Cisco HyperFlex Solutions

Running SAP Production Landscapes in a Cisco HyperFlex 2.0 All-Flash System with an SAP HANA Appliance

Author: Brian Everitt

4/10/2017

Executive Summary

SAP landscapes frequently are deployed on virtualization platforms, most often using virtualized application servers. In recent years, SAP has been encouraging its customers to migrate to SAP's own database platform of the future: SAP HANA. SAP HANA databases can be deployed on virtual servers or on physical machines. However, deployment of SAP HANA as a virtualized database has some strict limitations, and in some scenarios SAP may decline to offer full production-level support. A way to help ensure full production support is to deploy SAP HANA using a certified appliance model. With the launch of the Cisco HyperFlex™ system, Cisco offers a low-cost, easy-to-deploy, high-performance hyperconverged virtual server platform that is an excellent solution for SAP landscapes. Alongside the Cisco HyperFlex solution, customers can deploy other standalone servers within the same Cisco Unified Computing System™ (Cisco UCS®) domain, including certified SAP HANA appliances based on Cisco UCS C-Series Rack Servers. This combination of fully virtualized SAP application servers and SAP HANA running on a certified appliance is an excellent choice for production SAP landscapes. For systems not requiring production support, the Cisco HyperFlex solution can be used to deploy SAP application servers and fully virtualized SAP HANA servers, as described in separate white papers published by Cisco.

Solution Overview

This section defines current IT Infrastructure challenges and presents a high-level view of the Cisco HyperFlex and SAP HANA solution and its benefits.

Document Purpose
This document describes deployment strategies for an SAP landscape using virtualized SAP NetWeaver application servers running on a Cisco HyperFlex 2.0 all-flash cluster and an SAP HANA back-end database running on a certified Cisco UCS C460 M4 Rack Server appliance.

Solution Purpose
Cisco HyperFlex HX Data Platform solutions are built on the Cisco Unified Computing System™ (Cisco UCS®) platform. They offer faster deployment and greater flexibility and efficiency at a competitive price, while lowering risk for the customer. Proven components from Cisco are integrated to form a software-defined storage (SDS) platform. This approach eliminates or reduces the need for planning and configuration decisions, while allowing the levels of customization needed to address customer workload needs. The platform and management model adopted is an extension of established Cisco UCS data center strategy, in which familiar components are managed in a consistent manner through a policy-based framework with Cisco UCS Manager.

Business Challenge

Deployment of an SAP landscape in a typical data center environment presents several challenges:

- **Mixed hardware platforms:** Many landscapes are deployed in such a way that the NetWeaver application tier runs in a separate environment from the other components, such as the database tier, often on completely different hardware. This arrangement can also apply to storage; in this case, multiple storage systems may contain the application virtual machines, while the database data is stored elsewhere.
- **Multiple management silos:** Use of multiple hardware platforms and storage systems can lead to a disaggregated management model, in which only certain members of the IT staff can perform tasks on systems for which they have specialized knowledge.
- **Speed of deployment:** Deploying all the layers of a typical SAP environment can be time consuming, requiring setup of networking resources, computing hardware, storage arrays, SANs, and virtualization software.

The Solution

Cisco UCS can be used to deploy a Cisco HyperFlex system to rapidly implement an SAP landscape with virtualized application servers and a standalone SAP HANA database on standard Cisco UCS rack servers.

Solution Benefits

The Cisco HyperFlex and SAP HANA solution provides the following benefits to customers:

1. **Single hardware platform:** Cisco UCS is the base platform for Cisco HyperFlex systems, which provide a fully contained hyperconverged environment, combining networking, storage, and virtualization resources in a single system. Alongside the Cisco HyperFlex solution, additional Cisco UCS servers can be deployed in the same Cisco UCS domain to service other workloads.
2. **Simplified management:** A single administrator can manage all aspects of Cisco UCS and the Cisco HyperFlex system through Cisco UCS Manager and the VMware vCenter Web Client, making tasks much easier and faster to complete.
3. **Rapid deployment:** The programmability and ease of use of Cisco UCS Manager allow Cisco HyperFlex systems to be deployed quickly. These features also allow you to rapidly provision additional Cisco UCS servers for other workload requirements.

Customers who have already invested in Cisco® products and technologies have the opportunity to mitigate their risk further by deploying familiar and tested Cisco UCS technology.

Recommendations Summary

The following list provides a summary of the recommendations for setting up your system:

- Consult the SAP Quick Sizer to help determine the memory, storage, and processor requirements for the SAP HANA database and the application servers to be used in your specific deployment.
- Select an Intel® processor model for your Cisco HyperFlex converged nodes that offers an optimal balance of CPU cores, clock speed, and price for your deployment and budget.
- Set **numa.vcpu.preferHT** to **true** for the SAP application server virtual machines to help ensure that they use hyperthreading properly.

- Configure SAP application server virtual machines with a number of virtual CPUs (vCPUs) that fit within a single non-uniform memory access (NUMA) node. Review the virtual machine's vmware.log file to help ensure that the configuration is correct.
- Configure additional service profiles in Cisco UCS with the appropriate settings for a standalone Cisco UCS C460 M4 SAP HANA appliance.
- Install the Cisco precertified SAP HANA appliance according to the guidelines specified by SAP to gain full production support.
- Use VMware Distributed Resource Scheduler (DRS) groups and affinity rules to keep SAP NetWeaver virtual machines separate from each other.
- Configure SAP System Monitoring at the guest virtual machine level.
- Deploy the solution on a Cisco HyperFlex cluster configured with a replication factor of 3 (RF3), to provide the best possible level of data protection, or alternatively consider replication factor of 2 (RF2), if the additional redundancy is not required, in exchange for lower write latency.

Detailed descriptions of these recommendations follow in this solution guide.

Solution Guide

The following sections describe the system architecture used for the solution. The solution consists of an SAP landscape, with virtualized SAP application servers running on the Cisco HyperFlex 2.0 all-flash system, and SAP HANA used as the database platform running on a Cisco certified SAP HANA appliance. Additional sections present sizing guidelines for the CPU and memory resources, high-availability options, tuning options for virtualized SAP systems, and a reference system design tested by Cisco. The Cisco HyperFlex system provides a hyperconverged architecture, containing all the processor, memory, networking, and storage resources necessary to run a virtual server platform.

This document does not describe the design, installation, or configuration of the Cisco HyperFlex system; those details are covered in other Cisco Validated Design documents, which are listed in the references section at the end of this paper. In addition, the installation of the SAP HANA certified appliance must adhere to strict guidelines, which are also found in the references section. This document also does not include detailed instructions about the installation of the SAP software or OS-level tuning required by SAP, although references to this information can also be found at the end of this document.

Requirements
To configure an SAP landscape running on a Cisco HyperFlex 2.0 all-flash cluster you need:

- A functional and healthy running Cisco HyperFlex cluster
  - Cisco HyperFlex Release 2.0.1.a or later
  - Cisco UCS Firmware Release 3.1(2f) or later
  - VMware ESXi 6.0 Update 2 or later
- VMware vCenter Server Appliance or Microsoft Windows–based vCenter Server
- Appropriate software and licensing from SAP for SAP HANA and NetWeaver

To configure an SAP HANA database on a certified Cisco appliance model you need:

- A Cisco certified SAP HANA appliance server model as documented in the Certified SAP HANA Hardware Directory
- Installation of the SAP HANA database software by Cisco or by a third party that has achieved the appropriate SAP HANA Certified Installer certification.

Limitations

Based on information and support stances from SAP, VMware, and Cisco, a virtual SAP landscape with a standalone SAP HANA database appliance is subject to the following limitations:

- Only a single-node scale-up SAP HANA database can be deployed on a standalone Cisco UCS C460 M4 appliance with internal storage.
- Scale-out multinode database deployments require connectivity to an external shared storage array.

Solution Overview

For a virtualized SAP landscape, all the SAP application server instances are constructed as virtual machines within the Cisco HyperFlex system alongside a standalone SAP HANA database appliance running on a Cisco UCS C460 M4 server within the same Cisco UCS domain. Therefore, various support and sizing guidelines from SAP, Cisco, and VMware need to be followed to help ensure that the virtual applications and standalone database function at their full potential.

SAP Landscape Design

The recommended SAP landscape design (Figure 1) includes one or more SAP NetWeaver application servers, in our tested design running SAP Business Warehouse (BW), running as virtual machines within a Cisco HyperFlex cluster. The design uses a single-node SAP HANA server, also known as a scale-up SAP HANA appliance, running a bare-metal OS on a standalone Cisco UCS C460 M4 Rack Server. An optional component is an additional SAP HANA appliance, which can receive a replicated copy of the SAP HANA database using the SAP HANA System Replication feature.

