



# Reference Architecture: Lenovo Client Virtualization with Citrix Virtual Apps and Desktops on ThinkSystem and ThinkAgile VX Servers

Last update: 6 February 2020

---

Reference Architecture for  
Citrix Virtual Apps and  
Desktops

---

Contains performance data and  
sizing recommendations for  
servers, storage, and networking

---

Describes variety of storage  
models including SAN storage  
and hyper-converged systems

---

Contains detailed bill of materials  
for servers, storage, networking,  
and racks

Chandrakandh Mouleeswaran  
Cristian Ghetau  
Srihari Angaluri



# Table of Contents

<b>1</b>	<b>Introduction .....</b>	<b>1</b>
<b>2</b>	<b>Architectural overview .....</b>	<b>2</b>
<b>3</b>	<b>Component model .....</b>	<b>3</b>
3.1	Citrix Virtual Apps and Desktops provisioning.....	5
3.1.1	Provisioning Services .....	5
3.1.2	Machine Creation Services.....	6
3.2	Storage model.....	7
3.3	VMware vSAN.....	8
3.3.1	VMware vSAN storage policies .....	8
<b>4</b>	<b>Operational model .....</b>	<b>11</b>
4.1	Operational model scenarios .....	11
4.1.1	Enterprise operational model.....	12
4.1.2	Hyper-converged operational model .....	12
4.2	Hypervisor support.....	13
4.2.1	VMware ESXi.....	13
4.2.2	Citrix Hypervisor .....	13
4.3	Compute servers for virtual desktops.....	13
4.3.1	Lenovo ThinkSystem with Intel Xeon Scalable Processor Gen 1 (Skylake) .....	14
4.3.2	Lenovo ThinkSystem with AMD EPYC™ Processors (Rome) .....	17
4.4	Compute servers for VMware vSAN .....	21
4.4.1	Lenovo ThinkSystem with Intel Xeon Scalable Processor Gen 1 (Skylake) .....	21
4.4.2	Lenovo ThinkAgile VX3320 Appliance with Intel Xeon Scalable Processor Gen 2 (Cascadelake) .....	24
4.5	Graphics acceleration .....	26
4.6	Management servers .....	28
4.7	Systems management .....	31
4.8	Shared storage .....	32
4.8.1	Lenovo ThinkSystem DS6200 storage array.....	33
4.9	Networking .....	35
4.9.1	10 GbE networking .....	35
4.9.2	1 GbE administration networking.....	36

4.10	Racks.....	36
4.11	Deployment models .....	37
4.11.1	Deployment example 1: vSAN hyper-converged using ThinkSystem servers .....	37
<b>5</b>	<b>Appendix: Bill of materials.....</b>	<b>38</b>
5.1	BOM for enterprise and SMB compute servers.....	38
5.2	BOM for ThinkSystem SR635 and SR655 (AMD EPYC processor) based compute nodes.....	40
5.3	BOM for hyper-converged compute and Management servers .....	43
5.4	BOM for enterprise and SMB management servers .....	49
5.5	BOM for shared storage.....	51
5.6	BOM for networking .....	52
5.7	BOM for Flex System chassis .....	53
5.8	BOM for rack.....	53
	<b>Resources .....</b>	<b>54</b>

# 1 Introduction

---

The intended audience for this document is technical IT architects, system administrators, and managers who are interested in server-based desktop virtualization and server-based computing (terminal services or application virtualization) that uses Citrix Virtual Apps and Desktops. In this document, the term client virtualization is used as to refer to all of these variations. Compare this term to server virtualization, which refers to the virtualization of server-based business logic and databases.

This document describes the reference architecture for Citrix Virtual Apps and Desktops 7.18 and also supports the previous versions of Citrix Virtual Apps and Desktops 7.x. This document should be read with the Lenovo Client Virtualization (LCV) base reference architecture document that is available at this website: [lenovopress.com/lp0756](http://lenovopress.com/lp0756).

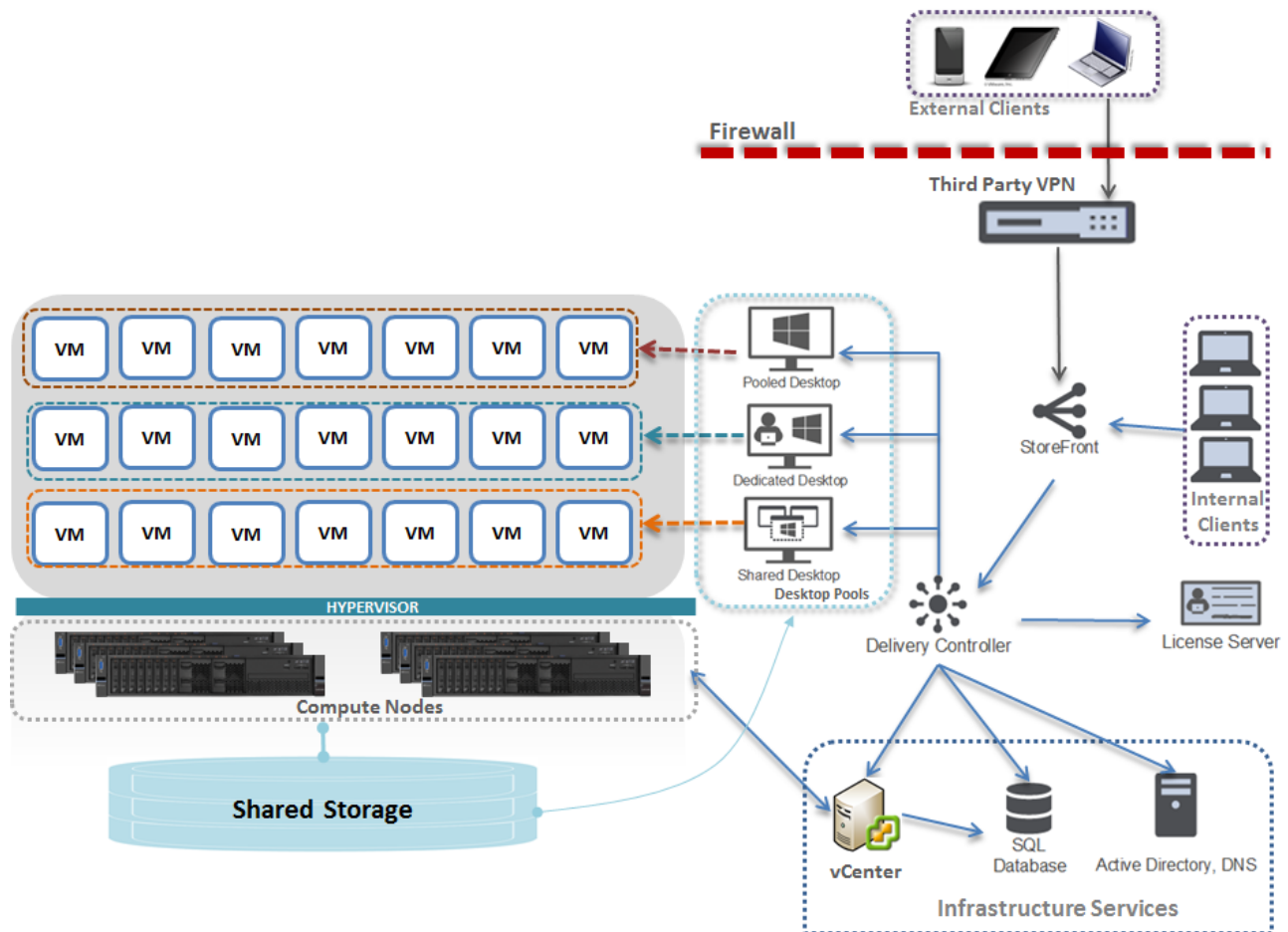
The business problem, business value, requirements, and hardware details are described in the LCV base reference architecture document and are not repeated here for brevity.

This document gives an architecture overview and logical component model of Citrix Virtual Apps and Desktops. The document also provides the operational model of Citrix Virtual Apps and Desktops by combining Lenovo® hardware platforms such as ThinkSystem Servers, ThinkAgile VX Servers and RackSwitch networking with storage such as ThinkSystem DS6200 storage or VMware vSAN. The operational model presents performance benchmark measurements and discussion, sizing guidance, and some example deployment models. The last section contains detailed bill of material configurations for each piece of hardware.

See also the “Reference Architecture for Workloads using the Lenovo ThinkAgile HX Series appliances” for Citrix Virtual Apps and Desktops specific information. This document is available at this website: [lenovopress.com/lp0665](http://lenovopress.com/lp0665).

## 2 Architectural overview

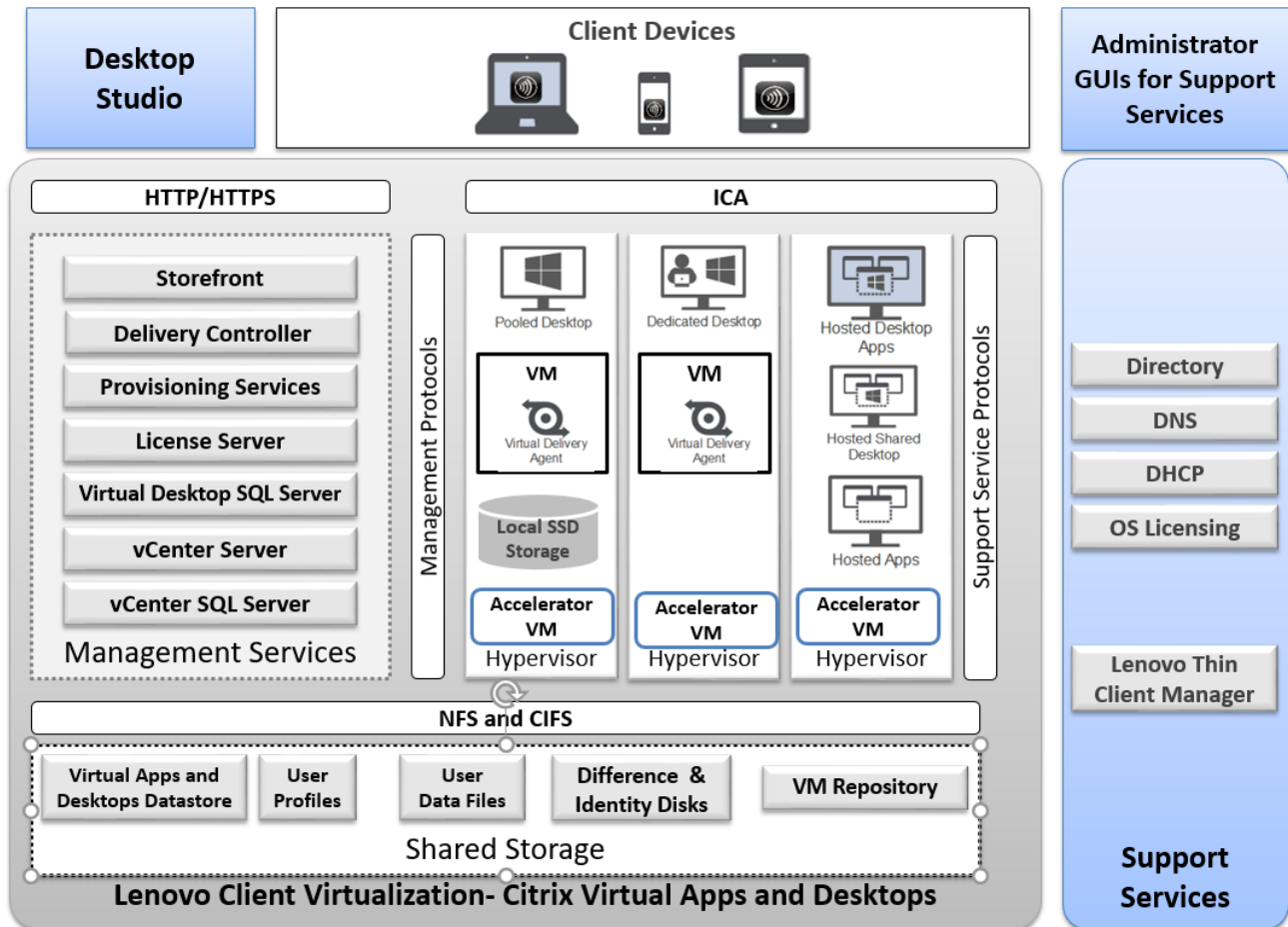
Figure 1 shows all of the main features of the Lenovo Client Virtualization reference architecture with Citrix Virtual Apps and Desktops. This reference architecture does not address the issues of remote access and authorization, data traffic reduction, traffic monitoring, and general issues of multi-site deployment and network management. This document limits the description to the components that are inside the customer's intranet.



**Figure 1: LCV reference architecture with Citrix Virtual Apps and Desktops**

### 3 Component model

Figure 2 is a layered view of the LCV solution that is mapped to the Citrix Virtual Apps and Desktops virtualization infrastructure.



