# Compute and Storage Infrastructure Capacity Planning



## Compute and Storage Infrastructure Capacity Planning

Monitoring a virtual environment is vastly different from monitoring a physical infrastructure. There is no one-to-one correspondence between an operating system instance and a physical server. A typical virtualization host provides shared server, network, and storage resources for multiple operating system instances, each running their own OS and application workload.

As previously mentioned, the UC Sizing Tools are critical to assist with accurate solution sizing. These tools factor in data from performance testing, individual product limits and performance ratings, advanced and new features in product releases, design recommendations from this Cisco Collaboration Systems Release Solution Reference Network Designs (SRND), and other factors. Based on input provided by the system designer, the tools apply their sizing algorithms to the supplied data to recommend a set of hardware resources.

To optimize a virtualized environment, monitoring must encompass virtual resource utilization from the VM's perspective, application service levels, and physical resource utilization on the hosts.

Virtualization monitoring ensures that a virtualized infrastructure performs optimally and that virtual resources are properly allocated. Virtualized infrastructure monitoring requires collecting and evaluating key performance indicators (KPIs) for both physical and virtual components. For example, VMware KPIs include:

- Datacenters
- Clusters
- Datastores
- Hosts
- Resource pools
- Virtual Machines

Because physical resources are shared between VMs, problems that would be localized on a physical server can cascade through the virtualized infrastructure and compromise multiple applications. To cope with virtualization interdependencies, a monitoring strategy is required to optimize resource utilization by recognizing and reacting to performance and availability issues early in a problem cycle.

It is highly recommended that you follow VMware's *Performance Best Practices for VMware vSphere* for the specific release you are planning to deploy.

You can access the most current versions of the vSphere documentation by going to:

http://www.vmware.com/support/pubs

You can access performance and other technical papers on the VMware Technical Papers page:

http://www.vmware.com/vmtn/resources

# Collaboration Virtualization Supported Hardware

Cisco UC virtualization hardware support is predominately dependent on:

- Intel CPU model
- VMware Hardware Compatibility Guide (HCG)

The specific server model requirements are:

1. Must align with the VMware HCG; see https://www.vmware.com/resources/compatibility
2. Must contain an Intel CPU model allowed for UC virtualization; see ***Processors / CPUs*** section at Collaboration Virtualization Hardware
3. Additional hardware component options must satisfy all other requirements of the UC policies; see ***Can I use this server?*** section at Collaboration Virtualization Hardware
    - For additional considerations, see TAC TechNote 115955.

Cisco UC employs three different support models.

- Cisco Unified Communications on Cisco Unified Computing System (UC on UCS) which includes:
    - UC on UCS Tested Reference Configuration (TRC), some of which are available as packaged collaboration solutions like Cisco Business Edition 6000 or Cisco Business Edition 7000. "TRC" used by itself means "UC on UCS Tested Reference Configuration (TRC)".
    - UC on UCS Specs-based - "Specs-based" used by itself refers to the common rules of UC on UCS Specs-based and Third-party Server Specs-based.
- Third-party Server Specs-based - "Specs-based" used by itself refers to the common rules of UC on UCS Specs-based and Third-party Server Specs-based

Allowed vs. Supported CPU policies are different for UC on UCS TRC, UC on UCS Specs-based and Third-party Server Specs-based. Required physical core speeds (base frequencies) are different for applications using 1vcpu:1pcore sizing approach (see Collaboration Virtualization Sizing) vs. applications that support CPU Reservations, such as

- Cisco Unified Communications Manager (CUCM) 11.5(1) or greater
- Cisco Unified Communications Manager - IM & Presence (CUCM IM&P) 11.5(1) or greater
- Cisco Unity Connection (CUC) 11.5(1) or greater

For additional information see Caveated Support for VMware CPU Reservations and Distributed Resource Scheduler).

# Collaboration Virtualized Hardware Oversubscription

Cisco HCS Management and Cisco UC applications do not support CPU or RAM oversubscription. Cisco does publish Caveated Support for VMware CPU Reservations and Distributed Resource Scheduler; however, this too is does not allow oversubscription of the CPU but provides guidance to modify the default OVA CPU Reservation value, for your specific environment and workloads, to enable Distributed Resource Scheduler (DRS) functionality and maintain 1:1 virtual CPU core (vcore) to one physical CPU (pcore) requirements without having to manually intervein.

Cisco Unity Connection virtual machines also require VMware CPU Affinity. If there is at least one live Unity Connection virtual machine on the physical server, then one CPU core per physical server must be left unused (it is being used by ESXi scheduler).

For additional information refer to *Cisco HCS Virtual Machine Requirements*.

# Shared Storage Engineering

Cisco does not recommend a specific vendor, model, or configuration for third-party, non-hyperconverged shared storage with HCS. The shared storage requirements are:

1. Must align with the VMware HCG; see https://www.vmware.com/resources/compatibility
2. Must align with the Cisco UCS Hardware and Software Compatibility when using Cisco UCS hardware; see https://ucshcltool.cloudapps.cisco.com/public/
3. Must meet UC IOPS (In/Out Operations per second) and Latency requirements; see https://www.cisco.com/c/dam/en/us/td/docs/voice_ip_comm/uc_system/virtualization/virtualization-collaboration-storage-design-requirements.html
4. Must meet all storage space requirements for each virtual appliance and virtual machine deployed.