Figure 1.   SAP Landscape Design
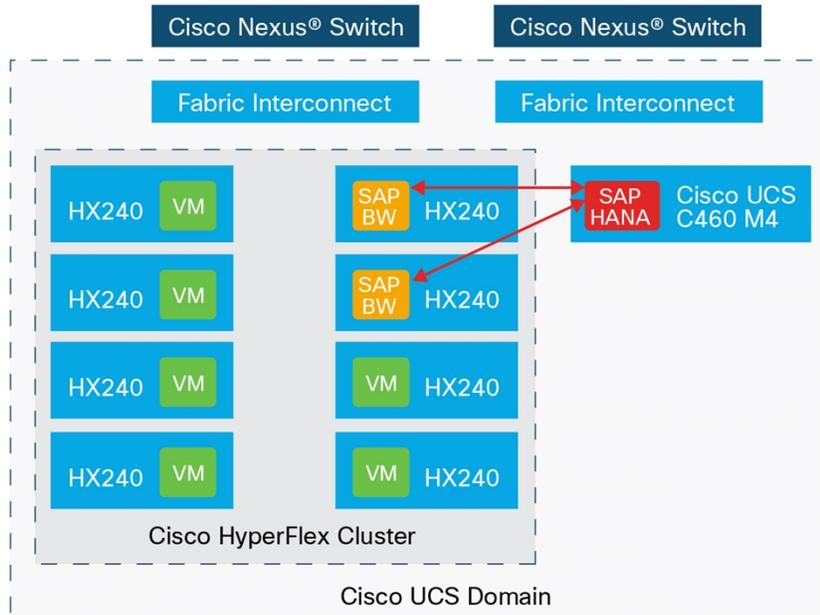
SAP HANA Deployment Options

SAP allows multiple types of SAP HANA database deployments, including single-node scale-up deployments and multinode scale-out deployments. Because all Cisco certified SAP HANA standalone server appliances are designed with internal storage disks, only single-node scale-up SAP HANA systems can be deployed on them. Although multinode SAP HANA systems can be deployed, this deployment methodology requires all SAP HANA nodes to connect to a shared external storage array. Although connectivity to an external array is allowed from the Cisco UCS domain, this deployment method is by default a custom deployment, known as an SAP HANA Tailored Datacenter Integration (TDI), and is outside the scope of this document. Also, multiple SAP HANA appliances can be deployed in an active-passive partnership through SAP HANA System Replication.

SAP System Landscape Sizing

An important first step in designing any SAP system is to properly size the system based on a number of factors:

- Applications to be used
- Number of users
- Relative performance requirements and resource utilization of the users
- Amount of data in the system

To facilitate this process, SAP provides various sizing resources at https://service.sap.com/sizing. One of the primary resources is the SAP Quick Sizer, which helps determine the memory and processing requirements of both the application servers and the database. To achieve accurate results using the Quick Sizer, you need:

- Clear definitions of the various software components that will be used in the SAP system
- Detailed estimates of the number of users and the characteristics of their daily use of the system
- Information about the size and type of data already stored in the database, if one exists

Given the potential complexity of this information, for the purposes of this document Cisco has defined a base workload of an SAP system using SAP BW powered by SAP HANA, a community of about 700 users, and a database with 250 GB of columnar data. For assistance with the proper sizing of your SAP system, please contact your Cisco account manager to discuss the services offered for SAP system landscape planning and sizing.

Tables 1 and 2 show sample Quick Sizer input and output.

Table 1.   Sample Quick Sizer for SAP BW on SAP HANA Input

Item	Sizing
Business planning and simulation 1 users	100
Business planning and simulation 2 users	50
Business planning and simulation 3 users	25
Information consumers	350
Business users	100
Expert users	50
Database column store footprint	250 GB

Table 2.   Sample Quick Sizer for SAP BW on SAP HANA Output

Item	Sizing
BW server SAP Application Performance Standard (SAPS) value required	11,000
BW server Advanced Business Application Programming (ABAP) memory	73 GB
BW server SCU Class	A
SAP HANA database server memory	235 GB
SAP HANA database server storage required	972 GB

Cisco HyperFlex Node Models

Cisco HyperFlex all-flash systems are available in two base models: the Cisco HyperFlex HXAF220c M4 Node and the Cisco HyperFlex HXAF240c M4 Node. The primary difference between the two models lies in the number of disks that can be installed in the converged storage nodes. The HXAF220c server is limited to a maximum of 6 solid state disks (SSDs) to provide storage for the overall capacity of the cluster, whereas the HXAF240c can scale up to 10 SSDs per server, as of HXDP release 2.0.1.a. The caching solid-state disk (SSD) capacity of the two models is identical; each server is built with an 800 GB high endurance SSD for write caching.

Cisco recommends the use of the HXAF240c for any virtualized SAP landscape on the Cisco HyperFlex system that also is intended for running other, additional virtual server workloads (Figure 2), although the HXAF220c is also an excellent choice in environments where the larger disk capacity of the HXAF240c is not needed.

Figure 2.   Cisco HyperFlex HXAF240c M4SX Server
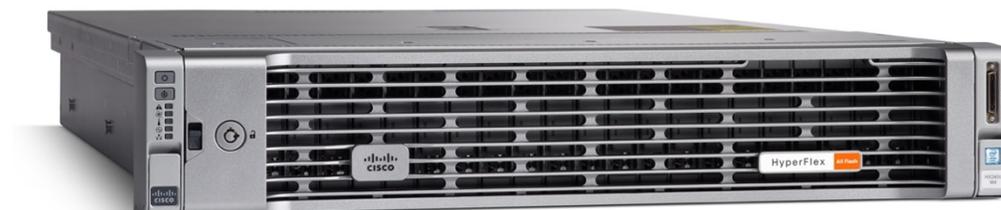
Physical CPU Selection

Processor performance results given by the SAP Quick Sizer need special attention, and for all virtualized server infrastructures picking the appropriate processor model for the servers is important. For the SAP HANA database system, SAP allows only the most current Intel® Xeon® processor E7-8880 v4 and E7-8890 v4 CPUs in a certified appliance. For the Cisco HyperFlex server nodes, any model of Intel Xeon processor E5-2600 v4 CPU is supported. Although this requirement allows a large amount of flexibility, it is important to choose a model that has an optimal combination of core count, clock speed, and cache memory to most effectively run the SAP application servers and other workloads.

For example, the highest-level CPU model, the Intel Xeon processor E5-2699 v4, has 22 cores per socket, but they run at a clock speed of only 2.2 GHz, and this model is nearly as expensive as the Intel Xeon processor E7-8880 v4. At the other end of the spectrum, the Intel Xeon processor E5-2637 v4 runs at a clock speed of 3.5 GHz, but it has only 4 cores per socket. To best accommodate the SAP application server workload along with other virtualized servers, you need to find the Intel Xeon processor E5 CPU model that offers the best combination of cores, clock speed, cache, and price for your needs. Another factor to consider is the SAPS requirement of the application servers reported in the Quick Sizer output. SAPS is a server performance benchmark standard measurement from SAP. In the example described here, a value of 11,000 SAPS is required by the application servers. All current Intel Xeon processor E5 v4 CPUs are capable of delivering that performance.

Table 3 presents four recommendations that each provide an excellent combination of all the attributes needed for a high-performance yet cost-effective Cisco HyperFlex cluster.

Table 3. Sample CPU Model Options for Cisco HyperFlex Cluster

Intel Xeon Processor Model	E5-2650 v4	E5-2660 v4	E5-2680 v4	E5-2690 v4
Number of Cores per Socket	12	14	14	14
Clock Speed	2.2 GHz	2.0 GHz	2.4 GHz	2.6 GHz
Cache Size	30 MB	35 MB	35 MB	35 MB
Estimated SAPS value per CPU	32,950	36,450	42,250	43,950

Note: All processors recommended support the highest DDR4 DIMM clock speed of 2400 MHz.