**Figure 2: Component model with Citrix Virtual Apps and Desktops**

Citrix Virtual Apps and Desktops features the following main components:

- |                |   |
|----------------|---|
| Desktop Studio | Desktop Studio is the main administrator GUI for Citrix Virtual Apps and Desktops. It is used to configure and manage all of the main entities, including servers, desktop pools and provisioning, policy, and licensing.   |
| Storefront     | Storefront provides the user interface to the Virtual Apps and Desktops environment. The Web Interface brokers user authentication, enumerates the available desktops and, upon start, delivers a .ica file to the Citrix Receiver on the user's local device to start a connection. The Independent Computing Architecture (ICA) file contains configuration information for the Citrix receiver to communicate with the virtual desktop. Because the Web Interface is a critical component, redundant servers must be available to provide fault tolerance. |

Delivery controller	<p>The Delivery controller is responsible for maintaining the proper level of idle desktops to allow for instantaneous connections, monitoring the state of online and connected desktops, and shutting down desktops as needed.</p> <p>The Virtual Apps and Desktops farm is a larger grouping of virtual machine servers. Each delivery controller in the Virtual Apps and Desktops acts as an XML server that is responsible for brokering user authentication, resource enumeration, and desktop starting. Because a failure in the XML service results in users being unable to start their desktops, it is recommended that you configure multiple controllers per farm.</p>
PVS and MCS	<p>Provisioning Services (PVS) is used to provision stateless desktops at a large scale. Machine Creation Services (MCS) is used to provision dedicated or stateless desktops in a quick and integrated manner. For more information, see “Citrix Virtual Apps and Desktops provisioning” section on page 5.</p>
License Server	<p>The Citrix License Server is responsible for managing the licenses for all Virtual Apps and Desktops components. Citrix Virtual Apps and Desktops has a 30-day grace period that allows the system to function normally for 30 days if the license server becomes unavailable. This grace period offsets the complexity of otherwise building redundancy into the license server.</p>
Virtual Apps and Desktops SQL Server	<p>Each Citrix Virtual Apps and Desktops site requires an SQL Server database to centralize farm configuration information and transaction logs. The database maintains all static and dynamic information about the Virtual Apps and Desktops environment. Because the SQL server is a critical component, redundant servers must be available to provide fault tolerance.</p>
vCenter Server	<p>By using a single console, vCenter Server provides centralized management of the virtual machines (VMs) for the VMware ESXi hypervisor. VMware vCenter can be used to perform live migration (called VMware vMotion), which allows a running VM to be moved from one physical server to another without downtime.</p> <p>Redundancy for vCenter Server is achieved through VMware high availability (HA). The vCenter Server also contains a licensing server for VMware ESXi.</p>
vCenter SQL Server	<p>vCenter Server for VMware ESXi hypervisor requires an SQL database. The vCenter SQL server might be Microsoft® Data Engine (MSDE), Oracle, or SQL Server. Because the vCenter SQL server is a critical component, redundant servers must be available to provide fault tolerance. Customer SQL databases (including respective redundancy) can be used.</p>

Client devices	Citrix Virtual Apps and Desktops supports a broad set of devices and all major device operating platforms, including Apple iOS, Google Android, and Google ChromeOS. Virtual Apps and Desktops enables a rich, native experience on each device, including support for gestures and multi-touch features, which customizes the experience based on the type of device. Each client device has a Citrix Receiver, which acts as the agent to communicate with the virtual desktop by using the ICA/HDX protocol.
VDA	Each VM needs a Citrix Virtual Desktop Agent (VDA) to capture desktop data and send it to the Citrix Receiver in the client device. The VDA also emulates keyboard and gestures sent from the receiver. ICA is the Citrix remote display protocol for VDI.
Hypervisor	Virtual Apps and Desktops has an open architecture that supports the use of several different hypervisors, such as VMware ESXi (vSphere), Citrix Hypervisor, and Microsoft Hyper-V.
Accelerator VM	The optional accelerator VM.
Shared storage	Shared storage is used to store user profiles and user data files. Depending on the provisioning model that is used, different data is stored for VM images. For more information, see “Storage model”. Shared storage can also be distributed storage as used in hyper-converged systems.

For more information, see the Lenovo Client Virtualization base reference architecture document that is available at this website: [lenovopress.com/lp0756](http://lenovopress.com/lp0756).

## 3.1 Citrix Virtual Apps and Desktops provisioning

Citrix Virtual Apps and Desktops features the following primary provisioning models:

- Provisioning Services (PVS)
- Machine Creation Services (MCS)

### 3.1.1 Provisioning Services

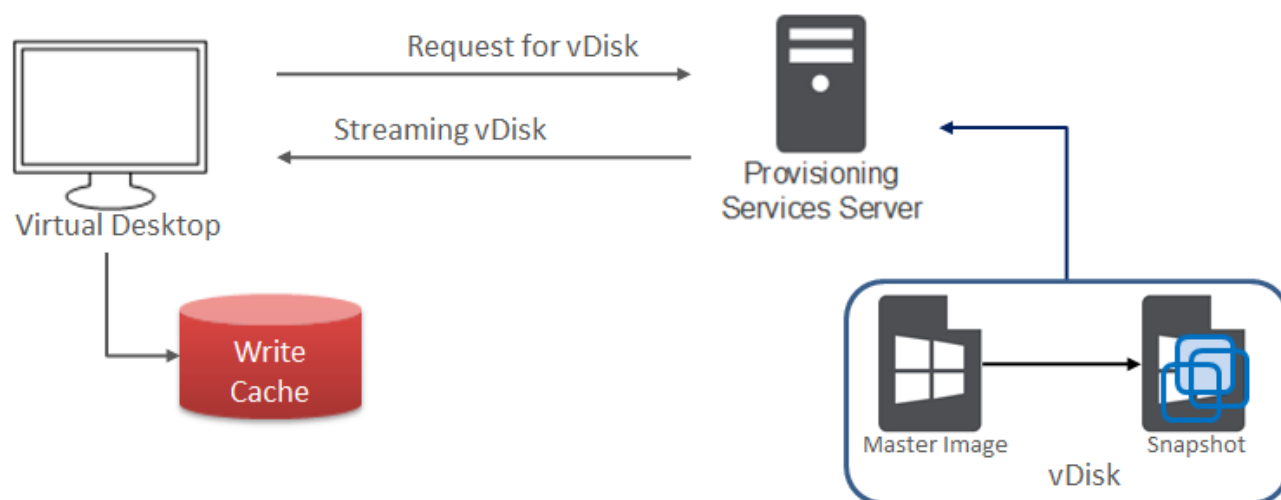
Hosted VDI desktops can be deployed with or without Citrix PVS. The advantage of the use of PVS is that you can stream a single desktop image to create multiple virtual desktops on one or more servers in a data center. Figure 3 shows the sequence of operations that are run by Virtual Apps and Desktops to deliver a hosted VDI virtual desktop.

When the virtual disk (vDisk) master image is available from the network, the VM on a target device no longer needs its local hard disk drive (HDD) to operate; it boots directly from the network and behaves as if it were running from a local drive on the target device, which is why PVS is recommended for stateless virtual desktops. PVS often is not used for dedicated virtual desktops because the write cache is not stored on shared storage.

PVS is also used with Microsoft Roaming Profiles (MSRPs) so that the user's profile information can be separated out and reused. Profile data is available from shared storage.



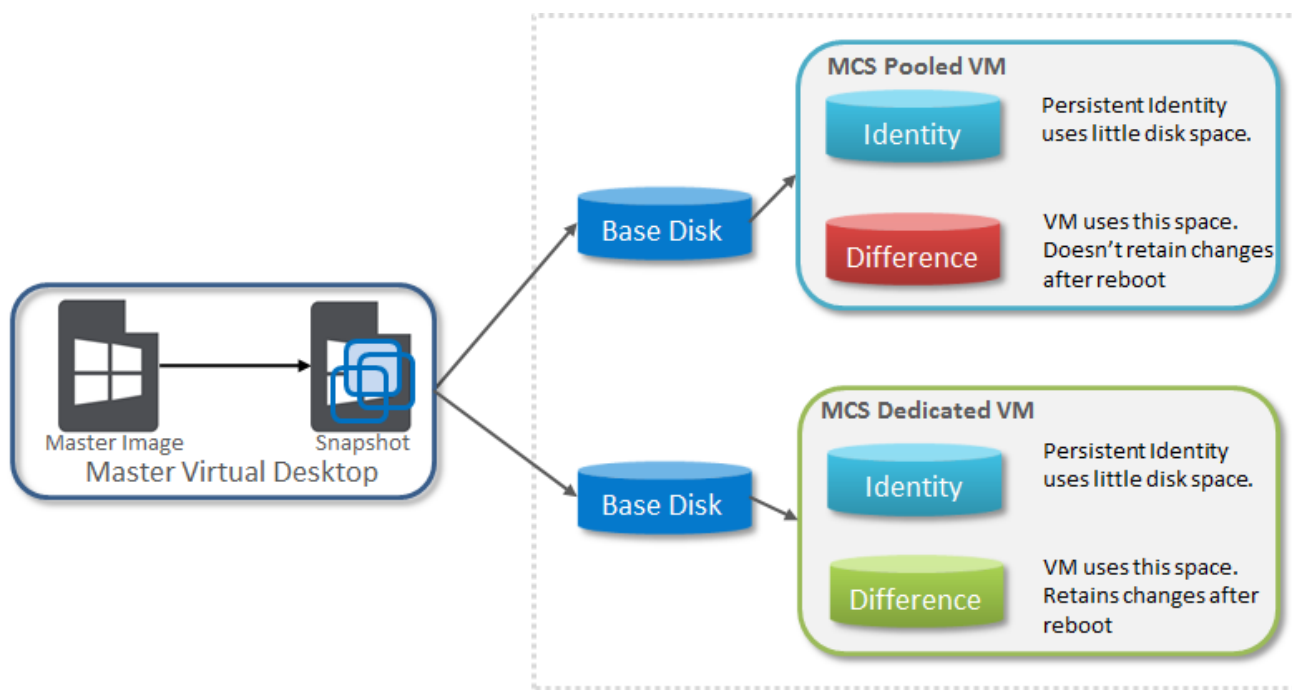
It is a best practice to use snapshots for changes to the master VM images and also keep copies as a backup.



**Figure 3: Using PVS for a stateless model**

### 3.1.2 Machine Creation Services

Unlike PVS, MCS does not require more servers. Instead, it uses integrated functionality that is built into the hypervisor (VMware ESXi, Citrix Hypervisor, or Microsoft Hyper-V) and communicates through the respective APIs. Each desktop has one difference disk and one identity disk (as shown in Figure 4). The difference disk is used to capture any changes that are made to the master image. The identity disk is used to store information, such as device name and password.



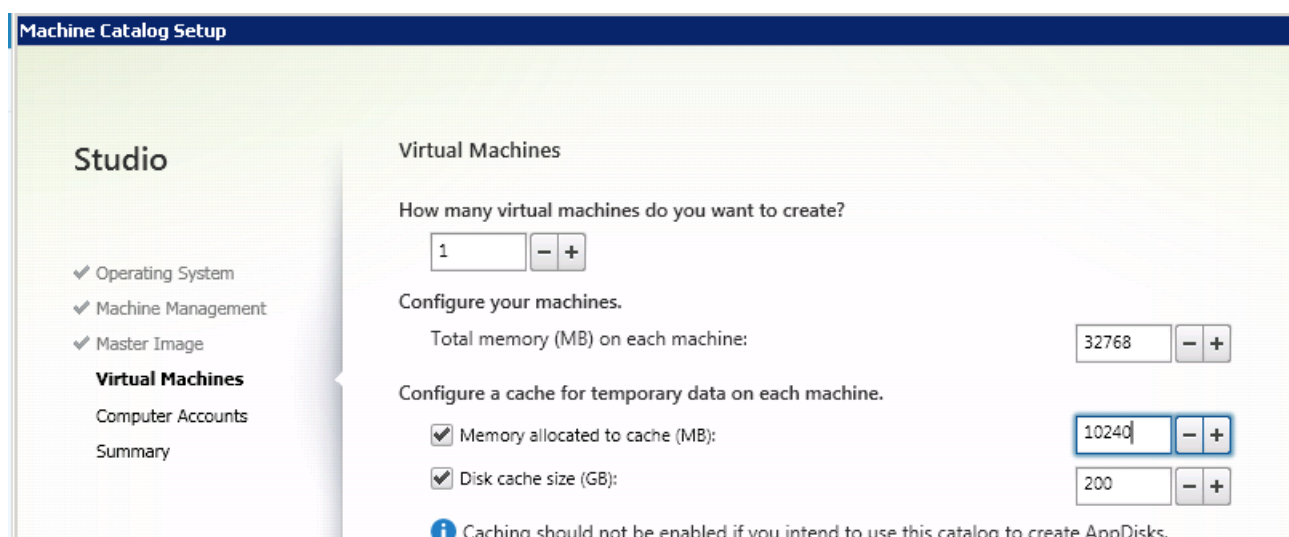
**Figure 4: MCS image and difference/identity disk storage model**

The following types of Image Assignment Models for MCS are available:

- Pooled-random: Desktops are assigned randomly. When they log off, the desktop is free for another user. When rebooted, any changes that were made are destroyed.
- Pooled-static: Desktops are permanently assigned to a single user. When a user logs off, only that user can use the desktop, regardless if the desktop is rebooted. During reboots, any changes that are made are destroyed.
- Dedicated: Desktops are permanently assigned to a single user. When a user logs off, only that user can use the desktop, regardless if the desktop is rebooted. During reboots, any changes that are made persist across subsequent restarts.

