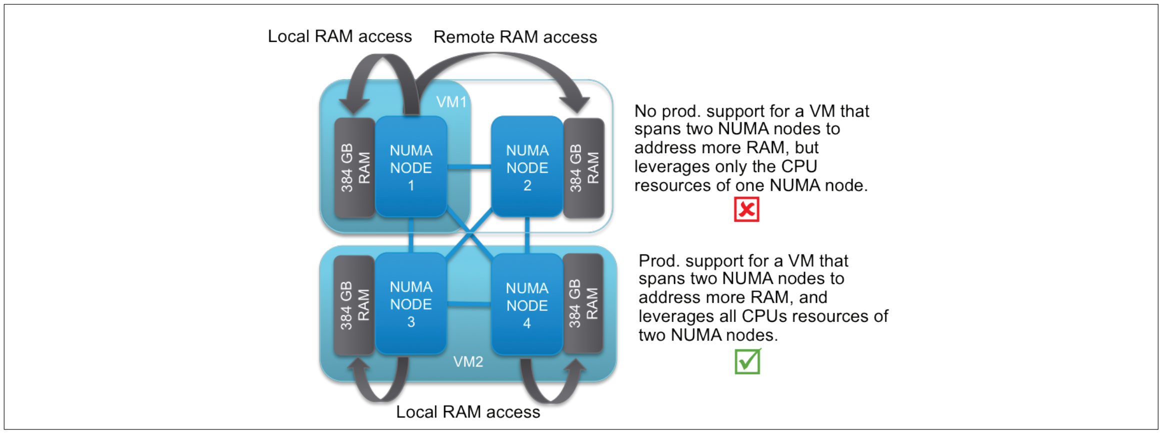


Figure 30. Example Maximum RAM Setting for VM

Figure 31 shows how to configure memory reservations for SAP HANA production-level VMs. The “reserve all guest memory” check box must be selected to ensure that the configured RAM is used exclusively by this VM.

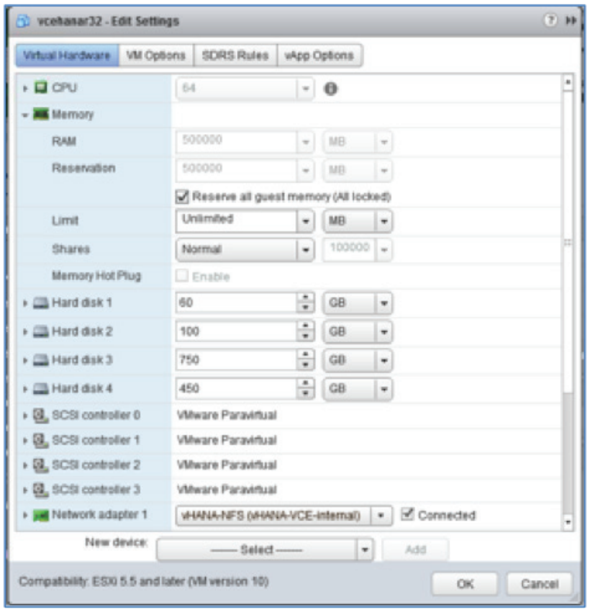


Figure 31. Setting Memory Reservation via vSphere Web GUI

Note that this memory overhead is visible to the memory reservations configured for SAP HANA virtual machines. In certain environments—like development and test environments where resource over-commitment is allowed—this memory overhead does not need to be calculated, and reservations do not need to get configured.

CPU, Hyperthreading and NUMA Node Considerations

CPU sizing for BWoH workloads is determined by the amount of the memory the BWoH system requires. SAP has defined with its hardware partners the maximum supported RAM-per-CPU socket (NUMA node) allowed for BWoH workload systems.

This ratio can be specified as a CPU socket-to-RAM ratio, and ensures that enough CPU resources are available for BW workloads. Taking into consideration the NUMA node architecture when configuring a VM helps leverage local memory instead of having to force an SAP HANA process to access the memory of a remote NUMA node. Just as RAM module sizes and CPU power levels change over time, their ratio will change over time as well. See changed ratios of Intel Xeon processor E7-4870 CPU (Westemere) and Intel Xeon processor E7-4890 v2 CPU (Ivy-Bridge) based CPUs as shown on page 47.

Figure 32 provides a view of the virtual HW an SAP HANA system can consume when running on an Intel Xeon processor E7-4870 CPU (Westemere). Or, as shown in figure 33, the same data can be seen with an Intel Xeon processor E7-4890 v2 CPU (Ivy-Bridge) based host system.

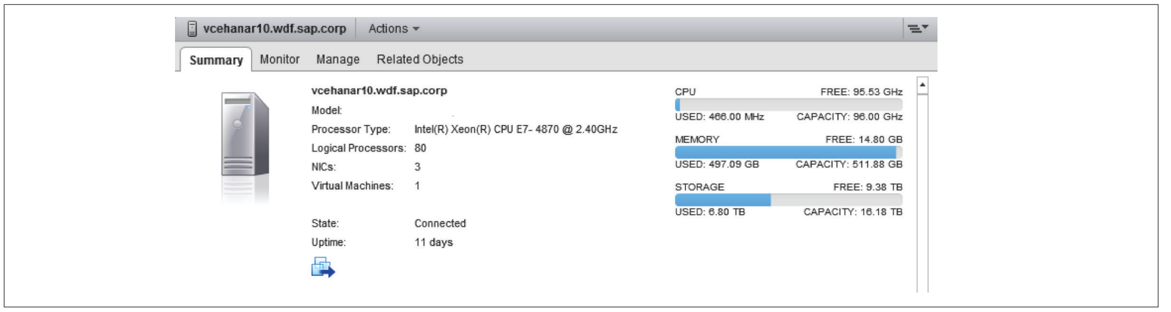


Figure 32. Virtualized SAP HANA 4-socket Intel Xeon Processor E7-4870 CPU (Westemere) System

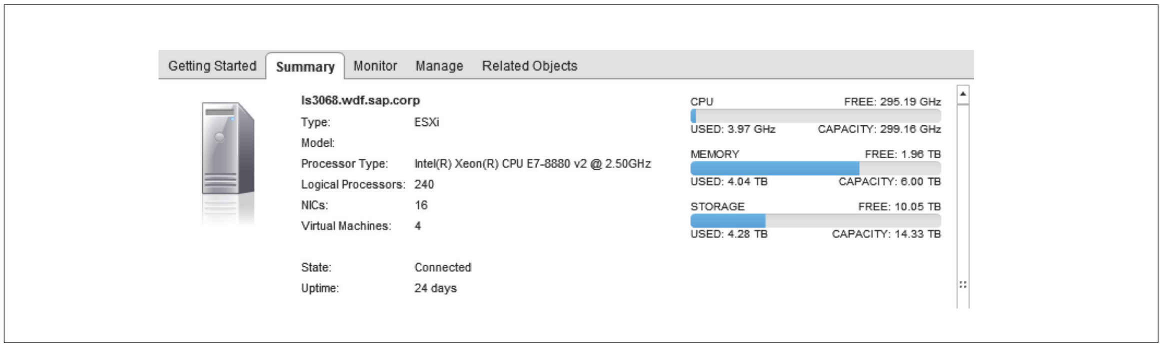


Figure 33. Virtualized SAP HANA 8-socket Intel Xeon Processor E7-4890 v2 CPU (Ivy-Bridge) System[liv]

CPU and Hyperthreading

Table 7 shows CPU figures that are important for SAP HANA and the SAP-defined CPU socket-to-RAM ratios. When designing a system it is important to keep these HW resource CPU socket-to-RAM ratios in mind to ensure optimal SAP HANA performance; this is true on both bare metal as well as virtualized on vSphere. Business Suite workloads are sized differently as BWoH workloads; this guide does not feature information on Suite on HANA workloads.

	INTEL E7 (WESTEMERE)	INTEL E7V2 (IVY-BRIDGE)	INTEL E7V3 (HASWELL)
CPU type	E7-4870[lv]	E7-4890V2[lvi]	E7-8890V3[lvii]
CPU frequency	2.4 GHz	2.8 GHz	2.5 GHz
CPU cores	10	15	18
Virtual HW capacity per CPU	24,000 MHz	42,000 MHz	45,000 MHz
Threads per CPU	20	30	36
QPI speed	6.4 gigatransfers per second (GT/s)	8 GT/s	9.6 GT/s
Third-level cache	30 MB	37.5 MB	45 MB
Memory type	DDR3	DDR3	DDR3 and DDR4
Average SAPS performance per CPU core	1,732	2,230	2,300
Max. CPU socket BWoH supports	8	8	8
BWoH CPU socket-to-RAM ratio	128 GB	256 GB	384 GB
BWoH CPU core-to-RAM ratio	12.8 GB	17.07 GB	21.33 GB

Table 7. BWoH CPU Socket-to-RAM Ratios

The Intel CPUs used for SAP HANA support hyperthreading. Hyperthreading allows a physical core to have two execution threads. These threads share the resources of the same physical core. This does not mean they double performance, rather, this feature provides up to a 10 to 20 percent increase in performance. The general guidance is to enable hyperthreading on the host and to consume hyperthreads in the guests by doubling the vCPUs to align them to the threads of the CPUs being used. Please keep the vCPU maximums of the vSphere versions in mind when configuring VMs that have more vCPUs configured than the NUMA node has installed.

A vCPU that is created for a virtual machine is scheduled to run on a logical thread. A vCPU maps directly to a logical thread on a physical core. The default behavior for vSphere is to schedule vCPUs on separate physical CPU cores, using a single execution thread on each physical core first. Once all of the physical cores have a vCPU scheduled and running, vSphere will begin to schedule vCPUs to use the second logical execution thread on each physical core.

From a sizing perspective, vCPUs are equivalent to physical cores without hyperthreading enabled in terms of performance. In order to gain up to an additional 10 to 20 percent of performance improvement with hyperthreading enabled, the number of vCPUs needs to be doubled. In other words, a physical SAP HANA server system with 64 physical cores, provides hyperthreading with 128 threads enabled. This results in 10 to 20 percent higher performance than a virtual SAP HANA server system (virtual machine) with 64 vCPUs and 64 execution threads, all mapped to physical cores of the example server with its 64 CPU cores or 128 threads. The VM configuration parameter **Numa.PreferHT=1**[lix] will ensure that the ESXi NUMA load balancer assigns a vCPU to a local NUMA node. However, the number of total vCPUs configured for a VM must be less or equal to the number of threads of NUMA node.

The VMware "*Guest and HA Application Monitoring Developer's Guide*[lx]" provides some information on how to determine if a VM was sized correctly by leveraging the vm.cpu.contention.cpu metric, which is described in this guide. As stated, a contention value of < 5 percent is normal "undercommit" operating behavior, representing minor hypervisor overheads. A contention value > 50 percent is "severe overcommit" and indicates CPU resource starvation: the workload would benefit from either adding CPUs or migrating VMs to different hosts. A contention value between 5 percent and 50 percent is "normal overcommit."