The SAP Quick Sizer output includes a BW Server SCU Class value. SAP defines SCU as a single computing unit, or a somewhat generic representation of the relative CPU single-thread performance. A value of "A" indicates good single-thread performance, "AA" indicates very good performance, and "AAA" indicates excellent performance. In this value, you can see how SAP is attempting to distinguish between a CPU model with a high core count and a model with a very high clock speed. Models with higher clock speeds will fall into the "AA" or "AAA" categories, because an E5-2637 v4 processor running at 3.5 GHz would obviously be able to finish computing a single thread more quickly than would an E5-2609 v4 running at 1.7 GHz. Certain SAP applications benefit from higher SCU ratings, so you should consider the Quick Sizer output regarding the application server SCU performance when choosing the processor model for the Cisco HyperFlex server nodes.

In the system example described here, the SAP BW application servers need only an "A" SCU rating, and so processors with the highest clock speeds are not necessary. The reference design built for this document uses E5-2690 v4 processors, which provide an excellent combination of core count, clock speed, and cache space, while still remaining highly affordable compared to the E7 models.

Tables 4 and 5 summarize the systems used in this reference design.

Table 4.   Cisco HyperFlex and SAP Landscape Reference System

Item	Description
Server model	Cisco HyperFlex HXAF240c M4SX
Processor model	Intel Xeon processor E5-2690 v4 CPU
RAM per server	512 GB
Caching SSD	One 800 GB SSD per server
Storage disks	Ten 960 GB SSDs per server
Cisco HyperFlex cluster size	8 nodes
Cisco UCS firmware	Release 3.1(2f)
Cisco HyperFlex software version	Release 2.0.1.a
VMware ESXi hypervisor version	Release 6.0 Update 2 Patch 4

Note: HyperFlex 2.0 all-flash systems can be ordered with 960-GB or 3.8-TB capacity SSDs. Choose the part that provides the capacity needed in the cluster. Performance results of the two models of disks is nearly identical.

Table 5.   Cisco Certified SAP HANA Appliance Reference System

Item	Description
Server model	Cisco UCS C460 M4
Processor model	4 Intel Xeon processor E7-8880 v4 CPUs
RAM	512 GB
Storage disks	12 x 1.6-TB SDDs

Virtual Machine CPU Configuration

The selection of the physical CPUs in the Cisco HyperFlex nodes is related to the virtual CPU (vCPU) configuration of the SAP application server virtual machines. In particular, you need to consider the number of physical cores per CPU socket when sizing the virtual machines and configuring the number of vCPUs they use.

Hyperthreading

Intel Xeon processor E5 family CPUs use Intel Hyper-Threading Technology, which enables each CPU core to accept a new thread for processing before the processing of the previous thread is complete. The hyperthreading feature commonly improves overall CPU processing throughput by up to 20 to 25 percent and therefore should be enabled in most circumstances. To enable the hypervisors or operating systems to submit these additional threads to the CPU cores, the underlying hardware presents each core as two cores. For example, to an OS or hypervisor, a CPU that has 10 cores and uses hyperthreading will look as though it is a CPU with 20 cores. Because the processor does not actually have 20 physical cores, these are commonly referred to as logical processors. This approach allows the OS or hypervisor to submit the additional threads for processing, and the server hardware will handle the processing of these threads and return the results.

Non-Uniform Memory Access

Modern Intel processor–based servers are designed using the non-uniform memory access (NUMA) architecture, in which each physical CPU socket has its own set of physical RAM resources to which that CPU has a direct physical connection. Memory that is controlled by another CPU socket is accessible from all other CPU sockets, however, the memory calls have to traverse the server's front-side CPU bus. Although these buses have extremely high bandwidth, this additional hop can add to the latency of the memory calls. A single CPU socket and the memory it directly controls are known as a NUMA node. In most circumstances, a virtual machine will achieve the best performance if all of its vCPUs are scheduled within a single NUMA node, because this approach will make

calls to remote memory on other NUMA nodes unnecessary. If the vCPU settings of a virtual machine cannot fit within a single NUMA node, this virtual machine is called a wide virtual machine.

vCPU Settings

To optimize the vCPU settings of the SAP application servers, you need to consider the number of physical CPU cores, hyperthreading, and the size of the NUMA nodes.
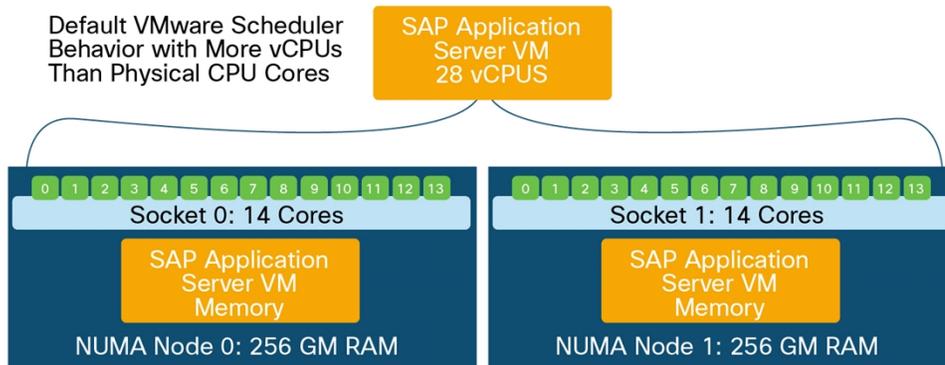
SAP Application Server Virtual Machines

Take particular care to configure all the vCPUs for the SAP application server virtual machines so that they are contained in a single NUMA node. A basic sizing guideline is to configure an application server virtual machine with as many vCPUs as there are cores in a single physical CPU socket. For example, the reference server described here with Intel Xeon processor E5-2690 v4 CPUs has 14 cores per socket. Configuring the application server virtual machines with 14 vCPUs would be a safe initial sizing.

Configuring a virtual machine with no more vCPUs than a single socket has cores helps ensure that all the virtual machine vCPUs and memory will be scheduled within one NUMA node. However, this configuration does not provide the full potential performance of the system because it does not take advantage of hyperthreading. Another guideline for sizing the number of vCPUs for an SAP application server virtual machine is to configure it with as many vCPUs as there are logical processors of a single physical CPU socket. For example, the reference server described here with Intel Xeon processor E5-2690 v4 CPUs has 14 cores per socket with hyperthreading, so there are 28 logical processors per CPU socket. Configuring the application server virtual machines with 28 vCPUs would provide better performance and offers the option to create additional ABAP dialog processes on the server to take advantage of the additional vCPUs.

When a virtual machine is configured with more vCPUs than the number of physical cores of a single NUMA node, the default CPU scheduler in VMware ESXi will schedule those vCPUs across multiple NUMA nodes. The scheduler does not take into account the hyperthreading feature, which allows each core to process two threads. As a result, the vCPUs are not all contained within a single NUMA node, and therefore the memory of the virtual machine is also spread across multiple NUMA nodes. In the example described here, a virtual machine with 28 vCPUs on a server with 14-core physical processors would have the vCPUs scheduled across both physical sockets. This configuration does not result in the best performance for that virtual machine, because it can result in a much higher percentage of remote memory calls across NUMA nodes: for example, a vCPU scheduled on a core of socket 0 may have to access memory pages in the RAM of socket 1 (Figure 3).
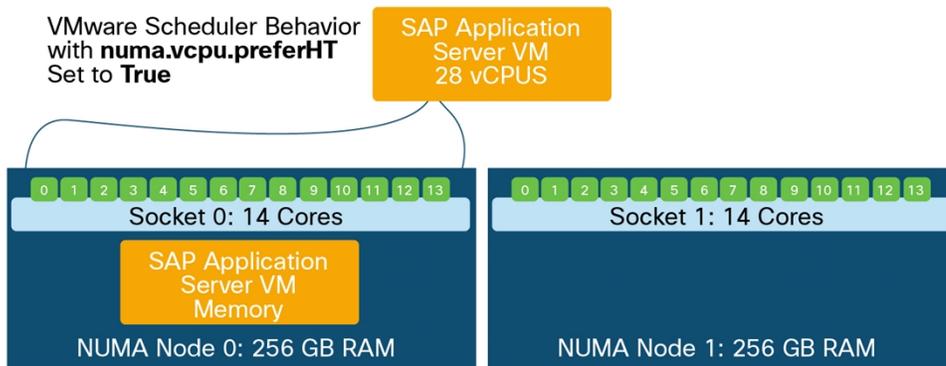
Figure 3. Default CPU Scheduler

Default VMware Scheduler Behavior with More vCPUs Than Physical CPU Cores

SAP Application Server VM 28 vCPUS

| 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 |
Socket 0: 14 Cores
SAP Application Server VM Memory
NUMA Node 0: 256 GM RAM

| 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 |
Socket 1: 14 Cores
SAP Application Server VM Memory
NUMA Node 1: 256 GM RAM

The memory of the single virtual machine is spread more across multiple NUMA nodes, resulting in more remote calls and lower performance.