MCS thin provisions each desktop from a master image by using built-in technology to provide each desktop with a unique identity. Only changes that are made to the desktop use more disk space. For this reason, MCS dedicated desktops are used for dedicated desktops.

There is a new caching option in Citrix Virtual Apps and Desktops 7.9 onwards for Pooled and Hosted Shared desktops. Figure 5 shows a screenshot of the option.



**Figure 5: Caching option for Pooled and Hosted Shared desktops**

## 3.2 Storage model

This section describes the different types of shared or distributed data stored for stateless and dedicated desktops. Stateless and dedicated virtual desktops should have the following common shared storage items:

- The master VM image and snapshots are stored by using Network File System (NFS) or block I/O shared storage.
- The paging file (or vSwap) is transient data that can be redirected to NFS storage. In general, it is recommended to disable swapping, which reduces storage use (shared or local). The desktop memory size should be chosen to match the user workload rather than depending on a smaller image and swapping, which reduces overall desktop performance.

- User profiles (from MSRP) are stored by using Common Internet File System (CIFS).
- User data files are stored by using CIFS.

Dedicated virtual desktops or stateless virtual desktops that need mobility require the following items to be on NFS or block I/O shared storage:

- Difference disks are used to store user's changes to the base VM image. The difference disks are per user and can become quite large for dedicated desktops.
- Identity disks are used to store the computer name and password and are small.
- Stateless desktops can use local solid-state drive (SSD) storage for the PVS write cache, which is used to store all image writes on local SSD storage. These image writes are discarded when the VM is shut down.

### 3.3 VMware vSAN

VMware vSAN is a Software Defined Storage (SDS) solution embedded in the ESXi hypervisor. vSAN pools flash caching devices and magnetic disks across three or more 10 GbE connected servers into a single shared datastore that is resilient and simple to manage.

vSAN can be scaled to 64 servers, with each server supporting up to 5 disk groups, with each disk group consisting of a single flash caching device (SSD) and up to 7 HDDs. Performance and capacity can easily be increased simply by adding more components: disks, flash or servers.

The flash cache is used to accelerate both reads and writes. Frequently read data is kept in read cache; writes are coalesced in cache and destaged to disk efficiently, greatly improving application performance.

vSAN manages data in the form of flexible data containers that are called objects and the following types of objects for VMs are available:

- VM Home
- VM swap (.vswp)
- VMDK (.vmdk)
- Snapshots (.vmsn)

Internally, VM objects are split into multiple components that are based on performance and availability requirements that are defined in the VM storage profile. These components are distributed across multiple hosts in a cluster to tolerate simultaneous failures and meet performance requirements. vSAN uses a distributed RAID architecture to distribute data across the cluster. Components are distributed with the use of the following two storage policies:

- Number of stripes per object. It uses RAID 0 method.
- Number of failures to tolerate. It uses either RAID-1 or RAID-5/6 method. RAID-5/6 is currently supported for the all flash configuration only.

#### 3.3.1 VMware vSAN storage policies

VMware vSAN uses the Storage Policy-based Management (SPBM) function in vSphere to enable policy driven virtual machine provisioning, and uses vSphere APIs for Storage Awareness (VASA) to expose vSAN's storage capabilities to vCenter.

This approach means that storage resources are dynamically provisioned based on requested policy, and not pre-allocated as with many traditional storage solutions. Storage services are precisely aligned to VM boundaries; change the policy, and vSAN will implement the changes for the selected VMs.

VMware vSAN has the following storage policies:

Storage Policy	Description	Default	Maximum
Failure Tolerance Method	Defines a method used to tolerate failures. RAID-1 uses mirroring and RAID 5/6 uses parity blocks (erasure encoding) to provide space efficiency. RAID-5/6 is supported only for All Flash configurations. RAID 5 requires minimum 4 hosts and RAID 6 requires minimum 6 hosts. When RAID 5/6 is chosen, RAID 5 is used when FTT=1 and RAID 6 is used when FTT=2.	RAID-1	N/A
Number of failures to tolerate	Defines the number of host, disk, or network failures a VM object can tolerate. For n failures tolerated, n+1 copies of the VM object are created and 2n+1 hosts with storage are required.  For example with a FTT=1, RAID-1 uses 2x the storage and RAID-5/6 uses 1.33x the storage. When FTT=2, RAID-1 uses 3x the storage and RAID-5/6 uses 1.5x the storage.	1	3
Number of disk stripes per object	The number of HDDs across which each replica of a VM object is striped. A value higher than 1 might result in better performance, but can result in higher use of resources.	1	12
Object space reservation	Percentage of the logical size of the object that should be reserved (or thick provisioned) during VM creation. The rest of the storage object is thin provisioned. If your disk is thick provisioned, 100% is reserved automatically. When deduplication and compression is enabled, this should be set to either 0% (do not apply) or 100%.	0%	100%
Flash read cache reservation	SSD capacity reserved as read cache for the VM object. Specified as a percentage of the logical size of the object. Should be used only to address read performance issues. Reserved flash capacity cannot be used by other objects. Unreserved flash is shared fairly among all objects.	0%	100%
Force provisioning	If the option is set to Yes, the object is provisioned, even if the storage policy cannot be satisfied by the data store. Use this parameter in bootstrapping scenarios and during an outage when standard provisioning is no longer possible. The default of No is acceptable for most production environments.	No	N/A
IOPS limit for object	Defines IOPS limit for a disk and assumes a default block size of 32 KB. Read, write and cache operations are all considered equivalent. When the IOPS exceeds the limit, then IO is throttled.	0	User Defined

Storage Policy	Description	Default	Maximum
Disable object checksum	Detects corruption caused by hardware/software components including memory, drives, etc. during the read or write operations. Object checksums carry a small disk IO, memory and compute overhead and can be disabled on a per object basis.	No	Yes

The following storage policies and default values are recommended for Citrix Virtual Apps and Desktops linked clones and full clones. Table 1 lists the default storage policies for linked clones.

**Table 1: Default storage policy values for linked clones**

Storage Policy	Dedicated				Stateless		
	VM_HOME	Replica	OS Disk	Persistent Disk	VM_HOME	Replica	OS Disk
Number of disk stripes per object	1	1	1	1	1	1	1
Flash-memory read cache reservation	0%	10 %	0%	0%	0%	10 %	0%
Number of failures to tolerate (FTT)	1	1	1	1	1	1	1
Failure Tolerance Method	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Force provisioning	No	No	No	No	No	No	No
Object-space reservation	0%	0%	0%	100%	0%	0%	0%
Disable object Checksum	N/A	N/A	N/A	N/A	N/A	N/A	N/A
IOPS limit for object	N/A	N/A	N/A	N/A	N/A	N/A	N/A

Table 2 lists the default storage policies for full clones.

**Table 2: Default storage policy values for full clones**

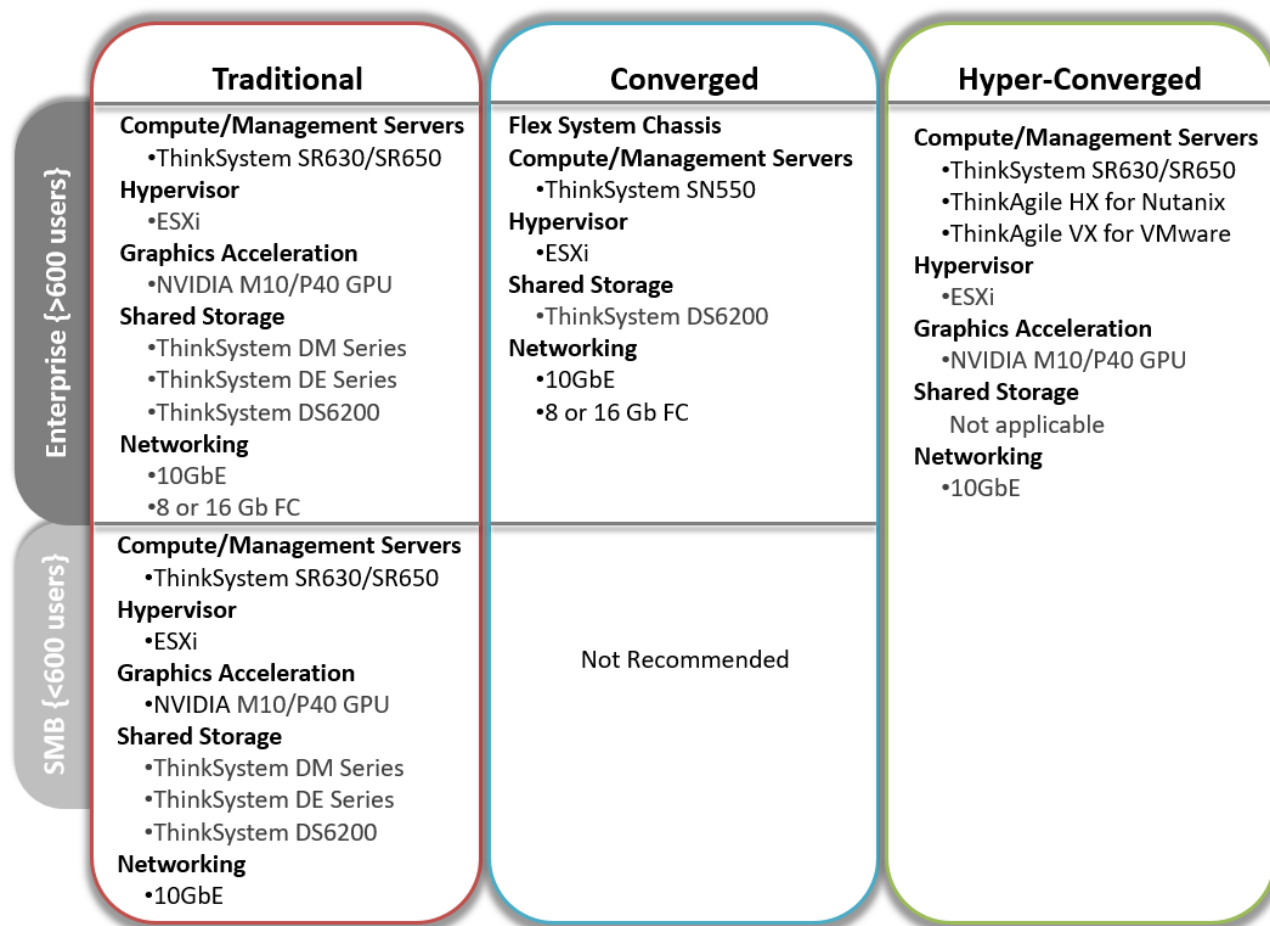
Storage Policy	Dedicated	
	VM_HOME	Full Clone Disk
Number of disk stripes per object	1	1
Flash-memory read cache reservation	0%	0%
Number of failures to tolerate (FTT)	1	1
Failure Tolerance Method	N/A	N/A
Force provisioning	No	No
Object-space reservation	0%	100%
Disable object Checksum	N/A	N/A
IOPS limit for object	N/A	N/A

## 4 Operational model

This section describes the options for mapping the logical components of a client virtualization solution onto hardware and software. The “Operational model scenarios” section gives an overview of the available mappings and has pointers into the other sections for the related hardware. Each subsection contains performance data, has recommendations on how to size for that particular hardware, and a pointer to the BOM configurations that are described in section 5 on page 38. The last part of this section contains some deployment models for example customer scenarios.

### 4.1 Operational model scenarios

Figure 6 shows the following operational models (solutions) in Lenovo Client Virtualization: enterprise, small-medium business (SMB), and hyper-converged.



**Figure 6: Operational model scenarios**

The vertical axis is split into two halves: greater than 600 users is termed Enterprise and less than 600 is termed SMB. The 600 user split is not exact and provides rough guidance between Enterprise and SMB. The last column in Figure 6 (labelled “hyper-converged”) spans both halves because a hyper-converged solution can be deployed in a linear fashion from a small number of users (100) up to a large number of users (>4000).

The horizontal axis is split into three columns. The left-most column represents traditional rack-based systems with top-of-rack (TOR) switches and shared storage. The middle column represents converged systems where the compute, networking, and sometimes storage are converged into a chassis, such as the Flex System. The right-most column represents hyper-converged systems and the software that is used in these systems. For the purposes of this reference architecture, the traditional and converged columns are merged for enterprise solutions; the only significant differences are the networking, form factor, and capabilities of the compute servers.

Converged systems are not generally recommended for the SMB space because the converged hardware chassis can be more overhead when only a few compute nodes are needed. Other compute nodes in the converged chassis can be used for other workloads to make this hardware architecture more cost-effective.

### **4.1.1 Enterprise operational model**

For the enterprise operational model, see the following sections for more information about each component, its performance, and sizing guidance:

- 4.2 Hypervisor support
- 4.3 Compute servers for virtual desktops
- 4.5 Graphics acceleration
- 4.6 Management servers
- 4.7 Systems management
- 4.8 Shared storage
- 4.9 Networking
- 4.10 Racks

To show the enterprise operational model for different sized customer environments, four different sizing models are provided for supporting 300, 600, 1500, and 3000 users.

The SMB model is the same as the Enterprise model for traditional systems.