VMware best practices describe the available CPU capacity of a host as equal to the number of cores *not including hyperthreads.* Thus, a 10-core Intel Xeon processor E7-4870 CPU (Westemere) host with 2.4 GHz processors has 10 cores x 2,400 MHz/core = 24,000 MHz of available compute capacity. When using all CPUs of a 4-socket server, the maximum available compute capacity would be 4 x 24,000 MHz = 96,000 MHz. When configuring CPU reservations with the maximum compute resources selectable would already take into account the compute resources needed by the hypervisor kernel, thus resulting in a little bit lower available compute resources for the VM. Figure 34 shows an example CPU reservation setting and the maximum selectable compute resources. In this example, 94,992 MHz represents a CPU overhead for the kernel of around one percent, compared to the maximum compute resources of 96,000 MHz. The actual reservation for this SAP HANA VM is set to 90,000, as the VM did not use more compute resources during runtime, and this allowed the ESXi kernel to consume more compute resources.

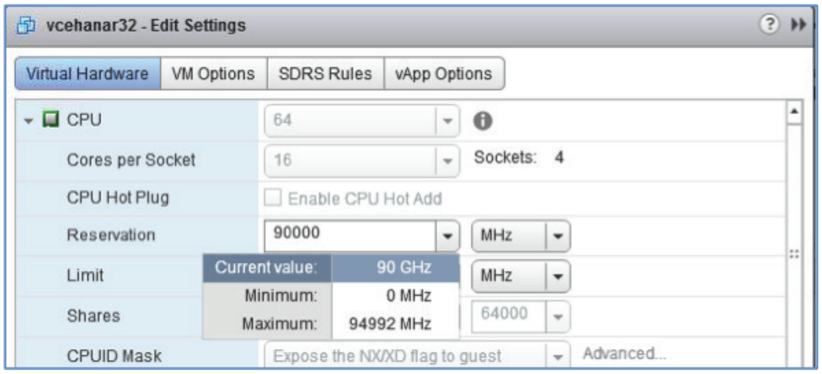


Figure 34. CPU Reservation Setting

The VMware "*Guest and HA Application Monitoring Developer's Guide*" further describes how when actual usage is below compute capacity, the hypervisor is "under committed." In this case, the hypervisor is scaling linearly with the load applied, and there is wasted capacity. As actual usage exceeds available compute capacity, the hypervisor begins utilizing hyperthreads for running virtual machines in order to ensure performance degradation is not disruptive; maximum aggregate utilization occurs during this "normal overcommit" (between 5 percent and 50 percent contention) where each virtual machine sees somewhat degraded performance, but overall system throughput is still increasing. In this "normal overcommit" region, adding load still improves overall efficiency, though at a declining rate. Eventually, all hyperthreads are fully utilized. Efficiency peaks and begins to degrade; this "severe overcommit" (> 50 percent contention) indicates the workload would be more efficient if spread across more hosts for better throughput.

When, as in with SAP HANA workloads very often, a virtual machine gets configured to match the number of host cores (not including hyperthreads), it will peak at the capacity of those cores (with < 5 percent contention), but at a performance between 10 percent and 20 percent lower than an equivalent physical machine utilizing all cores and hyperthreads at all times. A virtual machine configured to match the number of host threads (2 x host cores) will peak at a performance level more analogous to a physical machine. A recent SAP HANA BW-EML benchmark has shown that the difference can be as little as under four percent, but will show ~40 percent contention (the upper end of "normal overcommit") from running half the cores on hyperthreads. This contention metric indicates the load would run better when run on a larger host with additional cores, so it is technically "overcommitted," even though performance is better than a hypervisor running at full commit.[lxvii]

Typical SAP HANA VMs should be configured to be within the "normal overcommit" range as shown in figure 35. When very time-critical SAP HANA workloads are serviced by an SAP HANA system, it is recommended not to use hyperthreads to run this specific SAP HANA VM in the "undercommit" range.

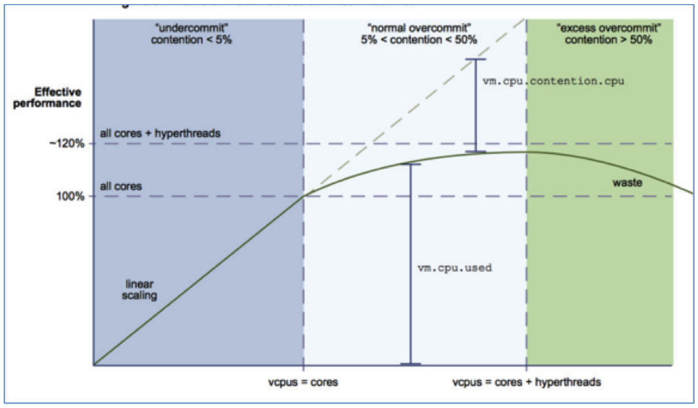


Figure 35. CPU Utilization Across all Virtual Machines[lx]

A recently performed SAP HANA BW-EML Scale-Out benchmark, with an SAP HANA VM configured to consume all available compute resources of a host, has shown an overhead of as little as 3.2 percent.[lxi]

Note: As a general guideline, it is recommended to enable hyperthreading on the host and configure as many vCPUs a NUMA node (CPU socket) provides on threads for an SAP HANA VM.

Impact of hyperthreading on SAP HANA BW systems running in a highly utilized benchmark environment

Recent SAP HANA BW-EML benchmarks have shown that the performance difference in these published benchmarks are largely explainable by the resources usable in a virtual machine compared to a natively running system that has leveraged all available CPU and memory resources.

Table 8 provides an overview of the recently published benchmark results and shows the impact when running a benchmark with fewer hyperthreads available in a system.

CERT	SAP HANA Scale-Up benchmark environment	CPU cores	Usable CPU threads	Virtualized	Result DS/hr with a two-billion record database	Difference in result	Difference in used threads	Performance gain per additional thread
2014009	4 x Intel Xeon Processor E7-4880 v2	60	120	No	126,980			0.21%
2014021	4 x Intel Xeon Processor E7-4880 v2	60	64	Yes	111850	-11,9%	-46.7%	
CERT	SAP HANA Scale-Out Benchmark Environment	CPU Cores	Usable CPU threads	Virtualized	Result DS/hr with a three-billion record database	Difference in result	Difference in used threads	Performance gain per additional thread
2013037	4 x Intel Xeon Processor E7-4870	40	80	No	128650			0.20%
2015016	4 x Intel Xeon Processor E7-4870	40	64	Yes	124,570	-3.2%	-20.0%	

Table 8. Overview of Benchmark Results and Impact of Fewer Hyperthreads in System

From the certified benchmarks of around a 0.2 percent performance gain per used CPU hyperthread, along with the BW-EML benchmark, workload can be calculated. Since the hypervisor prefers to schedule a thread exclusively on a CPU core, the maximum benefit hyperthreading will provide for this specific BW-EML workload would be:

- Around eight percent on a 4-socket Westemere-based server system (40 hyperthreads)
- Around 12 percent on Ivy-Bridge systems (60 hyperthreads)
- Around 14.4 percent on a Haswell system (72 hyperthreads)

These specific benchmark results are in line with the figures seen with general workloads described in the previous section.

Non-Uniform Memory Access

The Non-Uniform Memory Access architecture is common on servers with multiple processor sockets. It is important to understand this memory access architecture when sizing memory-bound and latency-sensitive systems.

A NUMA node is equivalent to one CPU socket. For a server with two sockets, there are two NUMA nodes, for a 4-socket server, there are four NUMA nodes. Therefore, the available number of physical CPU cores and RAM can be divided equally among NUMA nodes. This is critical when sizing virtual machines for optimal NUMA node placement and utilization. For example, a 4-socket, 40-core (10 cores per socket) Intel Westemere-based server with 512 GB RAM, has four NUMA nodes, each with 10 physical cores (CPUs), 20 CPU threads and 128 GB RAM (512 divided by four) or with Intel Ivy-Bridge CPUs, 15 cores, 30 threads and up to 256 GB RAM per NUMA node configured for BW workload. The new Intel Haswell EX E7 88x0v3-based server systems will provide even larger configurations and will support up to 18 cores and 36 threads per CPU socket.

When sizing virtual machines, it is important to carefully consider NUMA node boundaries. In the Intel Xeon processor E7-4870 CPU (Westemere) case, the boundaries are 10 vCPUs and 128 GB RAM. Exceeding the CPU and RAM size boundaries of a NUMA node causes the virtual machine to fetch memory from a remote location, and this can diminish performance. Figure 36 shows the node boundaries in a graphical view; SAP HANA VMs should get sized to stay within a NUMA node for as long as possible.

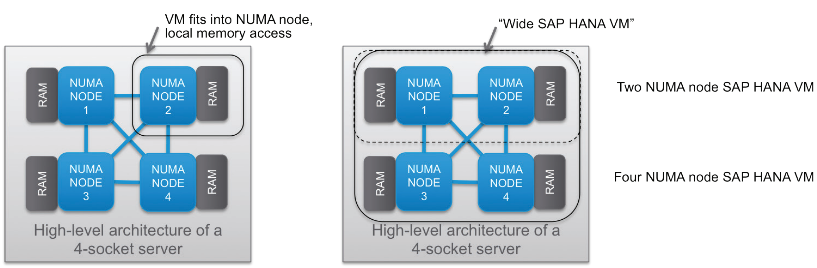


Figure 36. NUMA Node Boundaries

Whenever possible size an SAP HANA VM with the least amount of NUMA nodes. When an SAP HANA VM needs to be larger than a single NUMA node, allocate all resource of this NUMA node, and to maintain a high CPU cache hit ratio never share a NUMA node for production use cases with other workloads. Also, avoid allocation of more memory than a single NUMA node has connected to it, since this would force it to access memory that is not local to the SAP HANA processes scheduled on the NUMA node. Figure 37 provides an overview of NUMA node sharing; note that for production workloads NUMA node sharing is not supported.