To change the CPU scheduler behavior so that the scheduler will account for the additional threads possible per core with hyperthreading, set the advanced virtual machine parameter **numa.vcpu.preferHT** to **true** to override the default scheduling behavior. Changing the parameter will force the scheduler to schedule all 28 vCPUs on the same NUMA node, therefore avoiding remote memory calls and increasing performance (Figure 4).

Figure 4. Virtual Machine Setting to Use Hyperthreading

VMware Scheduler Behavior with **numa.vcpu.preferHT** Set to **True**

SAP Application Server VM 28 vCPUS

| 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 |
Socket 0: 14 Cores
SAP Application Server VM Memory
NUMA Node 0: 256 GB RAM

| 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 |
Socket 1: 14 Cores
NUMA Node 1: 256 GB RAM

The memory of the single virtual machine is contained in one NUMA node, eliminating remote calls and increasing performance.

For a virtual machine, you can configure the vCPU settings in a combination of virtual sockets with a number of virtual cores per virtual socket. You must ensure that the total number of vCPUs for an application server virtual machine does not exceed the number of logical processors in a single NUMA node. If the requirements of your SAP NetWeaver application servers do exceed the number of logical processors in a single NUMA node, refer the discussion of wide virtual machines later in this document. Alternatively, you can choose a processor model for the server with a higher physical core count, or simply deploy additional SAP NetWeaver application servers to spread the load across multiple virtual machines.

Even with the **numa.vcpu.preferHT** setting configured as true and the total number of vCPUs set properly, it is possible to configure a virtual machine with a combination of virtual sockets and virtual cores per socket that results in the vCPUs being scheduled across NUMA nodes. The recommended approach is to configure only one core per virtual socket and then configure the number of virtual sockets to match the total number of vCPUs needed by the virtual machine. To help ensure that the NUMA configuration is correct, you can search the vmware.log file of the virtual machine for entries containing "numa". The following example shows the vmware.log file of a virtual machine configured with the desired 28 vCPUs on the reference HXAF240c server, but with two virtual sockets and 14 cores per socket.

```
[root@hx240-01:/vmfs/volumes/ded01cd5-852ae5d5/NW75APP2] cat vmware.log | grep numa
2016-10-18T08:29:20.487Z| vmx| I125: DICT        numa.vcpu.preferHT = "true"
2016-10-18T08:29:20.544Z| vmx| I125: numa: Setting.vcpu.maxPerVirtualNode=14 to
match cpuid.coresPerSocket
2016-10-18T08:29:20.544Z| vmx| I125: numa: VCPU 0: VPD 0 (PPD 0)
2016-10-18T08:29:20.544Z| vmx| I125: numa: VCPU 1: VPD 0 (PPD 0)
2016-10-18T08:29:20.544Z| vmx| I125: numa: VCPU 2: VPD 0 (PPD 0)
2016-10-18T08:29:20.544Z| vmx| I125: numa: VCPU 3: VPD 0 (PPD 0)
2016-10-18T08:29:20.544Z| vmx| I125: numa: VCPU 4: VPD 0 (PPD 0)
2016-10-18T08:29:20.544Z| vmx| I125: numa: VCPU 5: VPD 0 (PPD 0)
2016-10-18T08:29:20.544Z| vmx| I125: numa: VCPU 6: VPD 0 (PPD 0)
2016-10-18T08:29:20.544Z| vmx| I125: numa: VCPU 7: VPD 0 (PPD 0)
2016-10-18T08:29:20.544Z| vmx| I125: numa: VCPU 8: VPD 0 (PPD 0)
2016-10-18T08:29:20.544Z| vmx| I125: numa: VCPU 9: VPD 0 (PPD 0)
2016-10-18T08:29:20.544Z| vmx| I125: numa: VCPU 10: VPD 0 (PPD 0)
2016-10-18T08:29:20.545Z| vmx| I125: numa: VCPU 11: VPD 0 (PPD 0)
2016-10-18T08:29:20.545Z| vmx| I125: numa: VCPU 12: VPD 0 (PPD 0)
2016-10-18T08:29:20.545Z| vmx| I125: numa: VCPU 13: VPD 0 (PPD 0)
2016-10-18T08:29:20.545Z| vmx| I125: numa: VCPU 14: VPD 1 (PPD 1)
2016-10-18T08:29:20.545Z| vmx| I125: numa: VCPU 15: VPD 1 (PPD 1)
2016-10-18T08:29:20.545Z| vmx| I125: numa: VCPU 16: VPD 1 (PPD 1)
2016-10-18T08:29:20.545Z| vmx| I125: numa: VCPU 17: VPD 1 (PPD 1)
2016-10-18T08:29:20.545Z| vmx| I125: numa: VCPU 18: VPD 1 (PPD 1)
2016-10-18T08:29:20.545Z| vmx| I125: numa: VCPU 19: VPD 1 (PPD 1)
2016-10-18T08:29:20.545Z| vmx| I125: numa: VCPU 20: VPD 1 (PPD 1)
2016-10-18T08:29:20.545Z| vmx| I125: numa: VCPU 21: VPD 1 (PPD 1)
2016-10-18T08:29:20.545Z| vmx| I125: numa: VCPU 22: VPD 1 (PPD 1)
2016-10-18T08:29:20.545Z| vmx| I125: numa: VCPU 23: VPD 1 (PPD 1)
2016-10-18T08:29:20.545Z| vmx| I125: numa: VCPU 24: VPD 1 (PPD 1)
```

```
2016-10-18T08:29:20.545Z| vmx| I125: numa: VCPU 25: VPD 1 (PPD 1)

2016-10-18T08:29:20.545Z| vmx| I125: numa: VCPU 26: VPD 1 (PPD 1)

2016-10-18T08:29:20.545Z| vmx| I125: numa: VCPU 27: VPD 1 (PPD 1)

2016-10-18T08:29:20.545Z| vmx| I125: numaHost: 2 virtual nodes, 2 virtual sockets, 2
physical domains
```

As can be seen in the log, the configuration resulted in 14 vCPUs being scheduled on one physical NUMA node, and the other 14 vCPUs scheduled on the second physical NUMA node, so this is not an ideal configuration. Changing the settings of this virtual machine to one core per virtual socket and 28 virtual sockets yields the following vmware.log file entries:

```
[root@hx240-01:/vmfs/volumes/ded01cd5-852ae5d5/NW75APP2] cat vmware.log | grep numa

2016-11-01T19:13:34.895Z| vmx| I125: DICT          numa.vcpu.preferHT = "true"

2016-11-01T19:13:34.954Z| vmx| A100: ConfigDB: Setting
numa.autosize.vcpu.maxPerVirtualNode = "28"

2016-11-01T19:13:34.954Z| vmx| A100: ConfigDB: Setting numa.autosize.cookie =
"280001"

2016-11-01T19:13:34.959Z| vmx| I125: numa: VCPU 0: VPD 0 (PPD 0)

2016-11-01T19:13:34.959Z| vmx| I125: numa: VCPU 1: VPD 0 (PPD 0)

2016-11-01T19:13:34.959Z| vmx| I125: numa: VCPU 2: VPD 0 (PPD 0)

2016-11-01T19:13:34.959Z| vmx| I125: numa: VCPU 3: VPD 0 (PPD 0)

2016-11-01T19:13:34.959Z| vmx| I125: numa: VCPU 4: VPD 0 (PPD 0)

2016-11-01T19:13:34.959Z| vmx| I125: numa: VCPU 5: VPD 0 (PPD 0)

2016-11-01T19:13:34.959Z| vmx| I125: numa: VCPU 6: VPD 0 (PPD 0)

2016-11-01T19:13:34.959Z| vmx| I125: numa: VCPU 7: VPD 0 (PPD 0)

2016-11-01T19:13:34.959Z| vmx| I125: numa: VCPU 8: VPD 0 (PPD 0)

2016-11-01T19:13:34.959Z| vmx| I125: numa: VCPU 9: VPD 0 (PPD 0)

2016-11-01T19:13:34.959Z| vmx| I125: numa: VCPU 10: VPD 0 (PPD 0)

2016-11-01T19:13:34.959Z| vmx| I125: numa: VCPU 11: VPD 0 (PPD 0)

2016-11-01T19:13:34.959Z| vmx| I125: numa: VCPU 12: VPD 0 (PPD 0)

2016-11-01T19:13:34.959Z| vmx| I125: numa: VCPU 13: VPD 0 (PPD 0)

2016-11-01T19:13:34.959Z| vmx| I125: numa: VCPU 14: VPD 0 (PPD 0)

2016-11-01T19:13:34.959Z| vmx| I125: numa: VCPU 15: VPD 0 (PPD 0)

2016-11-01T19:13:34.959Z| vmx| I125: numa: VCPU 16: VPD 0 (PPD 0)

2016-11-01T19:13:34.959Z| vmx| I125: numa: VCPU 17: VPD 0 (PPD 0)

2016-11-01T19:13:34.959Z| vmx| I125: numa: VCPU 18: VPD 0 (PPD 0)

2016-11-01T19:13:34.959Z| vmx| I125: numa: VCPU 19: VPD 0 (PPD 0)

2016-11-01T19:13:34.959Z| vmx| I125: numa: VCPU 20: VPD 0 (PPD 0)
```

```
2016-11-01T19:13:34.959Z| vmx| I125: numa: VCPU 21: VPD 0 (PPD 0)

2016-11-01T19:13:34.959Z| vmx| I125: numa: VCPU 22: VPD 0 (PPD 0)

2016-11-01T19:13:34.959Z| vmx| I125: numa: VCPU 23: VPD 0 (PPD 0)

2016-11-01T19:13:34.959Z| vmx| I125: numa: VCPU 24: VPD 0 (PPD 0)

2016-11-01T19:13:34.959Z| vmx| I125: numa: VCPU 25: VPD 0 (PPD 0)

2016-11-01T19:13:34.959Z| vmx| I125: numa: VCPU 26: VPD 0 (PPD 0)

2016-11-01T19:13:34.959Z| vmx| I125: numa: VCPU 27: VPD 0 (PPD 0)

2016-11-01T19:13:34.959Z| vmx| I125: numaHost: 1 virtual nodes, 28 virtual sockets,
1 physical domains
```