### **4.1.2 Hyper-converged operational model**

For the hyper-converged operational model, see the following sections for more information about each component, its performance, and sizing guidance:

- 4.2 Hypervisor support
- 4.4 Compute servers for VMware vSAN
- 4.5 Graphics acceleration
- 4.6 Management servers
- 4.7 Systems management
- 4.9.1 10 GbE networking
- 4.10 Racks

To show the hyper-converged operational model for different sized customer environments, four different sizing models are provided for supporting 300, 600, 1500, and 3000 users. The management server VMs for a hyper-converged cluster can either be in a separate hyper-converged cluster or on traditional shared storage.

## 4.2 Hypervisor support

This reference architecture was tested with VMware ESXi 6.5 U1, 6.7U3 and Citrix Hypervisor 7.3 hypervisors.

### 4.2.1 VMware ESXi

The hypervisor is convenient because it can start from a M.2 boot drive and does not require any additional local storage. Alternatively, ESXi can be booted from SAN shared storage. For the VMware ESXi hypervisor, the number of vCenter clusters depends on the number of compute servers.

To use the latest version ESXi, download the Lenovo custom image from the following websites:

[my.vmware.com/web/vmware/info/slug/datacenter\\_cloud\\_infrastructure/vmware\\_vsphere/6\\_5#custom\\_iso](https://my.vmware.com/web/vmware/info/slug/datacenter_cloud_infrastructure/vmware_vsphere/6_5#custom_iso)

[https://my.vmware.com/web/vmware/info/slug/datacenter\\_cloud\\_infrastructure/vmware\\_vsphere/6\\_7#custom\\_iso](https://my.vmware.com/web/vmware/info/slug/datacenter_cloud_infrastructure/vmware_vsphere/6_7#custom_iso)

### 4.2.2 Citrix Hypervisor

The hypervisor is convenient because it can start from a M.2 boot drive and does not require any additional local storage.

## 4.3 Compute servers for virtual desktops

This section describes stateless and dedicated virtual desktop models. In some customer environments, stateless and dedicated desktop models might be required, which requires a hybrid implementation. In this section stateless refers to MCS Pooled-Random and dedicated refers to MCS Dedicated.

Compute servers are servers that run a hypervisor and host virtual desktops. There are several considerations for the performance of the compute server, including the processor family and clock speed, the number of processors, the speed and size of main memory, and local storage options.




For stateless users, the typical range of memory that is required for each desktop is 2 GB - 6 GB. For dedicated users, the range of memory for each desktop is 2 GB - 8 GB. Designers and engineers that require graphics acceleration might need 8 GB - 16 GB of RAM per desktop. In general, power users that require larger memory sizes also require more virtual processors. The virtual desktop memory should be large enough so that swapping is not needed and vSwap can be disabled. This reference architecture standardizes on 2 GB per desktop as the absolute minimum requirement of a Windows 10 desktop.

Windows 10 was used for all of the performance testing. In general Windows 10 requires 10% to 20% more compute power than Windows 7. The following optimizations were applied to Windows 10 base image:

- Applied #VDILIKEAPRO Tuning Template (developed by LoginVSI) – see the following for more details:  
[loginvsi.com/blog/520-the-ultimate-windows-10-tuning-template-for-any-vdi-environment](https://loginvsi.com/blog/520-the-ultimate-windows-10-tuning-template-for-any-vdi-environment)
- Set Adobe acrobat as a default app for PDF files using steps in following webpage:  
[adobe.com/devnet-docs/acrobatetk/tools/AdminGuide/pdfviewer.html](https://adobe.com/devnet-docs/acrobatetk/tools/AdminGuide/pdfviewer.html)



- Disabled **Windows Modules installer** service on the base image because the CPU utilization can remain high after rebooting all the VMs. By default this service is set to manual rather than disabled.

	Windows Phone IP over US...	Enables co...	Running	Automatic	Local System...
	Windows Modules Installer	Enables inst...		Manual	Local System...
	Windows Mobile Hotspot S...	Provides th...		Manual (Trig...	Local Service

For more information, see “BOM for enterprise and SMB compute servers” section on page 38.

### 4.3.1 Lenovo ThinkSystem with Intel Xeon Scalable Processor Gen 1 (Skylake)

This section shows the performance results and sizing guidance for Lenovo ThinkSystem compute servers based on the Intel Scalable processor Gen 1 (Purley Skylake).

#### VMware ESXi 6.5 U1 performance results

Table 3 lists the VSI max performance results for different Login VSI 4.1 workloads with ESXi 6.5 U1 and Windows 10.

**Table 3: Login VSI Performance for ESXi**

Processor	Workload	Stateless	Dedicated
Two Scalable 6130 processors 2.10 GHz, 16C 125W	Office worker	276 users	271 users
Two Scalable 6130 processors 2.10 GHz, 16C 125W	Knowledge worker	260 users	258 users
Two Scalable 8176 processors 2.10 GHz, 28C 165W	Knowledge worker	418 users	420 users
Two Scalable 6130 processors 2.10 GHz, 16C 125W	Power worker	202 users	195 users
Two Scalable 8176 processors 2.10 GHz, 28C 165W	Power worker	347 users	347 users

These results indicate the comparative processor performance. The following conclusions can be drawn:

- The performance for stateless and dedicated virtual desktops is similar.
- The Scalable 8176 processor has significantly better performance than the Scalable 6130 processor and is directly proportional to the additional 12 cores but at an added cost.

#### Citrix Hypervisor 7.3 performance results

Table 4 lists the VSI max performance results for different Login VSI 4.1 workloads with Citrix Hypervisor 7.3 and Windows 10.

**Table 4: Login VSI Performance for Citrix Hypervisor**

Processor	Workload	Stateless	Dedicated
Two Scalable 6130 processors 2.10 GHz, 16C 125W	Office worker	240 users	244 users
Two Scalable 6130 processors 2.10 GHz, 16C 125W	Knowledge worker	223 users	224 users

## Compute server recommendation

The Xeon Scalable 6130 processor has a good cost/user density ratio. Many configurations are bound by memory; therefore, a faster processor might not provide any added value. Some users require a very fast processor and for those users, the Xeon Scalable 8176 processor is a good choice.

The default recommendation is two Xeon Scalable 6130 processors and 768 GB of system memory because this configuration provides the best coverage and density for a range of users.

For an office worker, Lenovo testing shows that 200 users per server is a good baseline and has an average of 65% usage of the processors in the server. If a server goes down, users on that server must be transferred to the remaining servers. For this degraded failover case, Lenovo testing shows that 240 users per server has an average of 78% usage of the processors. It is important to keep a 25% or more headroom on servers to cope with possible failover scenarios. Lenovo recommends a general failover ratio of 5:1.

For a knowledge worker, Lenovo testing shows that 165 users per server is a good baseline and has an average of 53% usage of the processors in the server. For the degraded failover case, Lenovo testing shows that 198 users per server have an average of 62% usage of the processors.

For a power user, Lenovo testing shows that 125 users per server is a good baseline and has an average of 57% usage of the processors in the server. For the degraded failover case, Lenovo testing shows that 150 users per server have an average of 69% usage of the processors. Moving up to two Xeon Scalable 8176 processors shows those densities increasing to 200 users and 240 users respectively with lower CPU utilizations.

Table 5 summarizes the processor usage with ESXi for the recommended user counts for normal mode and failover mode.

**Table 5: Processor usage for ESXi**

Processor	Workload	Users per Server	CPU Utilization
Two 6130	Office worker	200 users – Normal Mode	65%
Two 6130	Office worker	240 users – Failover Mode	78%
Two 6130	Knowledge worker	165 users – Normal Mode	53%
Two 6130	Knowledge worker	198 users – Failover Mode	62%
Two 6130	Power worker	125 users – Normal Mode	57%
Two 6130	Power worker	150 users – Failover Mode	69%

Table 6 summarizes the processor usage with Citrix Hypervisor for the recommended user counts for normal mode and failover mode.

**Table 6: Processor usage for Citrix Hypervisor**

Processor	Workload	Users per Server	CPU Utilization
Two 6130	Office worker	180 users – Normal Mode	71%
Two 6130	Office worker	216 users – Failover Mode	89%

Two 6130	Knowledge worker	165 users – Normal Mode	72%
Two 6130	Knowledge worker	198 users – Failover Mode	90%

Table 7 lists the recommended number of virtual desktops per server for different workload types and VM memory sizes. The number of users is reduced in some cases to fit within the available memory and still maintain a reasonably balanced system of compute and memory.

**Table 7: Recommended number of virtual desktops per server**

Workload	Office worker	Knowledge worker	Power worker
Processor	Two 6130	Two 6130	Two 6130
VM memory size	3 GB	4 GB	5 GB
System memory	768 GB	768 GB	768 GB
Desktops per server (normal mode)	200	165	125
Desktops per server (failover mode)	240	198	150

Table 8 lists the approximate number of compute servers that are needed for different numbers of users and Office worker workloads.

**Table 8: Compute servers needed for Office workers and different numbers of users**

Office workers	300 users	600 users	1200 users	3000 users
Compute servers @200 users (normal)	3	4	6	15
Compute servers @240 users (failover)	2	3	5	13

Table 9 lists the approximate number of compute servers that are needed for different numbers of users and Knowledge worker workloads.

**Table 9: Compute servers needed for Knowledge workers and different numbers of users**

Knowledge workers	300 users	600 users	1200 users	3000 users
Compute servers @165 users (normal)	3	4	8	18
Compute servers @198 users (failover)	2	3	6	15

Table 10 lists the approximate number of compute servers that are needed for different numbers of users and Power worker workloads.

**Table 10: Compute servers needed for Power workers and different numbers of users**

Power workers	300 users	600 users	1200 users	3000 users
Compute servers @125 users (normal)	3	5	10	24
Compute servers @150 users (failover)	2	4	8	20

### 4.3.2 Lenovo ThinkSystem with AMD EPYC™ Processors (Rome)

The new Lenovo ThinkSystem SR635 and ThinkSystem SR655 introduce the AMD EPYC 2<sup>nd</sup> generation 7002 series “Rome” processors. The EPYC processor is designed for high performing enterprise workloads which benefit from a high number of processor cores, high-speed memory, and scalable I/O via PCIe Gen 4 interconnect. Because of the EPYC processor’s high core density, systems with a single processor socket provide excellent compute capability, which supports a variety of workloads that typically require 2-socket servers. More information on the Lenovo ThinkSystem SR655 and 635 can be found at the following locations:

<https://lenovopress.com/lp1161-thinksystem-sr655-server>

<https://lenovopress.com/lp1160-thinksystem-sr635-server>

This section shows the performance results and sizing guidance for Lenovo ThinkSystem compute servers based on the AMD EPYC processor (Rome).

#### VMware ESXi 6.7 U3 performance results

We tested Login VSI benchmark with the 32-core AMD EPYC 7452 processor and the 64-core AMD EPYC 7742 processor.

Table 11 lists the VSI max performance results for different Login VSI 4.1 workloads with ESXi 6.7 U3 and Windows 10.

**Table 11: Login VSI Performance for ESXi**

Processor	Workload	Stateless	Dedicated
1x ThinkSystem AMD EPYC 7452 32C 155W 2.35 GHz	Office worker	150 users	150 users
	Knowledge worker	125 users	125 users
	Power worker	100 users	125 users
1x ThinkSystem AMD EPYC 7742 64C 225W 2.25 GHz	Office worker	280 users	280 users
	Knowledge worker	230 users	230 users
	Power worker	170 users	170 users

These results indicate the comparative processor performance. The following conclusions can be drawn:

- The performance for stateless and dedicated virtual desktops is similar.
- The 64-core AMD processor has significantly better performance over the 32-core processor, almost twice the performance when the number of cores is doubled, showing linear scaling with core count.

#### Compute server recommendation

The AMD EPYC 7002 series processor provides good cost/user density ratio. For moderate VDI desktops with office or knowledge worker profiles mostly, the low core count processors up to 32 cores will provide good desktop density. For high-end desktops with power user profiles requiring high compute and memory per desktop, the 64-core EPYC processor will provide good compute power as well as density.

The default recommendation is one AMD EPYC 7452, 2.35GHz, 32 core processor and 512 GB of system memory because this configuration provides the best coverage and density for a range of users.

For an office worker, Lenovo testing shows that 150 users per server is a good baseline and has an average of 71% usage of the processor in the server. If a server goes down, users on that server must be transferred to the remaining servers. For this degraded failover case, Lenovo testing shows that 180 users per server has an average of 81% usage of the processors. It is important to keep a 25% or more headroom on servers to cope with possible failover scenarios. Lenovo recommends a general failover ratio of 5:1.

For a knowledge worker, Lenovo testing shows that 125 users per server is a good baseline and has an average of 76% usage of the processors in the server. For the degraded failover case, Lenovo testing shows that 150 users per server have an average of 87% usage of the processors.

For a power user, Lenovo testing shows that 100 users per server is a good baseline and has an average of 72% usage of the processors in the server. For the degraded failover case, Lenovo testing shows that 125 users per server have an average of 83% usage of the processors. Moving up to the AMD EPYC 7742, 64-core, 2.25GHz processor shows those densities increasing to 280 users and 300 users respectively with lower CPU utilizations (office worker).

Table 12 summarizes the processor usage with ESXi for the recommended user counts for normal mode and failover mode for both stateless and dedicated desktops.