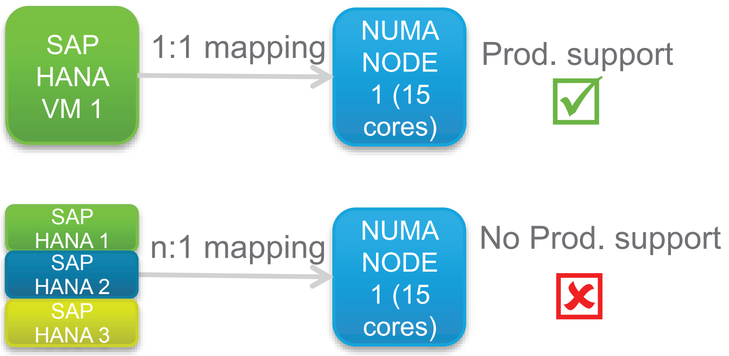


Figure 37. NUMA Node Sharing, Example with an Ivy-Bridge 15-Core System

Figure 38 shows a small 10 vCPU SAP HANA VM that fits into a NUMA node of an Intel Xeon processor E7-4870 CPU (Westemere)-based server system. When hyperthreads are used as well, 20 vCPUs have to get configured for this VM. As already noted, add the parameter **Numa.PreferHT=1** to the VM configuration to ensure only the threads of the local NUMA nodes are accessed, and not threads of another NUMA node; this will break the memory locality. The risk is that when all 20 vCPUs are heavily utilized, the CPU contention value may increase (vm.cpu.contention.cpu). See the previous chapter for details.

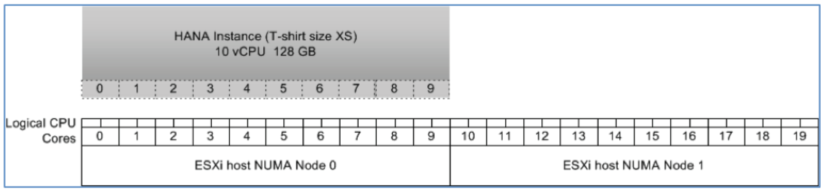


Figure 38. SAP HANA VM Sized to Fit into a Single NUMA Node (CPU socket)

Large SAP HANA VMs that consume more than two NUMA nodes (CPU sockets) may not be able to use all available CPU threads (64 and 128 vCPU limit). In this case it is always better to prefer the physical CPU core instead of using hyperthreads.

Storage Sizing for SAP HANA Virtualized Systems

All SAP HANA worker VMs have a database log, data, root, local SAP, and shared SAP volume. The storage capacity sizing calculation of these volumes is based on the overall amount of memory needed by SAP HANA's in-memory database. SAP has defined very strict performance KPIs that have to be met when configuring a storage subsystem. This might result in more storage capacity than is needed - even if the disk space is not needed – but the amount of spindles may be required to provide the required I/O performance and latency.

SAP has publishedl[xii] guidelines for calculating and sizing the correct storage requirements for SAP HANA TDI configurations. These guidelines change from time to time, therefore it is always recommended to download the latest SAP HANA storage requirements whitepaper. It is also recommended to involve the storage vendor offering an SAP HANA TDI storage solution for SAP HANA Scale-Out configurations to ensure they have followed the correct space and IP sizing guidelines.

The storage volume size figures used in this guide follow, for the most part, the SAP HANA storage requirements published in January 2015. These figures should only be used as an initial sizing guideline, which will have to be aligned to the available SAP sizing reports, Quick Sizer or storage vendor calculation. VMware cannot guarantee that the figures and formulas are 100 percent correct.

ROOT volume:	60 GB
Data volume:	1 .2 x RAM (net disk space)
Redo log:	Systems ≤ 512 GB = 0.5 x RAM
	Systems > 512 GB = 512 GB
SAP HANA shared:	1 x RAM
Backup volume:	Min. 0.5 - 0.6 x RAM

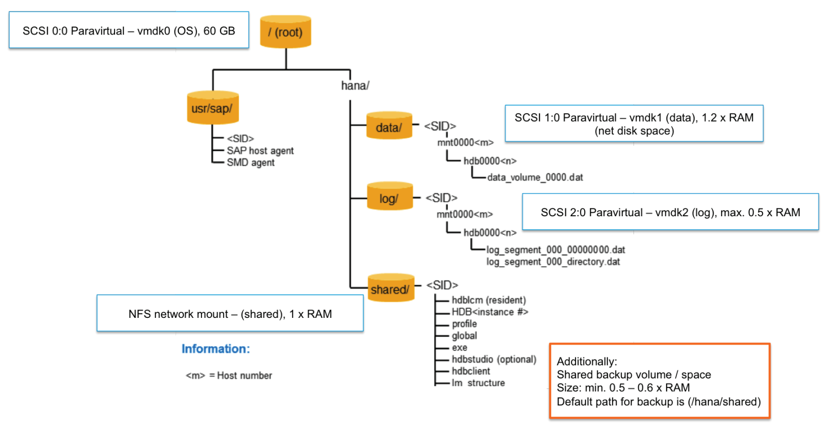
As a default, the backup volume will be stored under the /hana/shared volume. It is recommended to create a dedicated VMDK volume or NFS mount-point to store backup data. When calculating the capacity needs for an 8-node SAP HANA system, as an example, the capacity requirement must be multiplied by 8. To calculate the total disk capacity needed for such an 8-node SAP HANA Scale-Out system, you would follow the formula:

> Total disk capacity = (60 GB for root + 1.2 x RAM for DATA + max. 0.6 x RAM for LOG + 1 x RAM for shared) x number of active nodes + 60 GB for Standby node (when used)
>
> => ***3.3 x RAM + 60 GB for systems ≤ 512 GB or 2.7 x RAM + 572 GB for systems > 512 GB***, plus 60 GB for optional Standby node. The formula includes the space requirement for backup.

Note: Disk capacity sizing needs to be shown as the minimum requirement of the storage configuration. Actual storage sizing may require more disks to fulfill the I/O requirements of SAP HANA. If the I/O needs are fulfilled with a specific SAP HANA TDI configuration, the system must be verified with the SAP HANA HWCCT utility as described in the chapter "*Verifying an SAP HANA Scale-Out TDI on vSphere Configuration*" -on page 59.

In addition to the described volumes, a backup volume may also be required, which would need to be at least 0.5 x RAM of disk capacity. This volume could be a VMDK or an NFS mount. By default, the installation (/hana/shared) volume is also used for backup, but should be manually configured to align to the actual customer requirements - and must belong to a shared file system. For more details, refer to the backup information in the "*SAP HANA Administration Guide.*[lxiii]"

Besides sizing considerations, VMware also recommends using a dedicated para-virtualized SCSI adapter for any SAP HANA "disk." This would result in at least four adapters to support the storage volume configuration outlined on the previous page.



Picture © SAP AGI[lxiv]

Figure 39. Virtualized SAP HANA Scale-Out FC-SAN-based Storage Disk Layout

Network Sizing for SAP HANA Virtualized Systems

Network requirements were discussed in the SAP Scale-Out on vSphere Network Considerations section on page 25. The total amount of supported 1 and 10 Gb NIC ports is limited by the vSphere version being used. vSphere 5.5 supports 8 x 10 Gb and 4 x 1 Gb network ports, and vSphere 6.0 supports 16 x 10 Gb and 4 x 1 Gb network ports.

Table 9 summarizes network requirements and provides an overview of the dedicated networks a virtualized SAP HANA Scale-Out configuration needs. Please be aware that per HANA VM a dedicated 10 GB is needed, otherwise the SAP HANA network KPI may not get achieved.

	MGT, HA AND CLIENT NETWORK	SAP HANA SHARED (NFS MOUNT) NETWORK	VMOTION NETWORK	SAP HANA INTER-NODE NETWORK	SAP HANA NFS STORAGE NODES 2–8*
Network label	External	HANA-NFS-shared	vMotion	HANA-internal	SAP HANA-NFS-storage 2–8
Bandwidth	1 GbE	1 GbE	10 GbE	10 GbE	10 GbE
NIC cards	1		1	1	1
NIC ports**	2		2	2	2
NAS storage	✓				
FC storage	✓				†

Table 9. SAP HANA Scale-Out vSphere Network Sizing

* Minimal one dedicated 10 GbE NIC per SAP HANA VM, maximum of an additional six with vSphere 5.5, and 14 with vSphere 6.0 HANA NFS storage networks are possible.

** With network port redundancy

An 8-node NFS-based SAP HANA Scale-Out system requires a minimum of one 1 Gb dual port NIC and three 10 Gb dual port NICs. Adding more SAP HANA worker nodes to an NFS configuration requires an additional 10 Gb NFS storage network dual port NICs per eight worker nodes.

The theoretical maximum number of worker nodes of a vSphere 5.5-based SAP HANA Scale-Out system is 48 (when all server resources get consumed). vSphere 6.0 could support systems with up to 112 SAP HANA vSphere ESXi hosts. An SAP HANA system can span several vSphere clusters. Remember, SAP Scale-Out TDI configurations are generally supported from SAP for up to 16 hosts. As in any virtualized Scale-Out configuration, support for larger configurations can be requested by SAP. See CA-program for details (SAP HANA Tailored Datacenter Integration Support Process).

Virtualized SAP HANA BWoH Sizing

This section is based largely on the information published by the SAP HANA product management team for SAP HANA sizing.[lxxxii]

SAP has developed a comprehensive set of tools and resources to help customers size their SAP HANA workloads. SAP Quick Sizer tool is a quick and easy way for customers to determine the CPU, memory, and SAPS requirements for running their workloads on SAP HANA. Check the "SAP HANA Sizing Overview" document to quickly find what tools and resources to use to size the different SAP HANA workloads, including:

- **Sizing New Applications** - **"Initial Sizing" section** – Provides an overview of the steps and resources for sizing stand-alone SAP HANA, SAP Business Suite on HANA, SAP HANA Enterprise Search, Industry Solutions powered by SAP HANA, SAP NetWeaver BW powered by SAP HANA, and other sizing guidelines for using SAP HANA as a database.
- **Migrating to SAP HANA Applications** - **"Productive Sizing" section** – Contains the sizing guidelines to help you determine SAP HANA system requirements when migrating your existing applications to SAP HANA.
- **Sidecard Scenarios** – **"Sizing" section** – Describes the sizing process for running SAP HANA Enterprise Search, CO-PA Accelerator, and other SAP Applications on SAP HANA in a sidecard scenario.

The SAP HANA product management team has published seven tips for sizing SAP HANA; it is strongly recommended to follow these when performing an SAP HANA sizing. Once the sizing figures are known, the alignment of these sizing to a virtualized environment has to take place.

- **Tip #1**: Size your application before you look at the SAP HANA server.
- **Tip #2**: Determine the deployment model for SAP HANA that's best for your data center infrastructure strategy.
- **Tip #3**: Carefully explore your options for scale-up, before deciding to go with a scale-out deployment model.
- **Tip #4**: Fully understand the extra options available for reducing the cost of your non-production systems.
- **Tip #5**: Chose the right option among various High Availability/Disaster Recovery (HA/DR) models, including the choice of cost-optimized vs. performance-optimized system replication.
- **Tip #6**: Request a quote from at least two SAP HANA technology partners once you have finalized your landscape design.
- **Tip #7**: Before you go live, validate and take advantage of services offered with your software and hardware licenses.

When sizing a virtualized SAP HANA system, in addition to the SAP sizing guidelines you must take into account the VMware SAP HANA NUMA node recommendations, and VMware vSphere maximums.