This configuration would provide the best possible performance of the CPU and RAM for this virtual machine, because the 28 vCPUs are scheduled on a single physical NUMA node.

Wide Virtual Machines

In some situations, you may need to configure a virtual machine with more vCPUs than there are logical processors in a single NUMA node. To make the operating system in the virtual machine aware of the physical NUMA architecture, you must enable the VMware vNUMA feature. The vNUMA feature is turned on for a virtual machine by default whenever the virtual machine is configured with more than eight vCPUs. For example, the reference server described here uses Intel Xeon processor E5-2690 v4 CPUs with 14 cores each and hyperthreading. Each CPU therefore has 28 logical processors, and with two sockets a total of 56 logical processors are available. To get the best possible performance for a virtual machine on these servers you would configure the virtual machine with 56 vCPUs. Because this number of vCPUs is larger than the size of a single physical NUMA node, these virtual machines would be considered wide virtual machines.

Configuring the virtual machine setting **numa.vcpu.preferHT** as **true** is again recommended. You also should configure only one core per virtual socket, and you should configure the number of virtual sockets to match the total number of vCPUs needed by the virtual machine as the initial setting. However, in the case of a wide virtual machine, which spans NUMA nodes, this setting will not expose the true physical underlying NUMA architecture to the OS running on the virtual machine; therefore, it will not be the best-performing configuration. The following example shows the vmware.log file of a virtual machine configured with 56 virtual sockets and one core per socket.

```
[root@hx240-02:/vmfs/volumes/ded01cd5-852ae5d5/NW75APP3] cat vmware.log | grep numa

2016-10-31T20:49:23.429Z| vmx| I125: DICT        numa.vcpu.preferHT = "true"

2016-10-31T20:49:23.507Z| vmx| A100: ConfigDB: Setting
numa.autosize.vcpu.maxPerVirtualNode = "28"

2016-10-31T20:49:23.507Z| vmx| A100: ConfigDB: Setting numa.autosize.cookie =
"560001"

2016-10-31T20:49:23.512Z| vmx| I125: numa: Exposing multicore topology with
cpuid.coresPerSocket = 28 is suggested for best performance

2016-10-31T20:49:23.512Z| vmx| I125: numa: VCPU 0: VPD 0 (PPD 0)

2016-10-31T20:49:23.512Z| vmx| I125: numa: VCPU 1: VPD 0 (PPD 0)

2016-10-31T20:49:23.512Z| vmx| I125: numa: VCPU 2: VPD 0 (PPD 0)

2016-10-31T20:49:23.512Z| vmx| I125: numa: VCPU 3: VPD 0 (PPD 0)
```

```
2016-10-31T20:49:23.512Z| vmx| I125: numa: VCPU 4: VPD 0 (PPD 0)
2016-10-31T20:49:23.512Z| vmx| I125: numa: VCPU 5: VPD 0 (PPD 0)
2016-10-31T20:49:23.512Z| vmx| I125: numa: VCPU 6: VPD 0 (PPD 0)
2016-10-31T20:49:23.512Z| vmx| I125: numa: VCPU 7: VPD 0 (PPD 0)
2016-10-31T20:49:23.512Z| vmx| I125: numa: VCPU 8: VPD 0 (PPD 0)
2016-10-31T20:49:23.512Z| vmx| I125: numa: VCPU 9: VPD 0 (PPD 0)
2016-10-31T20:49:23.512Z| vmx| I125: numa: VCPU 10: VPD 0 (PPD 0)
2016-10-31T20:49:23.512Z| vmx| I125: numa: VCPU 11: VPD 0 (PPD 0)
2016-10-31T20:49:23.512Z| vmx| I125: numa: VCPU 12: VPD 0 (PPD 0)
2016-10-31T20:49:23.512Z| vmx| I125: numa: VCPU 13: VPD 0 (PPD 0)
2016-10-31T20:49:23.512Z| vmx| I125: numa: VCPU 14: VPD 0 (PPD 0)
2016-10-31T20:49:23.512Z| vmx| I125: numa: VCPU 15: VPD 0 (PPD 0)
2016-10-31T20:49:23.512Z| vmx| I125: numa: VCPU 16: VPD 0 (PPD 0)
2016-10-31T20:49:23.512Z| vmx| I125: numa: VCPU 17: VPD 0 (PPD 0)
2016-10-31T20:49:23.512Z| vmx| I125: numa: VCPU 18: VPD 0 (PPD 0)
2016-10-31T20:49:23.512Z| vmx| I125: numa: VCPU 19: VPD 0 (PPD 0)
2016-10-31T20:49:23.512Z| vmx| I125: numa: VCPU 20: VPD 0 (PPD 0)
2016-10-31T20:49:23.512Z| vmx| I125: numa: VCPU 21: VPD 0 (PPD 0)
2016-10-31T20:49:23.512Z| vmx| I125: numa: VCPU 22: VPD 0 (PPD 0)
2016-10-31T20:49:23.512Z| vmx| I125: numa: VCPU 23: VPD 0 (PPD 0)
2016-10-31T20:49:23.512Z| vmx| I125: numa: VCPU 24: VPD 0 (PPD 0)
2016-10-31T20:49:23.512Z| vmx| I125: numa: VCPU 25: VPD 0 (PPD 0)
2016-10-31T20:49:23.512Z| vmx| I125: numa: VCPU 26: VPD 0 (PPD 0)
2016-10-31T20:49:23.512Z| vmx| I125: numa: VCPU 27: VPD 0 (PPD 0)
2016-10-31T20:49:23.512Z| vmx| I125: numa: VCPU 28: VPD 1 (PPD 1)
2016-10-31T20:49:23.512Z| vmx| I125: numa: VCPU 29: VPD 1 (PPD 1)
2016-10-31T20:49:23.512Z| vmx| I125: numa: VCPU 30: VPD 1 (PPD 1)
2016-10-31T20:49:23.512Z| vmx| I125: numa: VCPU 31: VPD 1 (PPD 1)
2016-10-31T20:49:23.512Z| vmx| I125: numa: VCPU 32: VPD 1 (PPD 1)
2016-10-31T20:49:23.512Z| vmx| I125: numa: VCPU 33: VPD 1 (PPD 1)
2016-10-31T20:49:23.512Z| vmx| I125: numa: VCPU 34: VPD 1 (PPD 1)
2016-10-31T20:49:23.512Z| vmx| I125: numa: VCPU 35: VPD 1 (PPD 1)
2016-10-31T20:49:23.512Z| vmx| I125: numa: VCPU 36: VPD 1 (PPD 1)
2016-10-31T20:49:23.512Z| vmx| I125: numa: VCPU 37: VPD 1 (PPD 1)
2016-10-31T20:49:23.512Z| vmx| I125: numa: VCPU 38: VPD 1 (PPD 1)
```

```
2016-10-31T20:49:23.512Z| vmx| I125: numa: VCPU 39: VPD 1 (PPD 1)

2016-10-31T20:49:23.512Z| vmx| I125: numa: VCPU 40: VPD 1 (PPD 1)

2016-10-31T20:49:23.512Z| vmx| I125: numa: VCPU 41: VPD 1 (PPD 1)

2016-10-31T20:49:23.512Z| vmx| I125: numa: VCPU 42: VPD 1 (PPD 1)

2016-10-31T20:49:23.512Z| vmx| I125: numa: VCPU 43: VPD 1 (PPD 1)

2016-10-31T20:49:23.512Z| vmx| I125: numa: VCPU 44: VPD 1 (PPD 1)

2016-10-31T20:49:23.512Z| vmx| I125: numa: VCPU 45: VPD 1 (PPD 1)

2016-10-31T20:49:23.512Z| vmx| I125: numa: VCPU 46: VPD 1 (PPD 1)

2016-10-31T20:49:23.512Z| vmx| I125: numa: VCPU 47: VPD 1 (PPD 1)

2016-10-31T20:49:23.512Z| vmx| I125: numa: VCPU 48: VPD 1 (PPD 1)

2016-10-31T20:49:23.512Z| vmx| I125: numa: VCPU 49: VPD 1 (PPD 1)

2016-10-31T20:49:23.512Z| vmx| I125: numa: VCPU 50: VPD 1 (PPD 1)

2016-10-31T20:49:23.512Z| vmx| I125: numa: VCPU 51: VPD 1 (PPD 1)

2016-10-31T20:49:23.512Z| vmx| I125: numa: VCPU 52: VPD 1 (PPD 1)

2016-10-31T20:49:23.512Z| vmx| I125: numa: VCPU 53: VPD 1 (PPD 1)

2016-10-31T20:49:23.512Z| vmx| I125: numa: VCPU 54: VPD 1 (PPD 1)

2016-10-31T20:49:23.512Z| vmx| I125: numa: VCPU 55: VPD 1 (PPD 1)

2016-10-31T20:49:23.512Z| vmx| I125: numaHost: 2 virtual nodes, 56 virtual sockets,
2 physical domains
```