**Table 12: Processor usage for ESXi**

Processor	Workload	Stateless (# of users/ % CPU Util)		Dedicated (# of users/ % CPU Util)	
		Normal	Failover	Normal	Failover
1x ThinkSystem AMD EPYC 7452 32C 155W 2.35 GHz Processor	Office worker	150/ 71%	180/ 81%	150/ 72%	180/ 78%
	Knowledge worker	125/ 76%	150/ 87%	125/ 78%	150/ 85%
	Power worker	100/ 72%	125/ 83%	100/ 73%	125/ 80%
1x ThinkSystem AMD EPYC 7742 64C 225W 2.25 GHz Processor	Office worker	280/ 70%	300/ 82%	280/ 71%	300/ 80%
	Knowledge worker	230/ 75%	250/ 82%	230/ 73%	250/ 81%
	Power worker	170/ 74%	190/ 80%	170/ 73%	190/ 81%

Table 13 lists the recommended number of virtual desktops per server for different workload types and VM memory sizes with the AMD EPYC 7452, 32C processor. The number of users is reduced in some cases to fit within the available memory and still maintain a reasonably balanced system of compute and memory.

**Table 13: Recommended number of virtual desktops per server (7452, 32C CPU)**

Workload	Office worker	Knowledge worker	Power worker
Processor	1x 7452	1x 7452	1x 7452
VM vCPUs	1	2	4
VM memory size	2 GB	4 GB	6 GB
System memory	512 GB	1024 GB	1024 GB
Desktops per server (normal mode)	150	125	100
Desktops per server (failover mode)	180	150	125

Table 14 lists the approximate number of compute servers that are needed for different numbers of users and for various workload profiles using the AMD EPYC 7452, 32C processor.

**Table 14: Compute servers needed for various numbers of users (7452, 32C CPU)**

Office workers	300 users	600 users	1200 users	3000 users
Compute servers @150 users (normal)	2	4	8	20
Compute servers @180 users (failover)	2	4	7	17
<b>Knowledge workers</b>				
Compute servers @125 users (normal)	3	5	10	24
Compute servers @150 users (failover)	2	4	8	20
<b>Power workers</b>				
Compute servers @100 users (normal)	3	6	12	30
Compute servers @125 users (failover)	3	5	10	24

Table 15 lists the recommended number of virtual desktops per server for different workload types and VM memory sizes with the AMD EPYC 7742, 64C processor. The number of users is reduced in some cases to fit within the available memory and still maintain a reasonably balanced system of compute and memory.

**Table 15: Recommended number of virtual desktops per server (7742, 64C CPU)**

Workload	Office worker	Knowledge worker	Power worker
Processor	1x 7742	1x 7742	1x 7742
VM vCPUs	1	2	4
VM memory size	2 GB	4 GB	6 GB
System memory	1024 GB	1024 GB	1024 GB
Desktops per server (normal mode)	280	230	170
Desktops per server (failover mode)	300	250	190

Table 16 lists the approximate number of compute servers that are needed for different numbers of users and for various workload profiles using the AMD EPYC 7742, 64C processor.

**Table 16: Compute servers needed for various numbers of users (7742, 64C CPU)**

<b>Office workers</b>	<b>300 users</b>	<b>600 users</b>	<b>1200 users</b>	<b>3000 users</b>
Compute servers @280 users (normal)	2	3	5	11
Compute servers @300 users (failover)	2	2	4	10
<b>Knowledge workers</b>				
Compute servers @230 users (normal)	2	3	6	14
Compute servers @250 users (failover)	2	3	5	12
<b>Power workers</b>				
Compute servers @170 users (normal)	2	4	7	18
Compute servers @190 users (failover)	2	3	6	16

## 4.4 Compute servers for VMware vSAN

This section presents the compute servers for the hyper-converged system using VMware vSAN 6.6 with ESXi 6.5 U1 for Purley Skylake and VMware vSAN 6.7 with ESXi 6.7 U3 for Purley Cascade lake. In this section stateless refers to MCS Pooled-Random and dedicated refers to MCS Dedicated.

Additional processing and memory is required to support vSAN. Systems with two disk groups need approximately 32 GB and those with 4 disk groups need 50GB. See the following for more details on the calculation of vSAN system memory:

[kb.vmware.com/selfservice/microsites/search.do?cmd=displayKC&externalId=2113954](http://kb.vmware.com/selfservice/microsites/search.do?cmd=displayKC&externalId=2113954)

As the price per GB for flash memory continues to reduce, there is a trend to use all SSDs to provide the overall best performance for vSAN.

For more information, see “BOM for hyper-converged compute servers” on page 43.

### 4.4.1 Lenovo ThinkSystem with Intel Xeon Scalable Processor Gen 1 (Skylake)

VMware vSAN is tested by using Login VSI 4.1. Four Lenovo ThinkSystem SR630 servers with two Xeon Scalable 6130 processors and 768 GB of system memory were networked together by using a 10 GbE TOR switch with two 10 GbE connections per server.

Each server was configured with two all flash disk groups because a single disk group does not provide the necessary resiliency. Each disk group had one 800 GB SSD and two 3.84 GB SSDs. This provided more than enough capacity for linked clone or full clone VMs. Additional 3.84 GB SSDs could be added for more capacity but may not be needed when all flash de-duplication is used.

Table 11 lists the per server VSI max performance results for different Login VSI 4.1 workloads with ESXi 6.5 U1 and Windows 10. Linked clones were used with the VMware vSAN default storage policy of number of failures to tolerate (FTT) of 0 and stripe of 1. No de-duplication or compression was used.

**Table 17: Login VSI Performance with ESXi 6.5 U1**

Processor	Workload	Stateless	Dedicated
Two Scalable 6130 processors 2.10 GHz, 16C 125W	Office worker	266 users	269 users
Two Scalable 6130 processors 2.10 GHz, 16C 125W	Knowledge worker	210 users	210 users
Two Scalable 6130 processors 2.10 GHz, 16C 125W	Power worker	185 users	185 users

Table 12 lists the per server VSI max performance results for different Login VSI 4.1 workloads with ESXi 6.7, Windows 10 and the operating system and firmware patches for the side channel security issues. For more information, please see the Lenovo Product Security Advisories:

[support.lenovo.com/us/en/product\\_security/home](http://support.lenovo.com/us/en/product_security/home).

Linked clones were used with the VMware default storage policy of number of failures to tolerate (FTT) of 0 and stripe of 1. No de-duplication or compression was used.



**Table 18: Login VSI Performance with ESXi 6.7**

Processor	Workload	Stateless	Dedicated
Two Scalable 6130 processors 2.10 GHz, 16C 125W	Office worker	223 users	238 users
Two Scalable 6130 processors 2.10 GHz, 16C 125W	Knowledge worker	180 users	193 users
Two Scalable 6130 processors 2.10 GHz, 16C 125W	Power worker	162 users	172 users

The Xeon Scalable 6130 processor has a good cost/user density ratio. Many configurations are bound by memory; therefore, a faster processor might not provide any added value.

The default recommendation is two Xeon Scalable 6130 processors and 768 GB of system memory because this configuration provides the best coverage and density for a range of users.

For an office worker, Lenovo testing shows that 150 users per server is a good baseline and has an average of 66% usage of the processors in the server. If a server goes down, users on that server must be transferred to the remaining servers. For this degraded failover case, Lenovo testing shows that 180 users per server have an average of 80% usage of the processors. It is important to keep this 25% headroom on servers to cope with possible failover scenarios. Lenovo recommends a general failover ratio of 5:1. By using a target of 150 users per server, the maximum number of office workers is 9,600 in a 64 node cluster.

For a knowledge worker, Lenovo testing shows that 125 users per server is a good baseline and has an average of 68% usage of the processors in the server. For the degraded failover case, Lenovo testing shows that 150 users per server have an average of 81% usage of the processors. By using a target of 125 users per server, the maximum number of knowledge workers is 8,000 in a 64 node cluster.

For a power user, Lenovo testing shows that 110 users per server is a good baseline and has an average of 67% usage of the processors in the server. For the degraded failover case, Lenovo testing shows that 132 users per server have an average of 81% usage of the processors. By using a target of 110 users per server, the maximum number of knowledge workers is 7,040 in a 64 node cluster.

Table 13 lists the recommended number of virtual desktops per server for different workload types and VM memory sizes. The number of users is reduced in some cases to fit within the available memory and still maintain a reasonably balanced system of compute and memory.

**Table 19: Recommended number of virtual desktops per server**

Workload	Office worker	Knowledge worker	Power worker
Processor	Two 6130	Two 6130	Two 6130
VM memory size	3 GB	4 GB	5 GB
System memory	768 GB	768 GB	768 GB
Memory overhead of vSAN	32 GB	32 GB	32 GB
Desktops per server (normal mode)	150	125	110
Desktops per server (failover mode)	180	150	132

Table 14 lists the approximate number of compute servers that are needed for different numbers of users and Office worker workloads.

**Table 20: Compute servers needed for Office workers and different numbers of users**

Office workers	300 users	600 users	1200 users	3000 users
Compute servers @150 users (normal)	3	5	8	20
Compute servers @180 users (failover)	2	4	7	17

Table 15 lists the approximate number of compute servers that are needed for different numbers of users and Knowledge worker workloads.

**Table 21: Compute servers needed for Knowledge workers and different numbers of users**

Knowledge workers	300 users	600 users	1200 users	3000 users
Compute servers @125 users (normal)	3	5	10	24
Compute servers @150 users (failover)	2	4	8	20

Table 16 lists the approximate number of compute servers that are needed for different numbers of users and Power worker workloads.

**Table 22: Compute servers needed for Power workers and different numbers of users**

Power workers	300 users	600 users	1200 users	3000 users
Compute servers @110 users (normal)	4	6	11	27
Compute servers @132 users (failover)	3	5	9	23

The Login VSI measurement above use a default login interval of 30 seconds per server. Table 17 shows what happens when the login interval for a cluster of 4 all flash servers is increased as might happen during a login storm. In order to maintain the 4 second response time, the number of supported desktops would need to be reduced.

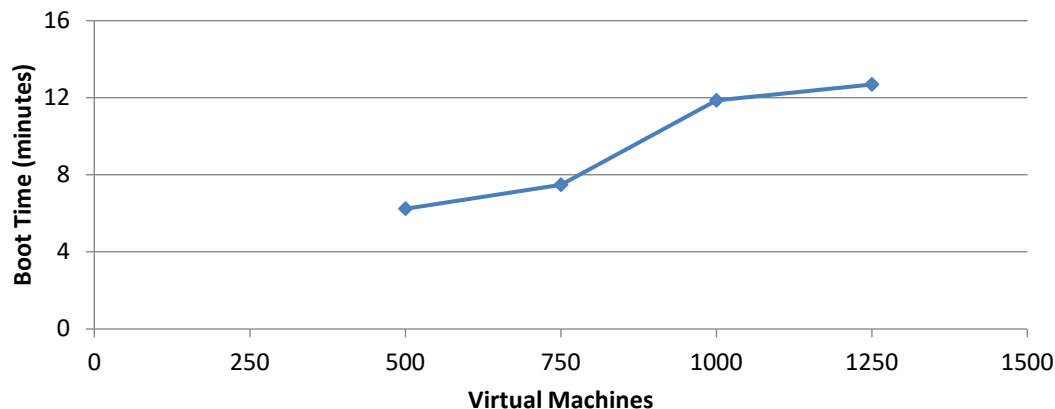
**Table 23: Performance during login storm with ESXI 6.5 U1**

Workload	Number of VMs	Login Interval	Stateless Linked Clones	Dedicated Linked Clones
Knowledge worker	940	Every 30 seconds	842 users	842 users
Knowledge worker	940	Every 10 seconds	742 users	835 users
Knowledge worker	940	Every 5 seconds	528 users	735 users
Knowledge worker	940	Every 2 seconds	243 users	456 users

A boot storm occurs when a substantial number of VMs are all started within a short period of time. Booting a large number of VMs simultaneously requires large IOPS otherwise the VMs become slow and unresponsive.

Different numbers of VMs were booted on a cluster of 4 vSAN all flash servers. The VMs were unpowered in vCenter and the boot storm created by powering on all of the VMs simultaneously. The time for all of the VMs to become visible was measured.

Figure 7 shows the boot times for different numbers of VMs in the 4 node cluster. With an even spread of VMs on each node, the boot time for the VMs on each node was similar to the overall cluster boot time. These results show that even a large number of VMs can be very quickly restarted using a vSAN all flash cluster.



**Figure 7: Boot storm performance**

#### 4.4.2 Lenovo ThinkAgile VX3320 Appliance with Intel Xeon Scalable Processor Gen 2 (Cascadelake)

VMware vSAN 6.7 is tested by using Login VSI 4.1.39. Four Lenovo ThinkAgile VX3320 Appliances with two Xeon Scalable Gold 6230 processors and 1536 GB (24x64GB DIMM) of system memory were networked together by using a 10 GbE TOR switch with two 10 GbE connections per server.

Each server was configured with two hybrid disk groups because a single disk group does not provide the necessary resiliency. Each disk group had one 800 GB SSD and three 1.84 GB HDDs. This provided more than enough capacity for linked clone or full clone VMs. Additional 1.84 GB HDDs could be added for more capacity.