As noted previously, SAP HANA BW Scale-Out sizing is mainly determined by the amount of RAM required by the database. Depending on this size, the number of required SAP HANA worker VMs and their resources can also be determined; network and storage configuration requirements are sized accordingly.

In the series of figures that follow, you will get an overview of the CPU and memory sizes for configuring vSphere

SAP HANA VMs. The sizes shown are the maximum values and may differ from the sizes usable for the actual VM configuration, as CPU and memory overhead are not shown.

Figure 40 shows how many vCPUs can be assigned to a VM per NUMA node. When hyperthreading is used, the number of vCPUs should correspond to the number of available physical threads. The horizontal axis shows the NUMA nodes and the vertical axis shows the CPU cores/threads.

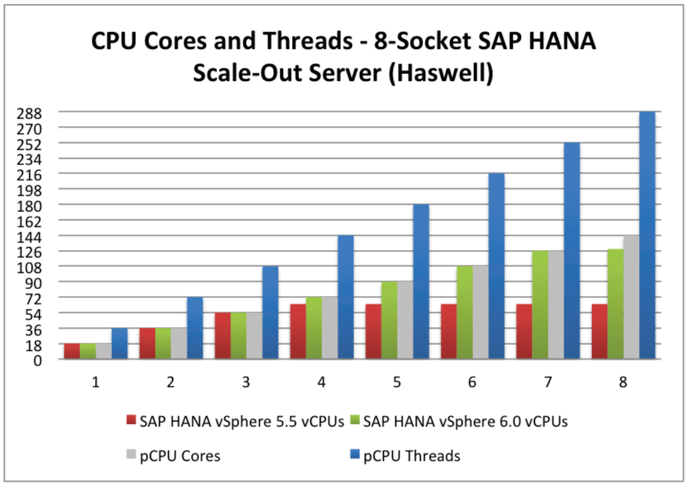


Figure 40. Intel Haswell EX E7 8880v3 CPU Cores and Threads – per NUMA Node and vSphere Version

X-axis – NUMA nodes, Y-axis – CPU threads

The graph in figure 40 can help you determine the maximum size of a VM when it spans one to eight NUMA nodes. vSphere 5.5 with its 64 vCPU support, for instance, can consume all 60 cores of a 4-CPU socket system, but could not leverage all the available hyperthreads (64) of these four NUMA nodes. Due to this limitation, consider vSphere 6.0, which can leverage all available compute resources, resulting in better overall performance.

Figure 41 shows maximum RAM sizes a VM can have by complying with the NUMA node boundaries and the SAP HANA RAM limitation of 256 GB per CPU socket (NUMA node). vSphere 5.5 can support the same amount of RAM per CPU socket as a bare metal or vSphere 6 system. When more than four CPU sockets are used in a host, only vSphere 6.0 could scale the same way a bare metal system does. Nevertheless, two vSphere 5.5 systems, with each 60 vCPUs and 1 TB of RAM, can be installed on the 8-socket server, and consume all available resources *in addition* to the hyperthreads. This would be possible with vSphere 6.0 and its 128 vCPU support.

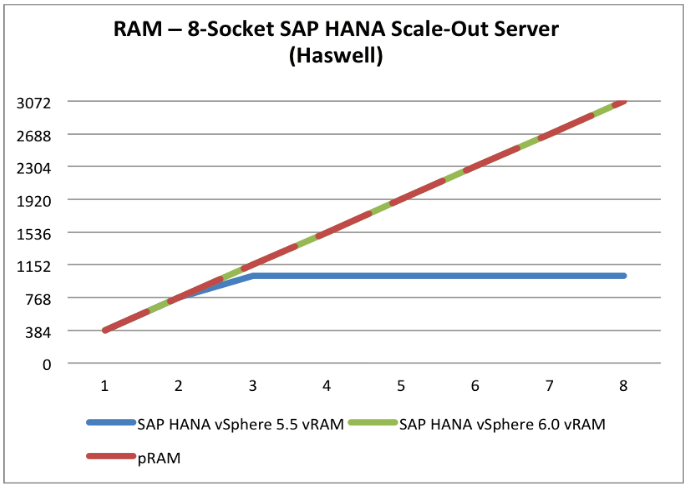
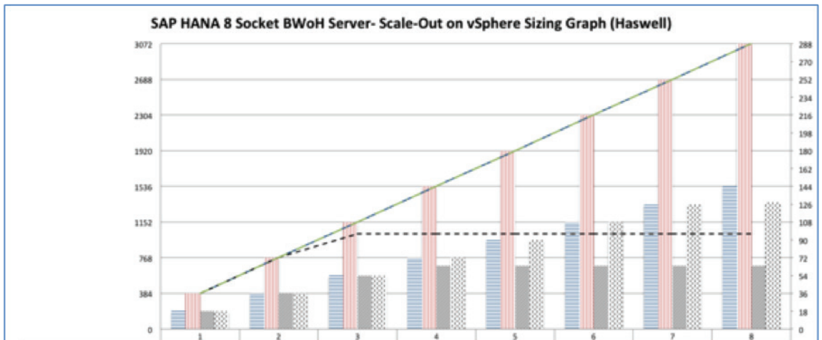


Figure 41. Intel Haswell EX E7 8880v3 CPU - RAM per VM by Maintaining NUMA Node Boundaries and vSphere Version

X-axis – NUMA nodes, Y-axis – memory

Figure 42 shows the sizes of SAP HANA Scale-Out VMs based on supported 8-socket SAP HANA Scale-Out servers. Using this table helps in selecting the right SAP HANA VM configuration needed to support a specific SAP HANA workload. Example SAP Scale-Out on vSphere configurations are discussed on page 60.



	1	2	3	4	5	6	7	8
pCPU Cores	18	36	54	72	90	108	126	144
pCPU Threads	36	72	108	144	180	216	252	288
SAP HANA vSphere 5.5 vCPUs	18	36	54	64	64	64	64	64
SAP HANA vSphere 6.0 vCPUs	18	36	54	72	90	108	126	128
pRAM	384	768	1152	1536	1920	2304	2688	3072
SAP HANA vSphere 5.5 vRAM	384	768	1024	1024	1024	1024	1024	1024
SAP HANA vSphere 6.0 vRAM	384	768	1152	1536	1920	2304	2688	3072

Figure 42. SAP HANA BW on vSphere Sizing Graph for up to Eight NUMA Node Intel Haswell EX E7 8880v3 Systems

X-axis – NUMA nodes, Y-axis – CPU threads and memory

The information in figure 42 summarizes the different SAP HANA BWoH VMs you can configure. If, for instance, a 768 GB RAM SAP HANA master node is needed, the graph provides the details on how many CPUs, threads and NUMA nodes the VM will need to run. In this example, a vSphere 5.5 VM would have the following sizes:

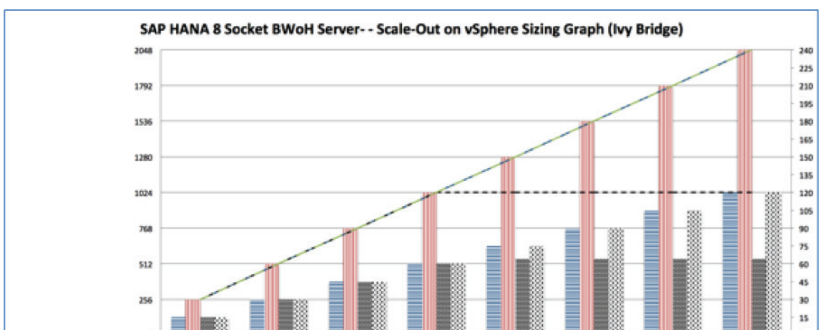
- vRAM 768
- Two CPU sockets (NUMA nodes)
- 36 CPU cores
- Configured with 64 vCPUs to allow this VM to benefit from hyperthreading.

This VM would not be able to utilize eight hyperthreads, but would gain the performance benefits from 28 hyperthreads. A vSphere 6.0 VM would be able to access all available resources, which will provide a slightly better overall performance, since it will be able to use all available 36 hyperthreads.

If you select—for the same workload—an Ivy-Bridge-based server, the result would be a different SAP HANA BWoH VM. Figure 43 provides the details. The 768 GB RAM BWoH VM would need:

- Ivy-Bridge-based CPUs
- Three CPU sockets (NUMA nodes)
- 45 CPU cores
- Configured with 64 vCPUs

This system would benefit from some of the available hyperthreads, but is not able to address all 45 of them.



	1	2	3	4	5	6	7	8
pCPU Cores	15	30	45	60	75	90	105	120
pCPU Threads	30	60	90	120	150	180	210	240
SAP HANA vSphere 5.5 vCPUs	15	30	45	60	64	64	64	64
SAP HANA vSphere 6.0 vCPUs	15	30	45	60	75	90	105	120
pRAM	256	512	768	1024	1280	1536	1792	2048
SAP HANA vSphere 5.5 vRAM	256	512	768	1024	1024	1024	1024	1024
SAP HANA vSphere 6.0 vRAM	256	512	768	1024	1280	1536	1792	2048

Figure 43. SAP HANA BW on vSphere Sizing Graph for up to Eight NUMA Node Intel E7-4890V2 Systems

X-axis – NUMA nodes, Y-axis – CPU threads and memory

Verifying an SAP HANA Scale-Out TDI on vSphere Configuration

To determine if the SAP HANA configuration was configured properly, SAP provides the "*SAP HANA Hardware Configuration Check Tool (HWCCT).*[lxv]" The HWCCT tool allows checking the infrastructure SAP HANA will run on, and will verify adherence to the SAP HANA Tailored Datacenter Integration guidelines and KPIs:

- Measurement of data throughput rates and latency between the SAP HANA computing nodes and the external storage system.
- Measurement of the minimum bandwidth requirement of the SAP HANA intra-node network to ensure it meets the bandwidth requirements of an SAP HANA multi-node/Scale-Out system.

The SAP-defined HANA TDI storage and network KPIs are specified in the *"HWCCT tool documentation,"* which can be downloaded from the SAP support website.

Before a virtualized SAP HANA system can be used for production workloads, it is strongly recommended to verify the configuration with the HWCCT utility to verify that the specified KPIs are met.

Example SAP Scale-Out on vSphere Configurations

This chapter provides example sizing and configurations for SAP HANA Scale-out systems on vSphere with NFS- and FC-based storage solutions. These examples provide information on the required compute, network and storage configurations for 4 to 16-node SAP HANA Scale-Out systems.