As highlighted in the log, while the number of physical NUMA nodes in use is correct at two, ESXi recommends that the number of cores per virtual socket be changed to 28 to expose the true NUMA architecture to the OS and achieve the best performance. This change means that for the best performance of a wide virtual machine, you need to set the number of virtual sockets to match the number of physical CPU sockets in the server, and to set the number of cores per socket to match the number of logical processors in each physical NUMA node (Figure 5).

Figure 5.   Wide Virtual Machine CPU Scheduling
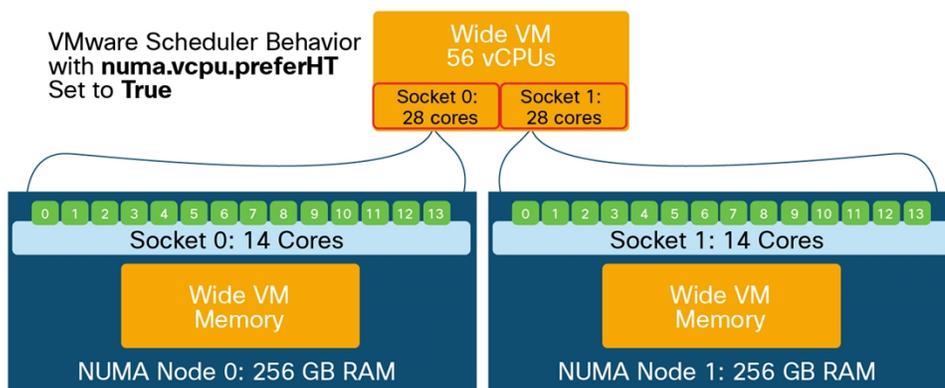
The memory of the single virtual machine is spread across both NUMA nodes with full OS awareness, resulting in the best performance.

The following example shows a vmware.log file after the virtual machine is changed to two virtual sockets with 28 cores per socket.

```
[root@hx240-02:/vmfs/volumes/ded01cd5-852ae5d5/NW75APP3] cat vmware.log | grep numa
2016-10-31T20:52:38.731Z| vmx| I125: DICT        numa.vcpu.preferHT = "true"
2016-10-31T20:52:38.731Z| vmx| I125: DICT numa.autosize.vcpu.maxPerVirtualNode =
"28"
2016-10-31T20:52:38.731Z| vmx| I125: DICT     numa.autosize.cookie = "560001"
2016-10-31T20:52:38.812Z| vmx| I125: numa: Setting.vcpu.maxPerVirtualNode=28 to
match cpuid.coresPerSocket
2016-10-31T20:52:38.812Z| vmx| I125: numa: VCPU 0: VPD 0 (PPD 0)
2016-10-31T20:52:38.812Z| vmx| I125: numa: VCPU 1: VPD 0 (PPD 0)
2016-10-31T20:52:38.812Z| vmx| I125: numa: VCPU 2: VPD 0 (PPD 0)
2016-10-31T20:52:38.812Z| vmx| I125: numa: VCPU 3: VPD 0 (PPD 0)
2016-10-31T20:52:38.812Z| vmx| I125: numa: VCPU 4: VPD 0 (PPD 0)
2016-10-31T20:52:38.812Z| vmx| I125: numa: VCPU 5: VPD 0 (PPD 0)
2016-10-31T20:52:38.812Z| vmx| I125: numa: VCPU 6: VPD 0 (PPD 0)
2016-10-31T20:52:38.812Z| vmx| I125: numa: VCPU 7: VPD 0 (PPD 0)
2016-10-31T20:52:38.812Z| vmx| I125: numa: VCPU 8: VPD 0 (PPD 0)
2016-10-31T20:52:38.812Z| vmx| I125: numa: VCPU 9: VPD 0 (PPD 0)
2016-10-31T20:52:38.812Z| vmx| I125: numa: VCPU 10: VPD 0 (PPD 0)
2016-10-31T20:52:38.812Z| vmx| I125: numa: VCPU 11: VPD 0 (PPD 0)
2016-10-31T20:52:38.812Z| vmx| I125: numa: VCPU 12: VPD 0 (PPD 0)
2016-10-31T20:52:38.812Z| vmx| I125: numa: VCPU 13: VPD 0 (PPD 0)
2016-10-31T20:52:38.812Z| vmx| I125: numa: VCPU 14: VPD 0 (PPD 0)
2016-10-31T20:52:38.812Z| vmx| I125: numa: VCPU 15: VPD 0 (PPD 0)
2016-10-31T20:52:38.812Z| vmx| I125: numa: VCPU 16: VPD 0 (PPD 0)
2016-10-31T20:52:38.812Z| vmx| I125: numa: VCPU 17: VPD 0 (PPD 0)
2016-10-31T20:52:38.812Z| vmx| I125: numa: VCPU 18: VPD 0 (PPD 0)
2016-10-31T20:52:38.812Z| vmx| I125: numa: VCPU 19: VPD 0 (PPD 0)
2016-10-31T20:52:38.812Z| vmx| I125: numa: VCPU 20: VPD 0 (PPD 0)
2016-10-31T20:52:38.812Z| vmx| I125: numa: VCPU 21: VPD 0 (PPD 0)
2016-10-31T20:52:38.812Z| vmx| I125: numa: VCPU 22: VPD 0 (PPD 0)
2016-10-31T20:52:38.812Z| vmx| I125: numa: VCPU 23: VPD 0 (PPD 0)
2016-10-31T20:52:38.812Z| vmx| I125: numa: VCPU 24: VPD 0 (PPD 0)
2016-10-31T20:52:38.812Z| vmx| I125: numa: VCPU 25: VPD 0 (PPD 0)
2016-10-31T20:52:38.812Z| vmx| I125: numa: VCPU 26: VPD 0 (PPD 0)
```

```
2016-10-31T20:52:38.812Z| vmx| I125: numa: VCPU 27: VPD 0 (PPD 0)

2016-10-31T20:52:38.812Z| vmx| I125: numa: VCPU 28: VPD 1 (PPD 1)

2016-10-31T20:52:38.812Z| vmx| I125: numa: VCPU 29: VPD 1 (PPD 1)

2016-10-31T20:52:38.812Z| vmx| I125: numa: VCPU 30: VPD 1 (PPD 1)

2016-10-31T20:52:38.812Z| vmx| I125: numa: VCPU 31: VPD 1 (PPD 1)

2016-10-31T20:52:38.812Z| vmx| I125: numa: VCPU 32: VPD 1 (PPD 1)

2016-10-31T20:52:38.812Z| vmx| I125: numa: VCPU 33: VPD 1 (PPD 1)

2016-10-31T20:52:38.812Z| vmx| I125: numa: VCPU 34: VPD 1 (PPD 1)

2016-10-31T20:52:38.812Z| vmx| I125: numa: VCPU 35: VPD 1 (PPD 1)

2016-10-31T20:52:38.812Z| vmx| I125: numa: VCPU 36: VPD 1 (PPD 1)

2016-10-31T20:52:38.812Z| vmx| I125: numa: VCPU 37: VPD 1 (PPD 1)

2016-10-31T20:52:38.812Z| vmx| I125: numa: VCPU 38: VPD 1 (PPD 1)

2016-10-31T20:52:38.812Z| vmx| I125: numa: VCPU 39: VPD 1 (PPD 1)

2016-10-31T20:52:38.812Z| vmx| I125: numa: VCPU 40: VPD 1 (PPD 1)

2016-10-31T20:52:38.812Z| vmx| I125: numa: VCPU 41: VPD 1 (PPD 1)

2016-10-31T20:52:38.812Z| vmx| I125: numa: VCPU 42: VPD 1 (PPD 1)

2016-10-31T20:52:38.812Z| vmx| I125: numa: VCPU 43: VPD 1 (PPD 1)

2016-10-31T20:52:38.812Z| vmx| I125: numa: VCPU 44: VPD 1 (PPD 1)

2016-10-31T20:52:38.812Z| vmx| I125: numa: VCPU 45: VPD 1 (PPD 1)

2016-10-31T20:52:38.812Z| vmx| I125: numa: VCPU 46: VPD 1 (PPD 1)

2016-10-31T20:52:38.812Z| vmx| I125: numa: VCPU 47: VPD 1 (PPD 1)

2016-10-31T20:52:38.812Z| vmx| I125: numa: VCPU 48: VPD 1 (PPD 1)

2016-10-31T20:52:38.812Z| vmx| I125: numa: VCPU 49: VPD 1 (PPD 1)

2016-10-31T20:52:38.812Z| vmx| I125: numa: VCPU 50: VPD 1 (PPD 1)

2016-10-31T20:52:38.812Z| vmx| I125: numa: VCPU 51: VPD 1 (PPD 1)

2016-10-31T20:52:38.812Z| vmx| I125: numa: VCPU 52: VPD 1 (PPD 1)

2016-10-31T20:52:38.812Z| vmx| I125: numa: VCPU 53: VPD 1 (PPD 1)

2016-10-31T20:52:38.812Z| vmx| I125: numa: VCPU 54: VPD 1 (PPD 1)

2016-10-31T20:52:38.812Z| vmx| I125: numa: VCPU 55: VPD 1 (PPD 1)

2016-10-31T20:52:38.812Z| vmx| I125: numaHost: 2 virtual nodes, 2 virtual sockets, 2
physical domains
```