Table 18 lists the per server VSI max performance results for different Login VSI 4.1 workloads with ESXi 6.7 U3 and Windows 10. Linked clones were used with the VMware vSAN default storage policy of number of failures to tolerate (FTT) of 0 and stripe of 1. No de-duplication or compression was used.

**Table 24: Login VSI Performance with ESXi 6.7 U3**

Processor	Workload	Stateless	Dedicated
Two Scalable Gold 6230 processors 2.10 GHz, 20C 125W	Office worker	262 users	264 users
Two Scalable Gold 6230 processors 2.10 GHz, 20C 125W	Knowledge worker	236 users	235 users
Two Scalable Gold 6230 processors 2.10 GHz, 20C 125W	Power worker	195 users	195 users

The Xeon Scalable Gold 6230 processor has a good cost/user density ratio. Many configurations are bound by memory; therefore, a faster processor might not provide any added value.

The default recommendation is two Xeon Scalable 6230 processors and 768 GB of system memory because this configuration provides the best coverage and density for a range of users.

For an office worker, Lenovo testing shows that 160 users per server is a good baseline and has an average of 74% usage of the processors in the server. If a server goes down, users on that server must be transferred to the remaining servers. For this degraded failover case, Lenovo testing shows that 192 users per server have an average of 83% usage of the processors. It is important to keep this 25% headroom on servers to cope with possible failover scenarios. Lenovo recommends a general failover ratio of 5:1. By using a target of 160 users per server, the maximum number of office workers is 10,240 in a 64 node cluster.

For a knowledge worker, Lenovo testing shows that 140 users per server is a good baseline and has an average of 73% usage of the processors in the server. For the degraded failover case, Lenovo testing shows that 168 users per server have an average of 86% usage of the processors. By using a target of 140 users per server, the maximum number of knowledge workers is 8,960 in a 64 node cluster.

For a power user, Lenovo testing shows that 120 users per server is a good baseline and has an average of 74% usage of the processors in the server. For the degraded failover case, Lenovo testing shows that 144 users per server have an average of 87% usage of the processors. By using a target of 120 users per server, the maximum number of knowledge workers is 7,680 in a 64 node cluster.

Table 19 lists the recommended number of virtual desktops per server for different workload types and VM memory sizes. The number of users is reduced in some cases to fit within the available memory and still maintain a reasonably balanced system of compute and memory.

**Table 25: Recommended number of virtual desktops per server**

Workload	Office worker	Knowledge worker	Power worker
Processor	Two 6230	Two 6230	Two 6230
VM memory size	3 GB	4 GB	5 GB
System memory	768 GB	768 GB	768 GB
Memory overhead of vSAN	32 GB	32 GB	32 GB
Desktops per server (normal mode)	160	140	120
Desktops per server (failover mode)	192	168	144

Table 20 lists the approximate number of compute servers that are needed for different numbers of users and Office worker workloads.

**Table 26: Compute servers needed for Office workers and different numbers of users**

Office workers	300 users	600 users	1200 users	3000 users
Compute servers @160 users (normal)	3	4	8	19
Compute servers @192 users (failover)	2	3	7	16

Table 21 lists the approximate number of compute servers that are needed for different numbers of users and Knowledge worker workloads.

**Table 27: Compute servers needed for Knowledge workers and different numbers of users**

Knowledge workers	300 users	600 users	1200 users	3000 users
-------------------	-----------	-----------	------------	------------

Compute servers @140 users (normal)	3	5	9	22
Compute servers @168 users (failover)	2	4	8	18

Table 22 lists the approximate number of compute servers that are needed for different numbers of users and Power worker workloads.

**Table 28: Compute servers needed for Power workers and different numbers of users**

Power workers	300 users	600 users	1200 users	3000 users
Compute servers @120 users (normal)	3	5	10	25
Compute servers @144 users (failover)	2	4	9	21

## 4.5 Graphics acceleration

The VMware ESXi hypervisor supports the following options for graphics acceleration:

- Dedicated GPU with one GPU per user, which is called virtual dedicated graphics acceleration (vDGA) mode.
- GPU hardware virtualization (vGPU) that partitions each GPU.
- Shared GPU with users sharing a GPU, which is called virtual shared graphics acceleration (vSGA) mode and is not recommended because of user contention for shared use of the GPU.

VMware also provides software emulation of a GPU, which can be processor-intensive and disruptive to other users who have a choppy experience because of reduced processor performance. Software emulation is not recommended for any user who requires graphics acceleration.

The Hypervisor hypervisor supports the following options for graphics acceleration:

- Pass-through mode (which is similar to vDGA under ESXi)
- GPU hardware virtualization (vGPU) that partitions each GPU (similar to vGPU under ESXi)

The vDGA (or pass-through) option has a low user density as it restricts a single user to access each very powerful GPU. This option is not flexible and is no longer cost effective even for high-end power users. Therefore vDGA is no longer recommended especially given that the performance of the equivalent vGPU mode is similar.

When using the vGPU option with ESXi 6.5 and the latest drivers from NVidia, it is necessary to change the default GPU mode from “Shared” (vSGA) to “Shared Direct” (vGPU) for each GPU using VMware vCenter. This enables the correct GPU support for the VMs which would otherwise result in the VM not powering on correctly and the standard “graphics resources not available” error message. The host needs to be rebooted for the changes to take effect.

There is also a known problem with VMware vSphere that may prevent the VM from booting correctly. Lenovo recommends always setting “pciPassthru.use64bitMMIO” option to TRUE regardless of the memory size of the VM by choosing “VM Options > Advanced”. See also the NVidia release notes for VMware vSphere:

[docs.nvidia.com/grid/latest/grid-vgpu-release-notes-vmware-vsphere/index.html](https://docs.nvidia.com/grid/latest/grid-vgpu-release-notes-vmware-vsphere/index.html).

The performance of graphics acceleration was tested using the Lenovo ThinkSystem SR650 servers. Each server supports up to two GPU adapters. The Heaven benchmark is used to measure the frames per second (FPS) rate per user for different GPUs, resolutions, and image quality. This benchmark is graphics-heavy and is fairly realistic for designers and engineers. Power users or knowledge workers usually have less intense graphics workloads and can achieve higher frame rates.

Table 23 lists the results of the Heaven benchmark as FPS that are available to each user with the NVidia M10 GPU using different vGPU profiles with Citrix XenDesktop 7.15, DirectX 11, and ESXi 6.5U1.

**Table 29: Performance of NVidia M10 vGPU profiles**

Quality	Tessellation	Anti-Aliasing	Resolution	M10-8Q	M10-4Q	M10-2Q	M10-4A	M10-2A
High	Normal	0	1280x1024	Untested	Untested	13.8	25.6	14.0
High	Normal	0	1680x1050	Untested	Untested	10.3	N/A	N/A
High	Normal	0	1920x1200	Untested	17.5	Untested	N/A	N/A
Ultra	Extreme	8	1280x1024	Untested	Untested	6.3	13.5	6.9
Ultra	Extreme	8	1680x1050	21.3	11.4	Untested	N/A	N/A
Ultra	Extreme	8	1920x1080	19.2	9.6	Untested	N/A	N/A
Ultra	Extreme	8	1920x1200	17.5	8.4	Untested	N/A	N/A
Ultra	Extreme	8	2560x1600	12.0	Untested	Untested	N/A	N/A

Table 24 lists the results of the Heaven benchmark as FPS that are available to each user with the NVidia P40 GPU using different vGPU profiles with Citrix XenDesktop 7.15, DirectX 11, and Citrix Hypervisor 7.3. The results with ESXi 6.7 are similar.

**Table 30: Performance of NVidia P40 vGPU profiles**

Quality	Tessellation	Anti-Aliasing	Resolution	P40-24Q	P40-12Q	P40-6Q	P40-4Q
High	Normal	0	1280x1024	Untested	63.0	59.6	43.1
High	Normal	0	1680x1050	Untested	62.8	51.9	40.0
High	Normal	0	1920x1200	Untested	62.0	44.9	30.0
Ultra	Extreme	8	1280x1024	Untested	60.4	38.9	25.8
Ultra	Extreme	8	1680x1050	62.7	53.7	31.8	21.1
Ultra	Extreme	8	1920x1080	62.4	49.6	28.6	19.0
Ultra	Extreme	8	1920x1200	62.3	40.6	26.7	17.6
Ultra	Extreme	8	2560x1600	57.8	32.8	17.4	Untested
Ultra	Extreme	8	3840x2160	34.0	16.9	Untested	Untested
Ultra	Extreme	8	4096x2160	32.0	15.5	Untested	Untested

Lenovo recommends that a medium to high powered CPU, such as the Xeon Scalable 6130, is used for accelerated graphics applications tend to also require extra load on the processor. For vGPU mode, Lenovo recommends at least 384GB of server memory. Because there are many variables when graphics acceleration is used, Lenovo recommends that testing is done in the customer environment to verify the performance for the required user workloads.

For more information about the bill of materials (BOM) for NVidia GPUs for Lenovo ThinkSystem SR650 servers, see the following corresponding BOMs:

- “BOM for enterprise and SMB compute servers” section on page 38.
- “BOM for hyper-converged compute servers” on page 43.

## 4.6 Management servers

Management servers should have the same hardware specification as compute servers so that they can be used interchangeably in a worst-case scenario. The Citrix Virtual Apps and Desktops management servers also use the same hypervisor, but have management VMs instead of user desktops.

Table 25 lists the VM requirements and performance characteristics of each management service.

**Table 31: Characteristics of Virtual Apps and Desktops and ESXi management services**

Management service VM	Virtual processors	System memory	Storage	Windows OS	HA needed	Performance characteristic
Delivery controller	4	8 GB	60 GB	2016	Yes	5000 user connections
Web Interface	4	4 GB	60 GB	2016	Yes	30,000 connections per hour
Citrix licensing server	2	4 GB	60 GB	2016	No	170 licenses per second
Virtual Apps and Desktops SQL server	2	8 GB	60 GB	2016	Yes	5000 users
PVS servers	4	32 GB	60 GB (depends on number of images)	2016	Yes	Up to 1000 desktops, memory should be a minimum of 2 GB plus 1.5 GB per image served
vCenter server	8	16 GB	60 GB	2016	No	Up to 2000 desktops
vCenter SQL server	4	8 GB	200 GB	2016	Yes	Double the virtual processors and memory for more than 2500 users

PVS servers often are run natively on Windows servers. The testing showed that they can run well inside a VM, if it is sized per Table 25. The disk space for PVS servers is related to the number of provisioned images.

Table 26 lists the number of management VMs for each size of users following the high availability and performance characteristics. The number of vCenter servers is half of the number of vCenter clusters because each vCenter server can handle two clusters of up to 1000 desktops.

**Table 32: Management VMs needed**

Virtual Apps and Desktops management service VM	300 users	600 users	1200 users	3000 users
Delivery Controllers	2 (1+1)	2 (1+1)	2 (1+1)	2 (1+1)
Includes Citrix Licensing server	Y	Y	N	N
Includes Web server	Y	Y	N	N
Web Interface	N/A	N/A	2 (1+1)	2 (1+1)
Citrix licensing servers	N/A	N/A	1	1
Virtual Apps and Desktops SQL servers	2 (1+1)	2 (1+1)	2 (1+1)	2 (1+1)



<b>Virtual Apps and Desktops management service VM</b>	<b>300 users</b>	<b>600 users</b>	<b>1200 users</b>	<b>3000 users</b>
PVS servers for stateless case only	2 (1+1)	2 (1+1)	2 (1+1)	2 (1+1)
vCenter servers	1	1	2	2
vCenter SQL servers	2 (1+1)	2 (1+1)	2 (1+1)	2 (1+1)

Each management VM requires a certain amount of virtual processors, memory, and disk. There is enough capacity in the management servers for all of these VMs. Table 27 lists an example mapping of the management VMs to the two physical management servers for 3000 users.

**Table 33: Management server VM mapping (3000 users)**

<b>Management service for 3000 stateless users</b>	<b>Management server 1</b>	<b>Management server 2</b>
vCenter servers (2)	1	1
vCenter SQL server (2)	1	1
Virtual Apps and Desktops SQL server (2)	1	1
Delivery controller (2)	1	1
Web Interface (2)	1	1
License server (1)		1
PVS servers for stateless desktops (2)	1	1

It is assumed that common services, such as Microsoft Active Directory, Dynamic Host Configuration Protocol (DHCP), domain name server (DNS), and Microsoft licensing servers exist in the customer environment.

For shared storage systems that support block data transfers only, it is also necessary to provide some file I/O servers that support CIFS or NFS shares and translate file requests to the block storage system. For high availability, two or more Windows storage servers are clustered.

Based on the number and type of desktops, Table 28 lists the recommended number of physical management servers. In all cases, there is redundancy in the physical management servers and the management VMs. Note for a small number of users the management VMs could be executed on the same servers as the virtual desktop VMs.

**Table 34: Management servers needed**

<b>Management servers</b>	<b>300 users</b>	<b>600 users</b>	<b>1200 users</b>	<b>3000 users</b>
Stateless desktop model	0	0	2	2
Dedicated desktop model	0	0	2	2
Windows Storage Server 2012	2	2	2	2

For more information, see “BOM for enterprise and SMB management servers” on page 49.