Example SAP HANA BWoH Sizing

General considerations:

1. The in-memory SAP HANA BW database size is already known and was verified by SAP and/or the HW partner.
2. Virtualized SAP BW-EML benchmark overhead as documented in benchmark certificate 2015016 < 4 percent,[lxvii] is negligible, as the primary sizing factor is RAM.
3. For memory sizing calculations, a three percent memory overhead is used. Overhead is not used for storage and networking for the sizing calculation, since the network and storage configurations will have to pass the HWCCT KPIs, and may need to change to be optimized to achieve the KPIs.
4. Available SAP HANA Intel Scale-Out hosts, as of June 2015, support RAM sizes from 256 GB for 2-socket systems, and up to 3 TB for 8-socket systems).[lxii]
5. vSphere 5.5 supports cluster sizes up to 32 vSphere hosts and VM RAM sizes of up to 1 TB RAM and 64 vCPUs are possible.
6. vSphere 6.0 supports cluster sizes up to 64 vSphere hosts and VM RAM sizes of up to 4 TB RAM and 128 vCPUs are possible.[lxviii]
7. SAP has currently restricted Scale-Out TDI systems to a maximum of 16 active worker hosts/VMs. More nodes are possible upon SAP approval. An SAP HANA system can span multiple vSphere clusters.
8. If NFS storage is used for log and data volumes, a dedicated 10 GbE NFS storage network will be needed for every eight SAP HANA worker nodes.
9. For production SAP HANA workloads, resource over commitments or CPU socket (NUMA node) sharing is not allowed.
10. In general, approximately 50 percent of available VM memory can be used to store static data; the other 50 percent must be available for SAP HANA dynamic data and working space.
11. Storage capacity needs: 3.3 x RAM + 60 GB for systems ≤ 512 GB, or 2.7 x RAM + 572 GB for systems > 512 GB, including backup, and a 60 GB optional Standby node is required.

vSphere RAM Sizing Example for Scale-Out BWoH Scenarios:

Example 1

The customer wants to migrate an existing SAP BW database from UNIX running on a RISC platform to SAP HANA Scale-Out running on standard x86 server systems with Linux, which offers as much as 20 TB more of storage capacity.

Step 1: Determine the SAP HANA in-memory data footprint

The customer reaches out to SAP and the selected HW partner to perform the sizing of the SAP HANA system. Once the memory need is known, the virtualized system configuration can be planned accordingly. For this example it is assumed that the SAP HANA memory sizing lecture provided that the compressed in-memory SAP HANA database will be around 5 TB.

Step 2: Determine the total RAM needed for the SAP HANA system

Total SAP HANA RAM = compressed in memory SAP HANA data x 2 -> 5 TB x 2 = 10 TB

Step 3: Select the proper SAP HANA server system configuration, and how many hosts are needed

Use the "*SAP HANA Certified Hardware Directory*" webpage to select the proper server configuration the organization wants to buy.

The following formula will help you calculate the number of hosts needed:
Total In-Memory need for SAP HANA / (host RAM – ESXi RAM) + n x standby VMs = Scale-Out VMs

When selecting a 1,024 GB, 4-CPU socket Ivy-Bridge system, the maximum size of an SAP HANA VM is 1,024 GB *minus* the ESXi RAM. Using this formula with a memory overhead of three percent, and by adding the resources to provide HA capabilities, 12 hosts—each with a 1 TB vSphere 5.5 VM—would be needed to support this 5 TB in-memory database.

10,240 GB / (1,024 GB -3 percent) + 1 = 11.31 -> 12 hosts (11 + 1 configuration)

The 11 active nodes will provide 11,264 GB of usable RAM for the SAP HANA in-memory stored database, which is 990 GB of RAM (minus ESXi overhead) more than is required, but can get used for future data growth.

Note: The twelfth host—which provides resources for HA, or for the running a Standby VM—can be used for other workloads as well. Consumed resources will be freed up prior to the start of the production of SAP HANA worker nodes. When selecting a 1.5 TB, 4 x Intel Haswell EX E7 8880v3 system, the maximum size of an SAP HANA VM can be 1,536 GB minus the ESXi RAM; eight 1.5 TB hosts, including HA resources, would be required with VMs as large as these hosts.

10,240 GB / (1,536 GB -3 percent) + 1 = 7.87 -> 8 hosts (7 + 1 configuration)

Step 4: Select the proper network configuration

The minimum network requirements for a virtualized SAP Scale-Out system running on vSphere are as follows *per host*:

- 1 x 1 GbE dual port NIC for management, client, HA and NFS shared storage
- 1 x 10 GbE dual port NIC for vMotion
- 1 x 10 GbE dual port NIC SAP HANA inter-node communication
- Optional: 1 x 10 GbE dual port NIC per worker VM for NFS storage communication, not needed when FC-SAN storage gets used

Depending on the selected storage subsystem, a single 10 GbE dual port NIC per eight worker nodes must be added when NFS-based storage is used. For FC-SAN-based storage the proper number of dual-port HBAs must be added as required by the storage vendor.

Step 5: Storage capacity and the proper storage configuration

The following storage capacity is needed *per host*:

Host storage capacity = 60 GB for ROOT + 1.2 x RAM for DATA + 0.5 x RAM (max. 512 GB) for log + 1 x RAM for shared + 0.5 x RAM for backup

Using this formula provides the storage capacity needs of a single SAP HANA worker VM. 2.7 x 1,024 GB + 572 GB = 3,337 GB per active host and 60 GB for a possible Standby VM. Example configuration: 11 x 3,337 GB + 60 GB (optional Standby VM) = 36,767 GB of overall storage capacity including an optional Standby VM and backup space. When using the Intel Haswell-based 1.5 TB hosts, 33,095 GB of storage capacity are required. Please note that in addition to storage capacity, IOPS requirements must also be fulfilled. Please refer to the storage vendor documentation to select the correct storage and SAP HANA TDI storage KPIs as discussed on page 60.

Example 2

The customer wants to migrate an existing SAP BW database—as large as 5 TB—from UNIX running on a RISC platform to SAP HANA Scale-Out running on a standard x86 server system with Linux.

Step 1: Determine the SAP HANA in-memory data footprint

Perform the SAP HANA in-memory database sizing as specified by SAP. The assumed result for this example is 1.25 TB of RAM.

Step 2: Determine the total amount RAM needed for SAP HANA

Total SAP HANA RAM = compressed SAP HANA data x 2 -> 1.25 TB x 2 = 2.5 TB

Step 3: Select the proper SAP HANA server system configuration, and how many hosts are needed
In this example we select a 512 GB, 2-CPU socket Ivy-Bridge server system for the SAP HANA Scale-Out system.

2,560 GB / (512 GB -3 percent) + 1 = 6.16 -> 7 hosts (6 + 1 configuration)

Step 4: Select the proper SAP HANA server system configuration, and how many hosts are needed

The minimum network requirements for a virtualized SAP Scale-Out system running on vSphere are as follows *per host*:

- 1 x 1 GbE dual port NIC for management, client, HA and NFS shared storage
- 1 x 10 GbE dual port NIC for vMotion
- 1 x 10 GbE dual port NIC SAP HANA inter-node communication
- Optional: 1 x 10 GbE dual port NIC per worker VM for NFS storage communication, not needed when FC-SAN storage gets used

Depending on the selected storage subsystem, a single 10 GbE dual port NIC per eight worker nodes must be added when NFS-based storage is used. For FC-SAN-based storage the proper number of dual-port HBAs must be added as required by the storage vendor.

Step 5: Storage capacity the proper storage configuration

The following storage capacity is needed *per host*:

3.3 x 512 GB + 60 GB = 1,750 GB per active host and 60 GB for a possible Standby VM.

Example configuration: 6 x 1,750 GB + 60 GB = 10,560 GB of overall storage capacity including a Standby VM and backup space. Please note that in addition to storage capacity, IOPS requirements must also be fulfilled. Please refer to the storage vendor documentation to select the correct storage.

Table 10 summarizes the two example sizing calculations

	SAP HANA BW 20 TB SOURCE DATABASE		SAP HANA BW 5 TB SOURCE DATABASE
HANA In-Memory sized DB size*	5,120 GB	5,120 GB	1,280 GB
SAP HANA system RAM needed (static + dynamic data)	10,240 GB	10,240 GB	2,560 GB
Selected host	4-CPU socket Intel **Haswell** with 1,536 GB RAM	4-CPU socket Intel **Ivy-Bridge** with 1,024 GB RAM	2-CPU socket Intel **Ivy-Bridge** with 512 GB RAM
Number of hosts	7	11	6
With Standby VM	8	12	7
Storage capacity with Standby VM and backup space	33,095 GB	36,767 GB	10,560 GB
Network	• 1 x 1 GbE dual port NIC for management, client, HA and NFS shared storage • 1 x 10 GbE dual-port NIC for vMotion • 1 x 10 GbE dual-port NIC for SAP HANA inter-node communication • Optional: 1 x 10 GbE dual port NIC per worker VM for NFS storage communication		

Table 10. Summary of Sizing Examples

*From SAP or SAP HW partner provided SAP HANA in-memory DB size

SAP HANA BWoH Memory Sizing Examples for up to 16 Host Systems

Table 11 provides an overview of the configuration components needed for a 4-, 8-, and 16-node n+1 SAP HANA Scale-Out configuration. The figures in this table can be used as a starting point when designing a custom VMware vSphere virtualized SAP Scale-Out system running on SAP HANA and VMware certified storage and server systems.

The table below shows the maximum memory available for a 4-, 8- or 16-node SAP HANA Scale-Out system. The shown memory is available for a single SAP HANA VM or for multiple-running SAP HANA VMs per node/ host system. For deployment examples please refer to page 9, SAP HANA Scale-Out option.

The actual memory sizing has to be performed by SAP or a selected SAP HW partner. The shown figures are the maximum values possible in such a virtualized SAP HANA environment.

Please note that VM sizes above 1 TB requires vSphere 6, which is in SAP HANA validation. For the latest support status of vSphere 6 please check out the referenced SAP HANA support notes for VMware.

	AVAILABLE TOTAL RAM (-3%) OF ALL ACTIVE SYSTEMS (GB)	TOTAL USABLE SAP HANA IN-MEMORY DATABASE RAM (GB)
	512 GB RAM host systems (2-socket Ivy-Bridge system)	
4-node (3+1) configuration	1,490	745
8-node (7+1) configuration	3,476	1,738
16-node (15+1) configuration**	7,450	3,725
	1,024 GB RAM host systems (4-socket Ivy-Bridge system)	
4-node (3+1) configuration	2,980	1,490
8-node (7+1) configuration	6,953	3,476
16-node (15+1) configuration**	14,899	7,450
	1,536 GB RAM host systems (4-socket Haswell system)	
4-node (3+1) configuration	4,470	2,235
8-node (7+1) configuration	10,429	5,215
16-node (15+1) configuration**	22,349	11,174

Table 11. Virtualized SAP HANA Scale-Out Memory Sizing Examples*

*The figures in the table include a three percent memory overhead for virtualization and assume VMware HA as the HA solution.