Storage thin provisioning (TP) at the VM Layer or storage array is not recommended and will only save space during initial deployment for Cisco UC applications. Cisco UC applications have been redesigned to utilize 100% of virtual disk (vDisk) either for UC features (such as Unity Connection message store or Contact Center reporting databases) or critical operations (such as spikes during upgrades, backups or statistics writes). Space allocated and will expand based upon Cisco Options Package (COP) files, (Call Detail Records) CDRs, Trace Files (debugging) and during the first application Upgrade. While thin provisioning does not introduce a performance penalty, not having physical disk space available when the app needs it can have the following harmful effects:

- Degrade UC app performance, crash the UC app and/or corrupt the vDisk contents
- Lock up all UC VMs on the same LUN within a SAN

With Cisco Business Edition 6000/7000 appliances or any of the Tested Reference Configurations (TRCs) with Direct-attached storage (DAS) or HyperFlex storage, the DAS configuration (e.g. disk technology, quantity, size, speed, RAID configuration) has already been properly designed to provide enough IOPS capacity and performance has been explicitly validated by Collaboration applications. Just follow the normal sizing rules described in Collaboration Virtualization Hardware and Sizing Guidelines such as CPU, memory, and storage capacity requirements.

It is important that you collect shared storage metrics to provide usage analytics. The shared storage should be engineered to handle the amount of IOPS and disk space required by the Cisco HCS system over a 24-hour period.

Usage patterns may vary for UC applications under normal operating conditions and at times of the day when backup and restore activities occur. The maintenance window might be the time of day when Shared storage usage is highest.

You can monitor IOPS loading based on the sum of IOPS loading that each application type generates on Cisco HCS. After you monitor IOPS and disk space utilization on an application-type basis, you can make future IOPS projections for growth planning. You must map the utilization that was experienced against the capacity specification of the purchased Shared storage to compare when a Shared storage upgrade is necessary.

- Cisco HCS recommends using shared storage vendor supplied or third-party monitoring software to monitor overall shared storage IOPS.

# UCS Storage Options

The following are the UCS Manager storage options and the benefits of each.

- **Direct Attached Storage (DAS)** – This is the storage available inside a server and is directly connected to the system through the motherboard within a parallel SCSI implementation. DAS is commonly described as captive storage. Devices in a captive storage topology do not have direct access to the storage network and do not support efficient sharing of storage. To access data with DAS, a user must go through a front-end network. DAS devices provide little or no mobility to other servers and little scalability.

  DAS devices limit file sharing and can be complex to implement and manage. For example, to support data backups, DAS devices require resources on the host and spare disk systems that other systems cannot use. The cost and performance of this storage depends upon the disks and RAID controller cards inside the servers. DAS is less expensive and is simple to configure; however, it lacks the scalability, performance, and advanced features provided by high-end storage.

- **Network Attached Storage (NAS)** – This storage is usually an appliance providing file system access. This storage could be as simple as an Network File System (NFS) or Common Internet File System (CIFS) share available to the servers. Typical NAS devices are cost-effective devices with not very high performance but have very high capacity with some redundancy for reliability. NAS is usually moderately expensive, simple to configure, and provides some advanced features; however, it also lacks scalability, performance, and advanced features provided by SAN.
- **Storage Area Network (SAN)** - A SAN is a specialized, high-speed network that attaches servers and storage devices. A SAN allows an any-to-any connection across the network by using interconnect elements, such as switches and directors. It eliminates the traditional dedicated connection between a server and storage, and the concept that the server effectively owns and manages the storage devices. It also eliminates any restriction to the amount of data that a server can access, currently limited by the number of storage devices that are attached to the individual server. Instead, a SAN introduces the flexibility of networking to enable one server or many heterogeneous servers to share a common storage utility. A network might include many storage devices, including disk, tape, and optical storage. Additionally, the storage utility might be located far from the servers that it uses. This type of storage provides maximum reliability, expandability, and performance. The cost of SAN is also very high compared to other storage options.

  SAN is the most resilient, highly scalable, and high-performance storage; however, it is also the most expensive and complex to manage.

# UCS Storage Design Considerations

UCS storage physical connectivity has a slightly different design consideration as compared to LAN physical connectivity. The following are some design considerations for SAN connectivity:

- Northbound storage physical connectivity does not support virtual port channels (vPCs) like LAN connectivity.
- Port channels or trunking is possible to combine multiple storage uplink ports that provide physical link redundancy.
- Redundancy of storage resources is handled by the storage itself and varies from vendor to vendor.
- Connect storage through northbound Cisco storage devices, such as Nexus or MDS Fabric Switches.
- It is possible to connect storage directly to UCS Fabric Interconnects, which is recommended for small implementations because of the fabric interconnect physical ports consumption and increased processing requirements.
- Software configuration including VSANs and zoning is required for providing access to storage resources.

For more information in regards to UCS Storage and Storage Management, see the *Cisco UCS Manager Storage Management Guide* for the specific release being deployed at:
https://www.cisco.com/c/en/us/support/servers-unified-computing/ucs-manager/products-installation-and-configuration-guides-list.html