Now the virtual machine has two virtual NUMA nodes, each with 28 virtual CPUs, scheduled on the two physical NUMA nodes, and there are no further recommendations from ESXi. The OS running in the virtual machine will see an accurate mapping of the virtual CPU resources that corresponds with the true underlying physical NUMA architecture.

Storage Controller Virtual Machines

The storage controller virtual machines (SCVMs) running on each Cisco HyperFlex node are each configured with a mandatory CPU reservation of 10,800 MHz (10.8 GHz). During normal operation of a Cisco HyperFlex system, the processor demands of the SCVMs are typically less than 50 percent of this value. The ESXi hosts will not perform any throttling of CPU use unless the system nears 100 percent overall CPU use. Only in those circumstances would the host set aside the 10.8 GHz of potential CPU performance, because this performance is otherwise guaranteed to the SCVMs. In most circumstances, the CPU time of the hosts is not being taxed to the limit, and therefore the reservation and CPU use of the SCVMs should have little overall impact on the performance of the virtual machines.

Memory Configuration

Memory sizing also plays a critical role in the overall sizing of an SAP system. In particular, SAP HANA is an in-memory database platform, in which the entire contents of the database will remain cached in memory while the system is operating. The SAP Quick Sizer output will provide details about the amount of memory recommended for operating the database without exhausting the system. The amount given for the example system described here is 235 GB. This value should be considered the minimum amount for the server that will run the SAP HANA database, because additional RAM can be consumed for scratch space and query processing.

Note the following guidelines when sizing the memory of the virtual machines in the Cisco HyperFlex system:

- Although the maximum amount of RAM that can be configured for a virtual machine in ESX 6.0 is 4 TB, the maximum amount of RAM that can be installed in a Cisco HyperFlex HXAF240c server is 1.5 TB.
- As noted by VMware, a portion of RAM should be set aside for the proper operation of ESXi itself. VMware recommends that you set aside approximately 3 or 4 percent. On a 1.5-TB server, this amount would further decrease the available physical memory by up to 62 GB, reducing the maximum virtual machine RAM to 1474 GB.
- All Cisco HyperFlex converged storage nodes run an SCVM that requires its own reserved virtual RAM. On the HXAF240c server, the Cisco HyperFlex SCVM has a memory reservation of 72 GB. This amount further reduces the maximum virtual machine RAM for a virtual machine to about 1400 GB before memory is overcommitted.
- For the best performance, the amount of RAM configured for a single application server virtual machine should not exceed the amount of RAM in a single NUMA node.

For the reference design discussed here, the HXAF240c servers have 512 GB of RAM each. For the NetWeaver Application Servers, the output of the SAP Quick Sizer provides the primary guidance regarding the amount of RAM that should be configured for the application server virtual machines. In the example here, the Quick Sizer specified only 73 GB, so for this reference design the application server virtual machines are configured with 128 GB of RAM. For the best performance, the amount of RAM configured for a single application server virtual machine should not exceed the amount of RAM in a single NUMA node. With the proper amount of memory and the proper vCPU configuration, all memory and CPU processing will take place inside a single NUMA node, helping ensure the best possible application server performance.

For a standalone SAP HANA appliance server, the memory configuration is fixed in size according to the Certified SAP HANA Hardware Directory entries. Various valid configurations exist with two to four CPUs and 512 GB to 4 TB of RAM, depending on the applications being used. Consult the Certified SAP HANA Hardware Directory to pick the configuration that supports the applications you are using and that you have enough RAM to contain the full SAP HANA database. For this reference design, the SAP Quick Sizer specified a database server with 235 GB of RAM. Therefore, the server described here is equipped with 512 GB of RAM.

Monitoring NUMA Memory Statistics

You can use the ESXTOP monitoring tool on Cisco HyperFlex hosts to monitor the NUMA statistics of the virtual machines. To monitor the NUMA statistics, do the following:

1. Using Secure Shell (SSH), log on to the ESXi host that is running the virtual machine whose statistics you want to monitor.

2. Type **esxtop** and press Enter.

3. Type the letter **M**, then the letter **F**, and then the letter **G**; then press Enter.

4. Monitor the statistics. When you are done, type the letter **Q**.

In the statistics, the field N%L indicates the percentage of NUMA memory that is operating local to the scheduled CPUs of the virtual machine (Figure 6). In general, this value should be 95 percent or higher, indicating that the virtual machine does not need to make many calls to a remote physical NUMA node for memory contents.

Figure 6.   ESXTOP NUMA Statistics

Storage Capacity Sizing

The overall storage capacity within a Cisco HyperFlex cluster is determined by the number of converged storage nodes in the cluster, and the number and size of the storage capacity disks in each node. The HXAF240 model server can be ordered with as few as 6 capacity disks, but can be configured with up to 10 capacity disks per node, using either 960-GB or 3.8-TB SSDs. Table 6 shows the overall cluster capacity for several scenarios.

Table 6.   Cisco HyperFlex Cluster Usable Capacity

Cisco HyperFlex Cluster Capacity (tebibits TiB)			
Number of Nodes Disks per Node	4	6	8
6 (960 GB/3.8 TB)	6.4 TiB / 25.7 TiB	9.6 TiB / 38.6 TiB	12.8 TiB / 51.4 TiB
10 (960 GB/3.8 TB)	10.7 TiB / 42.8 TiB	16.1 TiB / 64.3 TiB	21.4 TiB / 85.7 TiB

Note: All capacities are given at a replication factor of 3.