**Upon request, more SAP HANA worker nodes are possible.

*** 8-socket systems are in certification

Please note: As of the writing of this guide, no 2-socket Haswell system has been certified for SAP HANA Scale-Out. Therefore, the table may be obsolete once 2-socket Haswell-based server systems are deployed.

Best Practices and Configuration Settings

This section provides information on best practices and suggested settings for SAP HANA on vSphere. The listed reference documents provide more information on the best practices, which are non-SAP HANA specific, and are valid for all performance and latency critical applications.

Reference documents configuration guidelines:

- "SAP HANA Support for VMware vSphere Environments"
- "SAP HANA on VMware vSphere in production (controlled availability)"
- "SAP HANA Guidelines for Being Virtualized with VMware vSphere (v2)"
- "Best Practices and Recommendations for Scale-up Deployments of SAP HANA on VMware vSphere"
- "VMW-Tuning-Latency-Sensitive-Workloads"
- "Timekeeping in VMware Virtual Machines"

Optimizing Virtual Machines for SAP HANA Workloads on vSphere

In general, the testing and performance characterization of SAP HANA on VMware vSphere was conducted using the following VMware ESXi™ performance tunings. This section describes the additional optimizations that are tailored specifically for running SAP HANA in a VMware virtualization environment.

VMware vSphere can run multiple virtual machines in a heterogeneous clustered environment. In order to create SAP landscapes, application owners and virtual infrastructure administrators need to weigh the advantages of absolute performance versus heterogeneity.

These additional optimizations are described in the sections below:

- Workload optimization
- Pinning virtual NUMA to physical NUMA
- Low latency setting
- NIC optimization
- Linux guest operating system
- Virtual SCSI adapter optimization

Installation Steps for an SAP HANA Scale-Out Installation Without a Stand-by Host

Unlike in physical SAP HANA environments, it is possible to operate SAP HANA as highly available without a Standby host. Host/VM failover functionality can be substituted by VMware HA.

The current version of the SAP HANA installer always expects a Standby host configuration even if none is needed. To work around this installer issue please follow these installation steps:

1. Install SAP HANA as a single host installation on the first VM. After a successful installation, go to „/usr/sap/<SID>/SYS/global/hdb/custom/config" and add the following entries to the „global.ini":

   [communication]

   listeninterface = .global

   [persistence]

   basepath_shared = no

   ----

2. After editing the global.ini file restart this SAP HANA instance.
3. For the next step you need the ID number of the <SID>adm (uid) and the sapsys group (gid) of the first SAP HANA node. To find these IDs execute "id <SID>adm" on the first VM, for example:

   vm1 # id vmwadm
   uid=1001(vmwadm) gid=79(sapsys)
   groups=16(dialout),33(video),1001(vmwshm),79(sapsys)

4. On the next VM you want to add as a worker node to the SAP HANA installation, create "/hana/log/<SID>" and "/hana/data/<SID>". After the creation of the two directories, change the owner of those two directories to the previously retrieved <SID>adm (uid) and the sapsys group (gid) by executing "chown <UID>:<GID> <directory>":

   vm2 # chown 1001:79 /hana/data/VMW

5. Now go to "/hana/shared/<SID>/global/hdb/install/bin" on the VM you want to add as an additional worker and execute "./hdbaddhost --storage_partition=<X> -H <Hostname>" where <X> is a number greater than 1 and <Hostname> refers to the internal hostname (the hostname used for the SAP HANA interconnect link) of the host you want to add.

   vm2 # ./hdbaddhost --storage_partition=2 -H hana2int

6. Repeat these steps for every worker you want to add to your SAP HANA installation.

Enable Automatic Start of the SAP HANA Database

By default the SAP HANA database is configured so that it does not started automatically when the SAP HANA host is rebooted.

If VMware HA is selected as the HA solution for SAP HANA, it is recommended to enable the automatic start feature of SAP HANA.

To do this modify the SAP HANA database profile with the following procedure[lxxxiv]

Log on to the SAP HANA host as a user with root authorization.

Change to the system profile directory /usr/sap/<SID>/SYS/profile.

Edit the profile <SID>_HDB<instance number>_<host name>.

Change the parameter setting Autostart = 0 to Autostart = 1 .

Save the profile <SID>_HDB<instance number>_<host name>.

With this modification the SAP HANA database will start automatically when the SAP HANA VM is rebooted.

Workload Optimization

In order to achieve optimal performance for a productive SAP HANA virtual machine, use the settings described in the sections below. In terms of CPU scheduling and priority, these settings improve performance by reducing the amount of vCPU and vNUMA migration, while increasing the priority of the SAP HANA production virtual machine.

Pinning Virtual NUMA to Physical NUMA

vCPU pinning may be used to keep the virtual NUMA nodes from migrating to different physical NUMA nodes. Pin each vCPU to a physical NUMA node using the pattern of configuration settings in the virtual machine's .vmx file, or via the Advanced Configuration Parameters section in the vSphere client.

Here is an example:

```
sched.vcpu0.affinity = "0-19"
…
sched.vcpu9.affinity = "0-19"
sched.vcpu10.affinity = "20-39"
…
sched.vcpu19.affinity = "20-39"
sched.vcpu20.affinity = "40-59"

        sched.vcpu29.affinity = "40-59"
        sched.vcpu30.affinity = "60-79"
        …
        sched.vcpu39.affinity = "60-79"
```

Low Latency Setting

In order to achieve the absolute lowest possible latency for SAP HANA workloads, it is possible to set the latency sensitivity setting to high as follows:

- Go to the Virtual Machine Settings in vCenter Server for the SAP HANA virtual machine
- Click the Virtual Machine Options tab
- Select VM Options
- Select High from the Latency Sensitivity pull down menu as shown in figure 44

It is not generally recommended to use this feature on SAP HANA Scale-Out configurations, because the SAP HANA internode network communication benefits more from network bandwidth than from lower latency.

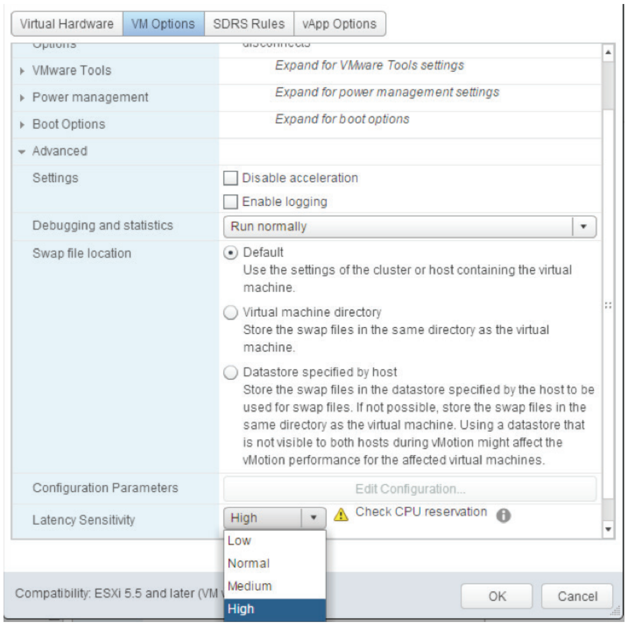


Figure 44. VM Latency Sensitivity Setting

This setting will automatically tune a VM for latency-critical workloads.

What the Latency Sensitivity feature does is:

- Gives exclusive access to physical resources
- Bypasses virtualization layers
- Tunes virtualization layers
- Pre-allocates memory

For details please review the VMware document "Deploying Extremely Latency-Sensitive Applications in VMware vSphere 5.5.[lxx]"

Attention! The low latency setting will lead to longer VM boot times due to the memory pre-allocation and longer vMotion VM switchover times. Longer switchover times may lead to a communication interruption between the application server and the database.

In vSphere 5.5 and 6.0 when setting low latency to high in conjunction with vCPU pinning, it is recommended to leave one core per-socket free for the hypervisor. This is only necessary when all physical cores are allocated to the SAP HANA virtual machine. For example:

- An 8-socket server with 15 cores per socket = 120 cores
- Maximum recommended SAP HANA virtual machine = 112 cores or 14 cores/28 vCPUs per socket

NIC Optimization

Client-facing network cards benefit from disabling interrupt coalescing, by reducing database-to-client response times. Unlike backend networks—like NFS storage or SAP HANA inter-node communication networks where network bandwidth is the critical factor—optimizing network latency is important for client-facing network interfaces

Most NICs provide a mechanism for disabling interrupt coalescing, which is usually available via the ethtool command and/or module parameters.

In order to determine if disabling the physical NIC interrupt moderation on the VMware ESXi host is needed, enter the following command:

```
# esxcli system module parameters set -m ixgbe -p "InterruptThrottleRate=0"
```

This example applies to the Intel 10 GbE driver, called ixgbe. In order to find the appropriate module parameter for the NIC, first find the driver using this ESXi command:

```
# esxcli network nic list
```

Then find the list of module parameters for the driver being used:

```
# esxcli system module parameters list -m <driver>
```

Note: Only disable interrupt coalescing of the client-facing network cards

Linux Guest Operating System

Large Receive Offload (LRO) is another feature of VMXNET 3 that helps deliver high throughput with lower CPU utilization. LRO aggregates multiple received TCP segments into a larger TCP segment before delivering it to the guest TCP stack.

However, for latency-sensitive applications that rely on TCP, the time spent aggregating smaller TCP segments into a larger one adds to the latency. It can affect TCP algorithms such as TCP delayed acknowledgment (ACK) which causes the TCP stack to delay an ACK until the two larger TCP segments are received. This also adds to the end-to-end latency of the application.

In order to determine if disabling LRO will benefit application stack requirements, reload the VMXNET 3 driver in:

```
# modprobe -r vmxnet3
```

Add the following line in /etc/modprobe.conf (Linux-version dependent):

```
options vmxnet3 disable_lro=1
```

Then reload the driver using:

```
# modprobe vmxnet3
```

Note: Only disable LRO of the client-facing network cards

Virtual SCSI Adapter Optimization

In order to increase the queue depth of the pvSCSI driver inside the Linux-based guest on which SAP HANA runs, add the following Linux kernel boot option:

```
vmw _ pvscsi.cmd _ per _ lun=1024 vmw _ pvscsi.ring _ pages=32
```

Attention! Please review VMware KB article 2088157[lxxi] to ensure that the minimum VMware patch level gets used to avoid possible virtual machine freezes under heavy IO load.

Recommended Configuration Settings for SAP HANA on vSphere

The following table provides an overview of the configuration parameters and settings recommended for SAP HANA workload on vSphere.