## 4.7 Systems management

Lenovo XClarity™ Administrator is a centralized resource management solution that reduces complexity, speeds up response, and enhances the availability of Lenovo® server systems and solutions.

The Lenovo XClarity Administrator provides agent-free hardware management for Lenovo's ThinkSystem, System x® rack servers and Flex System™ compute nodes and components, including the Chassis Management Module (CMM) and Flex System I/O modules. Figure 8 shows the Lenovo XClarity administrator interface, where Flex System components and rack servers are managed and are seen on the dashboard. Lenovo XClarity Administrator is a virtual appliance that is quickly imported into a virtualized environment server configuration.

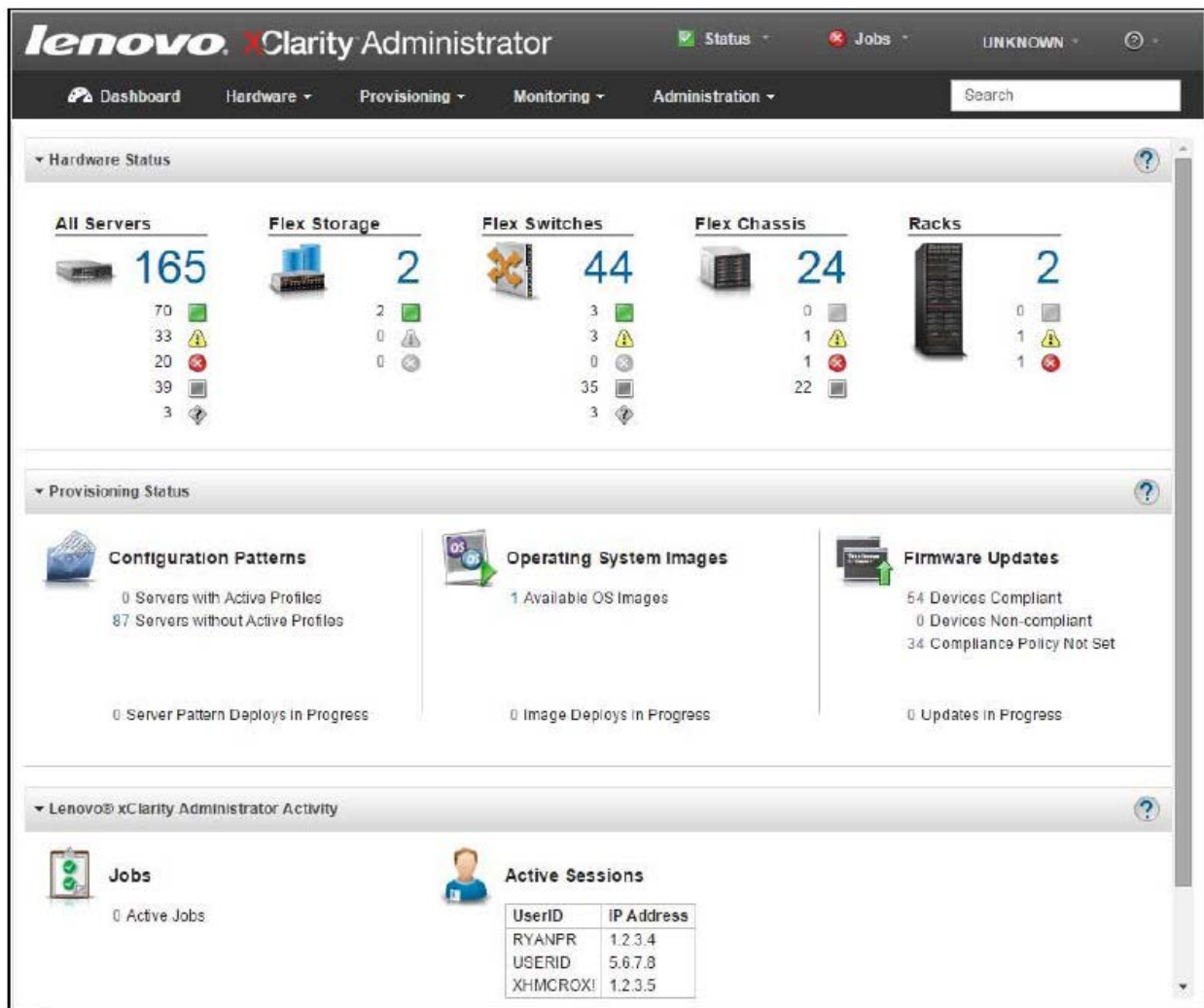


Figure 8: XClarity Administrator interface

## 4.8 Shared storage

VDI workloads, such as virtual desktop provisioning, VM loading across the network, and access to user profiles and data files place huge demands on network shared storage.

Experimentation with VDI infrastructures shows that the input/output operation per second (IOPS) performance takes precedence over storage capacity. This precedence means that more slower speed drives are needed to match the performance of fewer higher speed drives. Even with the fastest HDDs available today (15k rpm), there can still be excess capacity in the storage system because extra spindles are needed to provide the IOPS performance. From experience, this extra storage is more than sufficient for the other types of data such as SQL databases and transaction logs.

The large rate of IOPS, and therefore, large number of drives needed for dedicated virtual desktops can be ameliorated to some extent by caching data in flash memory or SSD drives. The storage configurations are based on the peak performance requirement, which usually occurs during the so-called “logon storm.” This is when all workers at a company arrive in the morning and try to start their virtual desktops, all at the same time.

It is always recommended that user data files (shared folders) and user profile data are stored separately from the user image. By default, this has to be done for stateless virtual desktops and should also be done for dedicated virtual desktops. It is assumed that 100% of the users at peak load times require concurrent access to user data and profiles.

In View 5.1, VMware introduced the View Storage Accelerator (VSA) feature that is based on the ESXi Content-Based Read Cache (CBRC). VSA provides a per-host RAM-based solution for VMs, which considerably reduces the read I/O requests that are issued to the shared storage. It is recommended to use the maximum VSA cache size of 2GB. Performance measurements by Lenovo show that VSA has a negligible effect on the number of virtual desktops that can be used on a compute server while it can reduce the read requests to storage by one-fifth.

Table 29 summarizes the peak IOPS and disk space requirements for virtual desktops on a per-user basis. Dedicated virtual desktops require a high number of IOPS and a large amount of disk space for the linked clones. Note that the linked clones also can grow in size over time.

**Table 35: Shared virtual desktop shared storage performance requirements**

Shared virtual desktops	Protocol	Size	IOPS	Write %
Master image	NFS or Block	30 GB	18	85%
Difference disks	NFS or Block	10 GB		
User “AppData” folder				
vSwap (recommended to be disabled)	NFS or Block	0	0	0
User files	CIFS/NFS	5 GB	1	75%
User profile (through MSRP)	CIFS	100 MB	0.8	75%

The sizes and IOPS for user data files and user profiles that are listed in Table 29 can vary depending on the customer environment. For example, power users might require 10 GB and five IOPS for user files because of

the applications they use. It is assumed that 100% of the users at peak load times require concurrent access to user data files and profiles.

Many customers need a hybrid environment of stateless and dedicated desktops for their users. The IOPS for dedicated users outweigh those for stateless users; therefore, it is best to bias towards dedicated users in any storage controller configuration.

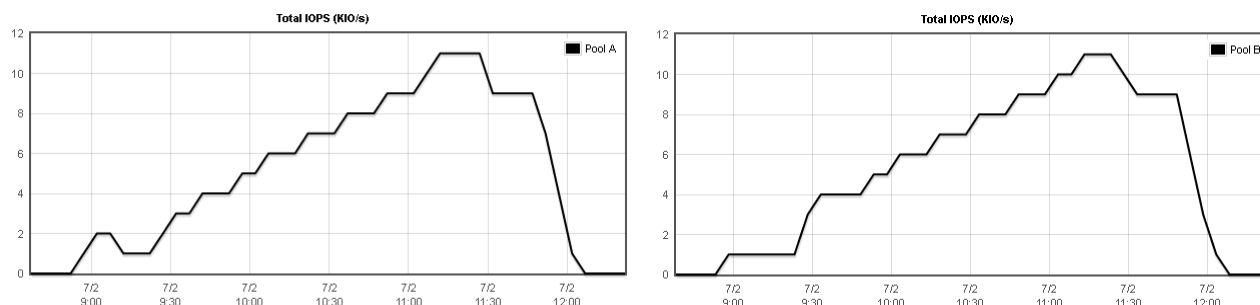
The storage configurations that are presented in this section include conservative assumptions about the VM size, changes to the VM, and user data sizes to ensure that the configurations can cope with the most demanding user scenarios.

### 4.8.1 Lenovo ThinkSystem DS6200 storage array

The Lenovo ThinkSystem DS6200 storage array supports up to 240 drives by using up to 9 expansion enclosures. Drives can be HDDs or SSDs in a variety of sizes and speeds. A SSD based cache can be provided for each of the redundant controllers. This tiered storage support also allows a mixture of different disk drives. Slower drives can be used for shared folders and profiles; faster drives and SSDs can be used for dedicated virtual desktops and desktop images.

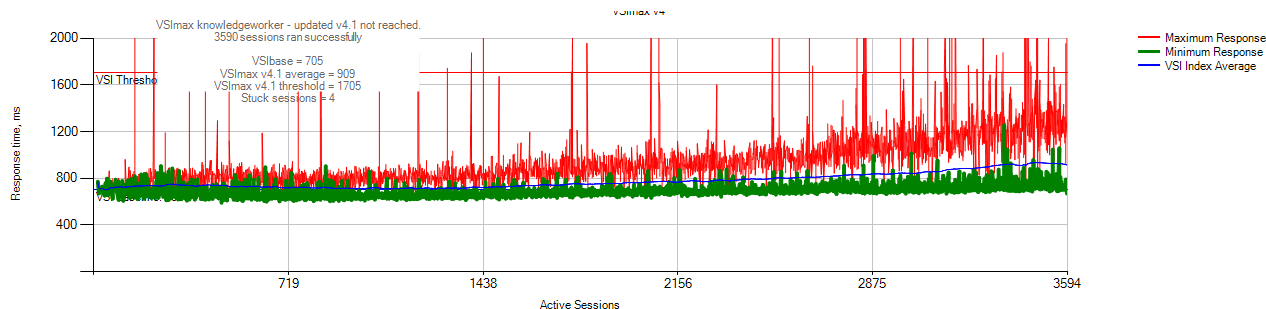
To support file I/O (CIFS and NFS) into Lenovo storage, Windows storage servers must be added, as described in “Management servers” on page 28.

Experimentation with the DS6200 shows that an array of 20 SSDs can support at least 3576 stateless users assuming a new login every 2 seconds. Figure 9 shows the IOP graphs for two the controller during the Login VSI test of 4000 users. The graphs show a maximum of 22,000 IOPS or about 6 per VM.



**Figure 9: DS6200 IOPs for all SSD array**

Figure 10 shows the Login VSI chart for the test of 3590 users. The response time linearly increases but never consistently exceeds the 4 second threshold. The maximum number of stateless desktops is approximately 4000.



**Figure 10: Login VSI Chart for all SSD array**

Testing with a hybrid array of SSDs and HDDs shows good performance until the SSD cache is full and then the performance drops sharply.

Lenovo recommends using all SSDs to provide the best performance both in terms of general usage and for boot and login storms. Note that de-duplication and compression features are not available. Therefore the required disk capacity can be determined from the storage capacity of the linked clone or full clone VMs, master images, and user storage.

Table 30 lists the storage configuration that is needed for each of the different numbers of stateless users. It is assumed that the master images are 40 GB (actual) with linked clones of 3GB each. Each user requires up to 5GB of persistent storage. The total storage includes 30% capacity for growth.

**Table 36: Storage configuration for stateless users**

Stateless storage	300 users	600 users	1200 users	3000 users
Master image count	1	1	2	4
Master image per controller with backup	160 GB	160 GB	320 GB	640 GB
Stateless linked clones (no growth)	1800 GB	3600 GB	7200 GB	18000 GB
User storage	1500 GB	3000 GB	3600 GB	15000 GB
Total storage (with 30% growth)	4.4 TB	8.6 TB	14 TB	33 TB
Number of 800 GB SSDs (RAID 5)	8	14	24	0
Number of 3.84 GB SSDs (RAID 5)	0	0	0	12
ThinkSystem DS6200 control enclosures	1	1	1	1

Table 31 lists the storage configuration that is needed for each of the different numbers of dedicated users. It is assumed that the master images are 40 GB (actual) with linked clone growth of up to 50%. Each user requires up to 5GB of persistent storage.

**Table 37: Storage configuration for dedicated users**

Stateless storage	300 users	600 users	1200 users	3000 users
Master image count	1	1	2	4
Master image per controller with backup	160 GB	160 GB	320 GB	640 GB
Stateless linked clones (50% growth)	6900 GB	13800 GB	27600 GB	69000 GB
User storage	1500 GB	3000 GB	3600 GB	15000 GB
Total storage (no growth)	9 TB	16.6 TB	31 TB	83 TB
Number of 800 GB SSDs (RAID 5)	16	24	0	0
Number of 3.84 GB SSDs (RAID 5)	0	0	10	24
ThinkSystem DS6200 control enclosures	1	1	1	1

Refer to the “BOM for shared storage” on page 51 for more details.