Storage Performance Considerations

In the original release of Cisco HyperFlex systems, the server nodes offered a combination of a caching SSD and traditional hard disk drives (HDDs) which is known as a hybrid configuration. The caching SSD serviced all write operations along with caching recently read blocks, in order to improve performance by mitigating the frequency that data must be retrieved from the HDDs. With HyperFlex 2.0 all-flash systems, the capacity HDDs have been replaced with SSDs, and all read requests are serviced by those SSDs directly. The caching SSD no longer maintains a read cache and now only acts as a high endurance write caching device, however some read caching does occur within the RAM of the SCVM.

SAP HANA is referred to as an in-memory database, however, this term is slightly misleading. In reality, SAP HANA operates with an extremely large amount of RAM, and it caches the data in RAM until either the entire database is cached or the amount of RAM is exhausted. Upon a start or restart of the SAP HANA database, only small portions of the overall database are cached in RAM by default, along with additional tables that the SAP database administrator (DBA) or Basis administrator designates to automatically load into memory. The remainder of the data is cached in the RAM of the SAP HANA server when a client requests to read the data. Therefore, as a result of this initial load sequence, upon a start or restart of the database, the initial performance of the database will be slightly lower. As the SAP HANA database remains online, data which is read will be cached in the RAM of the SAP HANA server, and performance will quickly increase to its maximum potential.

To optimize this initial ramp-up time, the dedicated SAP HANA appliance is configured with enough disks to pass the certification tests administered by SAP to qualify it as a valid and supportable platform for SAP HANA. The storage configuration of the SAP HANA appliance cannot be modified.

VMware Configuration and Recommendations
This section presents recommendations for configuring VMware.

Virtual Machine Placement
Given the frequently large sizes of SAP Netweaver application virtual machines, with virtual machines potentially consuming the entire RAM and CPU resources of the full node, you should use VMware DRS anti-affinity rules to keep the SAP virtual machines separate from the other virtual machines in the cluster. Normally, DRS would recognize the large size of a virtual machine and place it on a host by itself, but the use of DRS rules can help guarantee that this separation is always performed. This action helps ensure that a wide virtual machine will have no contention for CPU or memory resources. A DRS rule can also help mitigate the potential for losing multiple SAP application server virtual machines at one time as a result of a single host failure.

SAP System Monitoring
The SAP Host Agent, a software component that is installed by default on all NetWeaver Application Server virtual machines, can also view extended system information about the virtualization hosts from within the virtual machines. This feature is called SAP System Monitoring and is detailed in SAP Note 1606643. SAP System Monitoring is a mandatory feature for SAP support in virtualized environments. To enable the feature, configure the following advanced settings in the VMware vSphere Client or vSphere Web Client:

- Verify that the virtual machines are running VMware Tools.

Note: If generic Linux open-vm-tools are used, some libraries have different names. This issue can be fixed by entering the following Linux command to link the correct libraries:

```
ln -s /usr/lib64/libguestlib.so.0 /usr/lib64/libvmGuestLib.so
```

- On each ESXi host, change the advanced setting **Misc.GuestLibAllowHostInfo** to **1**.
- On each SAP virtual machine, insert the advanced setting **tools.guestlib.enableHostInfo** set to **true**.

Cisco HyperFlex Configuration and Recommendations

For the best level of redundancy for the data in the Cisco HyperFlex distributed file system, you should install the Cisco HyperFlex cluster using a replication factor of 3. This setting creates three copies of all data written to the cluster and allows simultaneous failures of multiple components throughout the cluster, even failure of entire nodes, without data loss. One of the most frequently failed components of a server or storage array is the legacy spinning hard drive (HDD). Studies have shown that solid state disks (SSDs) are anywhere from 2-10 times as reliable as HDDs on an annual replacement basis. Due to this, configuring the HyperFlex cluster to use replication factor 2 is a much more viable option in all-flash configurations. Consideration should be given to using RF2 if the extra redundancy is not considered to be necessary, because RF2 does benefit from 15-40% reduced write latency performance, plus it increases the usable capacity of the HyperFlex cluster.

As an option during the installation of a Cisco HyperFlex cluster, you can optimize the cluster for a virtual desktop infrastructure (VDI) deployment. However, this feature modifies the caching amounts and behaviors of the cluster and is not ideal for an environment with virtualized server application workloads. Therefore, optimizing Cisco HyperFlex for VDI deployment is not recommended.

Cisco UCS Configuration and Recommendations

Connect the Cisco UCS C460 M4 server to the Cisco UCS fabric interconnects and perform discovery. If necessary, approve the pending activities to allow the firmware of all the server components to upgrade.

Cisco UCS Organization

You should create a new Cisco UCS organization within Cisco UCS Manager to keep policies, templates, and service profiles used by the Cisco HyperFlex cluster separate from those of other systems. Create all of the following recommended policies and profiles within this new Cisco UCS organization.

BIOS Policy

Create a custom BIOS policy that disables all CPU C-state power savings settings.

Storage Policies

Create a custom disk group policy that specifies the following settings:

- RAID 5 striped parity
- Use of all available SSDs with no hot spares
- Strip size of 256 KB
- Access policy: Read-write
- Read policy: Read ahead
- Write policy: Always write back
- I/O policy: Cached

Storage Profiles

Create a custom storage profile to be used by the service profile for the server, with the following settings:

- Create a local logical unit number (LUN) definition using the disk group defined earlier.
- Set the size to 1 GB and check the box to allow the LUN to expand to use all available space.

Service Profile

Create the service profile for the Cisco UCS C460 M4 server for SAP HANA and make sure that it uses the BIOS policy and the storage profile created earlier.

Multiple SAP HANA Instances

If your use case requires multiple SAP HANA database instances, then you must follow these guidelines:

- Although SAP does offer support for multiple independent SAP HANA installations (multiple system IDs [SIDs]) installed on the same appliance, each database instance should have its memory global allocation limit set so that the two instances do not overcommit the physical RAM resources of the server.
- To retain full production support from SAP of an SAP HANA multiple SID installation, each HANA instance must be installed to its own independent certified storage subsystem. For example, the first HANA instance can be installed using the internal disks of the Cisco UCS C460 M4 server, and additional HANA instances would need to be installed using a different set of disks. Most often, this additional installation will use storage from an external array via NFS or fibre channel (FC).
- As an alternative to installing multiple database instances on a single server (multiple SIDs), consider using the SAP HANA Multi-Tenant Database Containers (MDC) feature, which allows multiple independent databases to operate under the management of a single installation of SAP HANA.
- A final alternative is to deploy additional standalone Cisco certified SAP HANA appliances.

Conclusion

This Cisco solution addresses the needs of many companies that use SAP in their data centers. It presents the unique considerations and recommendations for running an SAP landscape in a Cisco HyperFlex cluster while staying aligned with established practices. By following the guidelines in this solution, your SAP landscape can achieve its best possible performance in a Cisco HyperFlex system and take advantage of the ease of use and rapid deployment provided by the Cisco HyperFlex solution.

For More Information

The following documents, available online, provide additional relevant information. If you do not have access to a document, please contact your Cisco representative.

- Cisco HyperFlex Technical Support Documentation:
- Cisco UCS Scale-Up Solution for SAP HANA – Design and Deploy
- VMware Best Practices and Recommendation for Scale-Up Deployments of SAP HANA on VMware vSphere
- SAP on VMware Best Practices
- Performance Best Practices for VMware vSphere 6.0
- SAP Certified HANA Hardware Directory – Appliances

- [SAP on VMware vSphere Wiki](#)

- [SAP HANA on VMware vSphere Wiki](#)

- [SAP HANA TDI FAQ](#)

- [SAP HANA Help Portal](#)

SAP Notes

The following SAP Notes provide the latest up-to-date information regarding SAP HANA settings, requirements, and OS-specific settings to optimize the system. A valid logon with SAP support is required to view the content of the notes.

SAP Notes	
Note Number	**Description**
2240716	Recommended OS settings for SLES 11 and SLES for SAP Applications 11 SP4
2247020	Recommended OS settings for RHEL 6.7
2205917	Recommended OS settings for SLES 12 and SLES for SAP Applications 12
2292690	Recommended OS settings for RHEL 7.2
2235581	SAP HANA: Supported Operating Systems
1492000	General Support Statement for Virtual Environments
1681092	Multiple SAP HANA DBMSs (SIDs) on one SAP HANA system