SERVER BIOS SETTINGS	DESCRIPTION
Enable Intel VT technology	Enable all BIOS virtualization technology settings
Enable CPU Intel Turbo Boost	Allow Intel automatic CPU core overclocking technology (P-States)
Disable QPI Power Management	Static high power for QPI links
Enable hyper-threading	Double the logical CPU cores
Enable execute disable feature	Enable the Data Execution Prevention bit (NX-bit), required for vMotion
Disable node interleaving	
Disable C1E Halt state	Disable Enhanced C-States in Bios
Set Power Management to High Performance	No power saving modus (C-States)
Disable all unused BIOS features like:	Video BIOS Shadowable, Video RAM Cacheable, on-board audio, on-board modem, on-board serial ports, on-board parallel ports, on-board game port, floppy drive, CD-ROM, USB

Table 12. vSphere Physical Host Server Settings

VSPHERE ESXI HOST	DESCRIPTION
	Physical NIC Settings
Networking	Use: Virtual Distributed Switches (vDS) to connect all hosts that work together Define port groups that are dedicated to SAP HANA, management and vMotion traffic
	Storage
Log	Use dedicated HBAs/SCSI controller for DATA, LOG and ESXi Swap/Scratch partition
Data	
Swap	
	Memory
Configure RAM hemisphere mode	Distribute DIMM modules in a way to achieve best performance (hemisphere mode), use fastest DIMM modules available for this RAM size
	CPU
Populate all available CPU sockets, use glueless NUMA architecture	To avoid Timer Synchronization Issues, use a multi-socket server that ensures NUMA node timer synchronization. NUMA systems that do not run synchronized will need to synchronize the timers on the hypervisor area, which is very costly. Reference: "Timekeeping in VMware Virtual Machines"
SAP Monitoring[lxxii]	Enable SAP monitoring on the host -> **Misc.GuestLibAllowHostInfo** and set the value to 1.
Virtual Machine Monitor	Allow vSphere to choose the best virtual machine monitor based on the CPU and guest operating system combination.
CPU/MMU Virtualization Option = Automatic	Hardware Assisted Memory Virtualization

Table 13. vSphere ESXi Server Settings

SAP HANA VIRTUAL MACHINE	DESCRIPTION
VMXNET3	Use paravirtual VMXNET 3 virtual NICs for SAP HANA virtual machines
Set "ethernetX.coalescingScheme" to disable	Disable virtual interrupt coalescing for VMXNET 3 virtual NICs that communicate with the App servers or front end
Dedicated CPU and memory resources for HANA instances	Do not overcommit CPU or memory resources
Use as few NUMA nodes as possible	Optimize VM size to as few NUMA nodes as possible
Do not oversize the VM	
Align virtual CPU VM configuration to actual server HW	Example: A VM running on a 4-socket server with 10-core CPUs, should get configured with 10-cores per socket, like 2-CPU socket and 20 vCPUs.
numa.nodeAffinity = NUMA nodes (e.g. 0,1,2,3)	NUMA node localization, important for Multi-VM configurations (not needed for 1:1 configuration)
SAP monitoring[lxxiii]	Enable SAP monitoring on the guest VM tools.guestlib.enableHostInfo = true
Paravirtualized SCSI driver for IO devices	Use multiple and dedicated SCSI controller; for details see SAP HANA disk layout section
Use Virtual Machine File System - VMDKs	
Create dedicated and isolated datastores for SAP HANA data and log files	Ensure the storage configuration passes the SAP HANA HW check tool storage and file system requirements
Use Eagerzeroed thick virtual disks for data and log disk	This avoids lazy zeroing
Remove unused devices	Like floppy disks or CD-ROM

Table 14. Virtual Machine Settings

LINUX OS	DESCRIPTION
	Apply changes as documented in the following SAP and VMware notes:
SAP HANA DB: Recommended OS settings for SLES 11 / SLES for SAP Applications 11 SP2	http://service.sap.com/sap/support/notes/1824819
SAP HANA DB: Recommended OS settings for SLES 11 / SLES for SAP Applications 11 SP3	http://service.sap.com/sap/support/notes/1954788
Disable transparent HugePages in kernel	echo never > /sys/kernel/mm/transparent_hugepage/enabled verify with cat /sys/kernel/mm/transparent_hugepage/enabled always madvise (never) See SAP notes 2013638 (RHEL 6.5), 1954788 (SLES 11 SP3) and 1824819 (SLES 11 SP2)
Configure C-States for lower latency in Linux:	Disable in BIOS (e.g. static high performance or OS controlled) • If disabled in BIOS, optional: intel_idle.max_cstate=0 • If disabled in BIOS, optional: processor.max_cstate=0
GlibC update to avoid potential problem in the index server:	Update your GlibC at least to version glibc-2.11.3-17.56.2
Optional: Enabling LRO in Linux guest operating system to lower latency for client NIC adapter[lxxiv]	Optional use case to lower network latency of client traffic NIC adapter: "ethtool -K ethY lro on." Do not disable LRO for NFS or SAP HANA internode communication network! Works only with Linux kernel 2.6.24 and later and uses a VMXNET3.
Linux kernel	VMware strongly recommends using the latest kernel version
Minimal supported SUSE Linux kernel 3.0.101-0.35.1se[lxxv]	Otherwise customers may experience unplanned crashes/downtimes when many CPUs and much memory is used.
Do not set page cache limits[lxxvi]	This should only be done when the system experiences a performance problem caused by heavy paging/swapping, as is described in the referenced SAP note. Linux: VMware vSphere configuration guidelines[lxxvii]

Table 15. Virtual Machine Guest Operating System

LINUX OS	DESCRIPTION
Disable I/O scheduling	Set kernel parameter "elevator=noop" -> Change the following parameters Install latest VMware Tools™ echo never > /sys/kernel/mm/transparent_hugepage/enabled
Disable all unnecessary Linux services like:	anacron, apmd, atd, autofs, cups, cupsconfig, gpm, isdn, iptables, kudzu, netfs, and portmap Disable SUSE scheduler
Install the following packages:	gtk2 (use version as provided with operating system distribution) java-1_6_0-ibm or later (use version as provided with operating system distribution, necessary for the SAP HANA studio on the SAP HANA system) libicu (use version as provided with operating system distribution) mozilla-xulrunner192-1.9.2.xx-x.x.x (use version as provided with operating system, but at given minimum version) ntp, sudo syslog-ng (use version as provided with operating system distribution) tcsh, libssh2-1, expect, autoyast2-installation. yast2-ncurses
Disabling "Large Receive Offload"	Disabling LRO to lower latency for application that relies on TCP. Only for the App Server and front end interface. Add "ethtool -K ethX LRO off" to boot.local or edit the ethX setting through yast
Install the latest version of VMware Tools[lxxviii]	VMware Tools is a suite of utilities that enhance the performance of the virtual machine's guest operating system and improves management of the virtual machine.
Turn off the SLES kernel dump function	Turn off the SLES kernel dump function (kdump) if not needed for specific reasons.
Change the following parameters:	elevator=noop http://doc.opensuse.org/products/draft/SLES/SLES-tuning_sd_draft/cha.tuning.io.html vmw_pvscsi.cmd_per_lun=1024 set them through yast in the boot loader section, just copy them there separated through space vmw_pvscsi.ring_pages=32
/etc/sysctl.conf	net.core.rmem_default = 262144 net.core.wmem_max = 8388608 net.core.wmem_default = 262144 net.core.rmem_max = 8388608 net.ipv4.tcp_rmem = 4096 87380 8388608 net.ipv4.tcp_wmem = 4096 65536 8388608 net.ipv4.tcp_mem = 8388608 8388608 8388608 net.ipv4.tcp_slow_start_after_idle = 0
Adhere to the following shared memory settings:	Small, Shmmni value = 4096 if RAM > = 24 GB & < 64 GB Medium Shmmni value = 65536 if RAM > = 64 GB & < 256 GB Large Shmmni value = 53488 if RAM > 256 GB
Configure NTP time server[lxxix]	Linux VMware timing problem
Align file system offset – Ex: VNX 128K offset	

Table 15. Virtual Machine Guest Operating System (con't)

Conclusion

Virtualizing SAP HANA Scale-Out systems allows an organization to benefit from all supported VMware virtualization solutions and options—like live migration via vMotion—to increase SLAs and lower TCO. The recent joint SAP and partner testing and subsequent release of the VMware virtualized BW-EML benchmark of an SAP HANA Scale-Out system has shown reliable operation with a very small performance impact on overall system performance. SAP HANA Scale-Out support in controlled availability provides additional benefits for the customer by getting additional consultancy from SAP, VMware and the HW partner. This ensures that their virtualized SAP HANA configurations work well and run optimally in a virtualized environment.

Appendix: Troubleshooting vSphere-related Issues

Open an SAP Support Request Ticket

VMware is part of the SAP support organization, allowing VMware support engineers to work directly with SAP, SAP customers, and other SAP software partners, like SUSE, as well as with HW partners on solving issues needing escalation.

Before opening a VMware support ticket, it is recommended to open a support request within the SAP support system, when the SAP system runs virtualized with VMware.

This ensures that SAP VMware specialists will work on the case, and, if needed, escalate the issue to VMware product support (when it is a VMware product issue) or to SAP support (when it is an SAP HANA issue).

The following “components” are available for escalating SAP on vSphere issues:

- BC-OP-NT-ESX (Windows on VMware ESX)
- BC-OP-LNX-ESX (Linux on VMware ESX)

SAP HANA VMware vSphere-related issues should be escalated to BC-OP-LNX-ESX.

Figure 45 shows the support process workflow for VMware-related SAP issues.

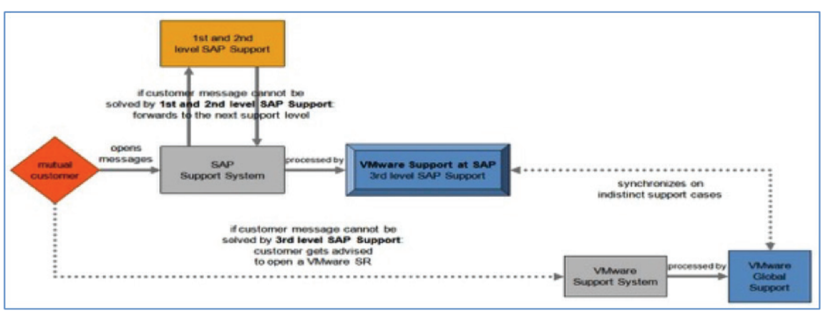


Figure 45. SAP Support Workflow for VMware-related Escalations

Non-VMware-related SAP escalations will be moved to the correct support component. If the issue, for instance, is a Linux kernel panic, or an SAP HANA product issue, it is recommend to first use the correct support component instead of using the VMware support component.

Open a VMware Support Request Ticket

If it appears that VMware vSphere is not configured optimally and is causing a bottleneck, file a support request on My VMware at: http://www.vmware.com/support/contacts/file-sr.html.

In addition:

- Follow the troubleshooting steps outlined in "*Troubleshooting ESX/ESXi virtual machine performance issues (2001003)*" at: http://kb.vmware.com/kb/2001003.
- Verify that all best practices have been applied. See "*Performance Best Practices for VMware vSphere 5.5*" at: http://www.vmware.com/pdf/Perf_Best_Practices_vSphere5.5.pdf.
- Run the vm-support utility, and then execute the following command at the service console: vm support-s
  - This command collects necessary information that VMware uses to help diagnose issues. It is best to run this command when symptoms occur.
- For more information, see the "*Resources*" section near the end of this guide.

Acknowledgements

Author:

- **Erik Rieger**, Technical Alliance Manager, SAP Global Alliances, VMware

The following individuals contributed to the creation of this guide:

- **Kevin Christopher**, Senior Staff, VMware
- **Bob Goldsand**, Staff Partner Architect Global Alliances, VMware
- **Sebastian Lenz**, Senior SAP Support Engineer, VMware
- **Todd Muirhead**, Staff Engineer Performance Team, SAP Alliances, VMware

Resources

For more information about the SAP HANA and VMware products and technologies discussed in this guide, review the links and references below.

Table of Tables

Table of Figures

Table of Resources/Footnotes

i. Scale-Up SAP HANA VMware Best Practices Guide: http://www.vmware.com/files/pdf/SAP_HANA_on_vmware_vSphere_best_practices_guide.pdf
ii. Running multiple HANA VMs on a single host is in controlled availability and must follow the sizing and configuration guidelines described in this guide and the relative SAP support notes.
iii. 3 TB is the current BWoH Maximum SAP allows with 8 CPU socket Haswell systems
iv. SAP HANA Master Guide, page 4
v. SAP HANA Server Installation and Update Guide, chapter 2.2
vi. 4 TB HANA VMs require VMware vSphere version 6 or later
vii. http://h30094.www3.hp.com/product/sku/10724004, list price $260 USD as of April 2015
viii. http://h30094.www3.hp.com/product/sku/10724005, list price $700 USD as of April 2015
ix. http://h30094.www3.hp.com/product/sku/10832312, list price $7,000 USD as of April 2015
x. SAP HANA Hardware Directory: http://global12.sap.com/community/ebook/2014-09-02-hana-hardware/enEN/appliances.html
xi. Hitachi Unified Compute Platform for the SAP HANA Platform on VMware vSphere: http://www.hds.com/assets/pdf/hitachi-ucp-for-the-sap-hana-on-vmware-vsphere-using-compute-blade-500-and-hus-vm.pdf
xii. http://scn.sap.com/docs/DOC-60328, PDF: SAPHANAtailoreddatacenterintegration_L1_Dec192013_final.pdf, pages 4–6
xiii. http://www.sdn.sap.com/irj/scn/go/portal/prtroot/docs/library/uuid/009bb21c-49b5-3210-afb5-d641f300d759?QuickLink=index&overridelayout=true&60125247185073, page 5
xiv. http://www.vmware.com/files/pdf/techpaper/vmw-wp-secrty-vsphr-hyprvsr-uslet-101.pdf, page 5+
xv. http://kb.vmware.com/selfservice/microsites/search.do?language=en_US&cmd=displayKC&externalId=2080735
xvi. https://blogs.saphana.com/2015/02/18/sap-hana-tailored-data-center-integration-tdi-overview/
xvii. http://service.sap.com/sap/support/notes/1943937
xviii. Solution brief: SAP HANA on VMware vSphere for Production Environments, http://www.vmware.com/files/pdf/SAPHANA-VMware-Brief-HIRES.pdf
xix. Document: Scale-Up SAP HANA VMware Best Practices Guide: http://www.vmware.com/files/pdf/SAP_HANA_on_vmware_vSphere_best_practices_guide.pdf
xx. https://pubs.vmware.com/vsphere-55/topic/com.vmware.ICbase/PDF/vsphere-esxi-vcenter-server-55-storage-guide.pdf
xxi. https://pubs.vmware.com/vsphere-55/topic/com.vmware.ICbase/PDF/vsphere-esxi-vcenter-server-55-storage-guide.pdf, page 25
xxii. https://www.suse.com/products/highavailability/
xxiii. SAP HANA Server Installation and Update Guide, pages 17 and 54
xxiv. http://help.sap.com/hana/SAP_HANA_Server_Installation_Guide_en.pdf, page 55
xxv. http://www.netapp.com/us/media/tr-4290.pdf
xxvi. http://germany.emc.com/collateral/white-paper/h13047-storage-config-best-practices-sap-hana-tdi-vnx-wp.pdf
xxvii. EMC IT, 02/14 EMC Perspective, H12853
xxviii. http://kb.vmware.com/selfservice/microsites/search.do?language=en_US&cmd=displayKC&externalId=1034165
xxix. http://help.sap.com/hana/SAP_HANA_Server_Installation_Guide_en.pdf, page 53
xxx. SAP Support for SAP HANA Scale-Out Configurations - http://scn.sap.com/docs/DOC-60470
xxxi. https://pubs.vmware.com/vsphere-55/topic/com.vmware.ICbase/PDF/vsphere-esxi-vcenter-server-55-networking-guide.pdf, page 18
xxxii. http://help.sap.com/hana/SAP_HANA_Server_Installation_Guide_en.pdf, page 15
xxxiii. http://www.netapp.com/us/media/tr-4338.pdf, page 7
xxxiv. http://scn.sap.com/docs/DOC-60334, guide: HANA_HA_2.1.pdf
xxxv. http://scn.sap.com/docs/DOC-60334, guide: HANA_HA_2.1.pdf
xxxvi. https://scn.sap.com/docs/DOC-62435
xxxvii. BUSINESS CONTINUITY BEST PRACTICES FOR SAP HANA TAILORED DATACENTER INTEGRATION, http://www.emc.com/collateral/white-papers/h13955-vnx-storage-mirrorview-snapview.pdf
xxxviii. http://www.vmware.com/products/vsphere/features/high-availability.html
xxxix. Symantec ApplicationHA, http://www.symantec.com/application-ha/data-sheets-white-papers/
xl. http://www.vmware.com/files/pdf/techpaper/vmware-emc-it-virtual-oracle-deploy-framework.pdf
xli. http://help.sap.com/hana/SAP_HANA_Server_Installation_Guide_en.pdf, page 53
xlii. http://kb.vmware.com/selfservice/microsites/search.do?language=en_US&cmd=displayKC&externalId=1034165
xliii. http://help.sap.com/hana/SAP_HANA_Server_Installation_Guide_en.pdf, page 54

xliv. http://www.vmware.com/business-continuity/disaster-recovery
xlv. http://www.vmware.com/files/pdf/products/SRM/VMware-vCenter-SRM-Datasheet.pdf
xlvi. 1781986 - Business Suite on SAP HANA Scale Out
xlvii. http://scn.sap.com/community/hana-in-memory/blog/2013/01/17/sap-hana-sizing
xlviii. 1514966 - SAP HANA 1.0: Sizing SAP In-Memory Database
xlix. 1855041 - Sizing Recommendation for Master Node in BW-on-HANA
l. http://www.sdn.sap.com/irj/scn/go/portal/prtroot/docs/library/uuid/009bb21c-49b5-3210-afb5-d641f300d759?QuickLink=index&overridelayout=true&60125247185073, page 5
li. http://global.sap.com/community/ebook/2014-09-02-hana-hardware/enEN/appliances.html#categories=Intel Ivy Bridge EX E7,Scale-out&order=MemoryDesc
lii. https://www.vmware.com/pdf/vsphere5/r55/vsphere-55-configuration-maximums.pdf
liii. https://www.vmware.com/pdf/vsphere6/r60/vsphere-60-configuration-maximums.pdf
liv. Shown configuration is a Scale-Up configuration, as this system is for BW workloads only supported with 2 TB of RAM.
lv. http://ark.intel.com/de/products/53579/Intel-Xeon-Processor-E7-4870-30M-Cache-2_40-GHz-6_40-GTs-Intel-QPI
lvi. http://ark.intel.com/de/products/75251/Intel-Xeon-Processor-E7-4890-v2-37_5M-Cache-2_80-GHz
lvii. http://ark.intel.com/products/84685/Intel-Xeon-Processor-E7-8890-v3-45M-Cache-2_50-GHz
lviii. http://global.sap.com/community/ebook/2014-09-02-hana-hardware/enEN/appliances.html
lix. http://kb.vmware.com/selfservice/microsites/search.do?language=en_US&cmd=displayKC&externalId=2003582
lx. http://pubs.vmware.com/vsphere-60/topic/com.vmware.ICbase/PDF/vs600_guest_HAappmon_sdk.pdf, page 25++
lxi. Native benchmark: http://global.sap.com/solutions/benchmark/pdf/Cert13037.pdf, virtual benchmark: http://global.sap.com/solutions/benchmark/pdf/Cert15016.pdf
lxii. http://1h9mo84f79012h5z12640ehs-wpengine.netdna-ssl.com/wp-content/uploads/2015/02/Storage-Whitepaper-2-55.pdf
lxiii. http://help.sap.com/hana/sap_hana_administration_guide_en.pdf
lxiv. SAP HANA Server Installation Guide, page 17
lxv. http://service.sap.com/sap/support/notes/1943937
lxvi. SAP_HANA_Hardware_Configuration_Check_Tool_1.6.pdf
lxvii. SAP BW-EML VCE vSphere SAP HANA Scale-Out Benchmark, certification number 2015016
lxviii. vSphere 6 is in SAP HANA certification
lxix. http://global.sap.com/community/ebook/2014-09-02-hana-hardware/enEN/appliances.html#order=CPU Desc&categories=Scale-out
lxx. http://www.vmware.com/files/pdf/techpaper/latency-sensitive-perf-vsphere55.pdf, page 7
lxxi. http://kb.vmware.com/kb/2088157
lxxii. http://service.sap.com/sap/support/notes/1409604
lxxiii. http://service.sap.com/sap/support/notes/1409604
lxxiv. http://kb.vmware.com/kb/2055140
lxxv. http://service.sap.com/sap/support/notes/1557506
lxxvi. http://service.sap.com/sap/support/notes/1557506
lxxvii. http://service.sap.com/sap/support/notes/1122388
lxxviii. http://kb.vmware.com/kb/1014294
lxxix. http://service.sap.com/sap/support/notes/989963
lxxx. http://service.sap.com/sap/support/notes/1736976
lxxxi. Quick Sizer for HANA webpage: http://www.service.sap.com/hanaqs
lxxxii. https://blogs.saphana.com/2014/11/06/7-tips-and-resources-for-cost-optimizing-sap-hana-infrastructure-2/
lxxxiii. http://scn.sap.com/docs/DOC-47702
lxxxiv. http://help.sap.com/saphelp_hanaplatform/helpdata/en/cb/2374f8bb571014b5a5853a3bf45c4c/content.htm

VMware, Inc. 3401 Hillview Avenue Palo Alto CA 94304 USA Tel 877-486-9273 Fax 650-427-5001 www.vmware.com