## 4.9 Networking

The main driver for the type of networking that is needed for VDI is the connection to storage. If the shared storage is block-based (such as the Lenovo ThinkSystem DS6200 storage array), it is likely that a SAN that is based on 8 or 16 Gbps FC, or 10 GbE iSCSI connection is needed. Other types of storage can be network attached by using 1 Gb or 10 Gb Ethernet. For vSAN 10 GbE networking should always be used.

Also, there is user and management virtual local area networks (VLANs) available that require 1 Gb or 10 Gb Ethernet as described in the Lenovo Client Virtualization reference architecture, which is available at this website: [lenovopress.com/lp0756](http://lenovopress.com/lp0756).

Automated failover and redundancy of the entire network infrastructure and shared storage is important. This failover and redundancy is achieved by having at least two of everything and ensuring that there are dual paths between the compute servers, management servers, and shared storage.

If only a single Flex System Enterprise Chassis is used, the chassis switches are sufficient and no other TOR switch is needed. For rack servers, more than one Flex System Enterprise Chassis TOR switches are required.

For more information, see “BOM for networking” on page 51.

### 4.9.1 10 GbE networking

For 10 GbE networking, the use of CIFS, NFS, or iSCSI, the Lenovo RackSwitch™ G8124E, and G8272 TOR switches are recommended because they support VLANs by using Virtual Fabric. Redundancy and automated failover is available by using link aggregation, such as Link Aggregation Control Protocol (LACP) and two of everything. For the Flex System chassis, pairs of the EN4093R switch should be used and connected to a G8124 or G8272 TOR switch. The TOR 10GbE switches are needed for multiple Flex chassis or external connectivity. iSCSI also requires converged network adapters (CNAs) that have the LOM extension. Table 32 lists the TOR 10 GbE network switches for each user size.

**Table 38: TOR 10 GbE network switches needed**

10 GbE TOR network switch	300 users	600 users	1200 users	3000 users
G8124E – 24-port switch	2	2	2	0
G8272 – 64-port switch	0	0	0	2

## 4.9.2 1 GbE administration networking

A 1 GbE network should be used to administer all of the other devices in the system. Separate 1 GbE switches are used for the IT administration network. Lenovo recommends that redundancy is also built into this network at the switch level. At minimum, a second switch should be available in case a switch goes down. Table 33 lists the number of 1 GbE switches for each user size.

**Table 39: TOR 1 GbE network switches needed**

1 GbE TOR network switch	300 users	600 users	1200 users	3000 users
G8052 – 48 port switch	2	2	2	2

Table 34 shows the number of 1 GbE connections that are needed for the administration network and switches for each type of device. The total number of connections is the sum of the device counts multiplied by the number of each device.

**Table 40: 1 GbE connections needed**

Device	Number of 1 GbE connections for administration
ThinkSystem rack server	1
Flex System Enterprise Chassis CMM	2
Flex System Enterprise Chassis switches	1 per switch (optional)
Lenovo ThinkSystem DS6200 storage array	4
TOR switches	1

## 4.10 Racks

The number of racks and chassis for Flex System compute nodes depends upon the precise configuration that is supported and the total height of all of the component parts: servers, storage, networking switches, and Flex System Enterprise Chassis (if applicable). The number of racks for ThinkSystem servers is also dependent on the total height of all of the components. For more information, see the “BOM” section on page 53.

## 4.11 Deployment models

This section contains the following example deployment models:

- vSAN hyper-converged using ThinkSystem servers

### 4.11.1 Deployment example 1: vSAN hyper-converged using ThinkSystem servers

This deployment example supports 1200 dedicated virtual desktops with VMware vSAN in hybrid mode (SSDs and HDDs). Each VM needs 3 GB of RAM. Six servers are needed assuming 200 users per server using the Office worker profile. There can be up to 1 complete server failure and there is still capacity to run the 1200 users on five servers (ratio of 5:1).

Each virtual desktop with Windows 10 takes 40 GB. The total capacity would be 126 TB, assuming full clones of 1200 VMs, one complete mirror of the data (FFT=1), space for swapping, and 30% extra capacity. However, the use of linked clones for vSAN or de-duplication for all flash means that a substantially smaller capacity is sufficient.

This example uses six ThinkSystem SR630 1U servers and two G8124E RackSwitch TOR switches. Each server has two Xeon Scalable 6130 processors, 768 GB of memory, two network cards, two 800 GB SSDs, and six 1.2TB HDDs. Two disk groups of one SSD and three HDDs are used for vSAN.

Figure 11 shows the deployment configuration for this example.



**Figure 11: Deployment configuration for 1200 dedicated users using vSAN on ThinkSystem servers**



## 5 Appendix: Bill of materials

This appendix contains the bill of materials (BOMs) for different configurations of hardware. There are sections for user servers, management servers, storage, networking switches, chassis, and racks that are orderable from Lenovo. The last section is for hardware orderable from an OEM.

The BOM lists in this appendix are not meant to be exhaustive and must always be double-checked with the configuration tools. Any discussion of pricing, support, and maintenance options is outside the scope of this document.

For connections between TOR switches and devices (servers, storage, and chassis), the connector cables are configured with the device. The TOR switch configuration includes only transceivers or other cabling that is needed for failover or redundancy.

### 5.1 BOM for enterprise and SMB compute servers

This section contains the bill of materials for enterprise and SMB compute servers. Table 35, Table 36, and Table 37 list the BOM for a server (a minimum of two servers per cluster is needed). See Table 54 on page 53 for the Lenovo Flex System chassis BOM.

**Table 41: ThinkSystem SR630 (Xeon SP Gen 1)**

Code	Description	Quantity
7X02CTO1WW	ThinkSystem SR630 - 3yr Warranty	1
AUW3	Lenovo ThinkSystem Mainstream MB - 1U	1
AUW0	ThinkSystem MS 1U 8x2.5" Chassis	1
AWEN	Intel Xeon Gold 6130 16C 125W 2.1GHz Processor	2
AUND	ThinkSystem 32GB TruDDR4 2666 MHz (2Rx4 1.2V) RDIMM	24
AUWB	Lenovo ThinkSystem SR630 8x2.5" SATA/SAS Backplane	1
AUNL	ThinkSystem 430-8i SAS/SATA 12Gb HBA	1
AUMV	ThinkSystem M.2 with Mirroring Enablement Kit	1
AUUV	ThinkSystem M.2 CV3 128GB SATA 6Gbps Non-Hot Swap SSD	2
AUWC	ThinkSystem SR530/SR570/SR630 x8/x16 PCIe LP+LP Riser 1 Kit	1
AUWQ	Lenovo ThinkSystem 1U LP+LP BF Riser BKT	1
AUKK	ThinkSystem 10Gb 4-port SFP+ LOM	1
AUPW	ThinkSystem XClarity Controller Standard to Enterprise Upgrade	1
AXCB	ThinkSystem Toolless Slide Rail Kit with 1U CMA	1
AVW9	ThinkSystem 750W (230V) Titanium Hot-Swap Power Supply	2
6311	2.8m, 10A/100-250V, C13 to IEC 320-C14 Rack Power Cable	2
5977	Select Storage devices - no configured RAID required	1
AXFT	VMware ESXi 6.5 (factory installed)	1

**Table 42: ThinkSystem SR650 (Xeon SP Gen 1)**

Code	Description	Quantity
7X06CTO1WW	ThinkSystem SR650 - 3yr Warranty	1
AUQB	Lenovo ThinkSystem Mainstream MB - 2U	1
AUVX	ThinkSystem MS 2U 2.5" Chassis, up to 16	1
AWEN	Intel Xeon Gold 6130 16C 125W 2.1GHz Processor	2
AUND	ThinkSystem 32GB TruDDR4 2666 MHz (2Rx4 1.2V) RDIMM	24
AURA	ThinkSystem 2U/Twr 2.5" SATA/SAS 8-Bay Backplane	1
AUNL	ThinkSystem 430-8i SAS/SATA 12Gb HBA	1
AUMV	ThinkSystem M.2 with Mirroring Enablement Kit	1
AUUV	ThinkSystem M.2 CV3 128GB SATA 6Gbps Non-Hot Swap SSD	2
AUKK	ThinkSystem 10Gb 4-port SFP+ LOM	1
AUPW	ThinkSystem XClarity Controller Standard to Enterprise Upgrade	1
AXCH	ThinkSystem Toolless Slide Rail Kit with 2U CMA	1
AUS8	ThinkSystem SR550/SR590/SR650 EIA Latch w/ VGA Upgrade Kit	1
AVWF	ThinkSystem 1100W (230V/115V) Platinum Hot-Swap Power Supply	2
6311	2.8m, 10A/100-250V, C13 to IEC 320-C14 Rack Power Cable	2
5977	Select Storage devices - no configured RAID required	1
AXFT	VMware ESXi 6.5 (factory installed)	1
Select optional GPU adapters (with both risers)		
AUR3	ThinkSystem SR550/SR590/SR650 x16/x8 PCIe FH Riser 1 Kit	1
AURC	ThinkSystem SR550/SR590/SR650 (x16/x8)/(x16/x16) PCIe FH Riser 2 Kit	1
B15U	ThinkSystem NVIDIA Tesla P40 24GB PCIe Passive GPU	2
B15V	ThinkSystem NVIDIA Tesla M10 32GB PCIe Passive GPU	2

**Table 43: ThinkSystem SN550 (Xeon SP Gen 1)**

Code	Description	Quantity
7X16CTO1WW	ThinkSystem SN550 - 3yr Warranty	1
5977	Select Storage devices - no configured RAID required	1
AXHY	Intel Xeon Gold 6130 16C 125W 2.1GHz Processor	2
AUND	ThinkSystem 32GB TruDDR4 2666 MHz (2Rx4 1.2V) RDIMM	24
AUXP	Lenovo ThinkSystem SN550 Server	1
ATBS	Flex System CN4052S 4-port 10Gb Virtual Fabric Adapter	1
AUMV	ThinkSystem M.2 with Mirroring Enablement Kit	1
AUUV	ThinkSystem M.2 CV3 128GB SATA 6Gbps Non-Hot Swap SSD	1
AXFT	VMware ESXi 6.5 (factory installed)	1

**Table 44: ThinkSystem SD530 (Xeon SP Gen 1)**

Code	Description	Quantity
7X21CTO1WW	SD530 : ThinkSystem SD530 - 3yr Warranty	1
AUXN	ThinkSystem SD530 Computing Node	1
AX6D	Intel Xeon Gold 6130 16C 125W 2.1GHz Processor	2
AUNE	ThinkSystem 64GB TruDDR4 2666 MHz (4Rx4 1.2V) LRDIMM	12
AUYG	ThinkSystem SD530 3x2 SAS/SATA BP	1
B0SS	ThinkSystem 430-8i SAS/SATA 12Gb Dense HBA	1
AUMV	ThinkSystem M.2 with Mirroring Enablement Kit	1
AUPW	ThinkSystem XClarity Controller Standard to Enterprise Upgrade	1
AUUL	ThinkSystem M.2 CV1 32GB SATA 6Gbps Non-Hot Swap SSD	2
5977	Select Storage devices - no configured RAID required	1
AXFT	VMware ESXi 6.5 (factory installed)	1

Note that 64GB LRDIMMs are slower than 32GB RDIMMs and this should be taken into account when sizing a solution with the ThinkSystem SD530 that is limited to 12 DIMMs and may require the larger 64GB LRDIMMs for the same amount of system memory as the other servers that can support 24 DIMMs.

**Table 45: ThinkSystem D2 Enclosure (for SD530)**

Code	Description	Quantity
7X20CTO1WW	Chassis1 : ThinkSystem D2 Enclosure -3yr Warranty	1
AUXM	ThinkSystem D2 Enclosure	1
AUY9	ThinkSystem D2 10Gb 8 port E10M SFP+	1
AUYC	ThinkSystem D2 Slide Rail	1
AUYD	ThinkSystem D2 CMA (Cable Management Arm)	1
AUZ2	ThinkSystem D2 2000W Platinum PSU	2
6201	1.5m, 10A/100-250V, C13 to IEC 320-C14 Rack Power Cable	2
AVR1	ThinkSystem Single Ethernet Port SMM	1
AUY7	ThinkSystem D2 8-slot x8 Shuttle ASM	1

## 5.2 BOM for ThinkSystem SR635 and SR655 (AMD EPYC processor) based compute nodes

Part number	Product Description	Qty
7Y99CTO1WW	ThinkSystem-SR635 : ThinkSystem SR635 - 3yr Warranty	1
B5VL	ThinkSystem SR635 10x2.5" Chassis	1
B6TU	ThinkSystem AMD EPYC 7452 32C 155W 2.35 GHz Processor	1
B5XC	ThinkSystem 64GB TruDDR4 3200MHz (2Rx4 1.2V) RDIMM-A	8
B5VS	ThinkSystem 1U 10x2.5" PCIe Gen4 AnyBay Backplane	1
5977	Select Storage devices - no configured RAID required	1
AUNM	ThinkSystem 430-16i SAS/SATA 12Gb HBA	1

AUM1	ThinkSystem 2.5" 1.2TB 10K SAS 12Gb Hot Swap 512n HDD	4
B5XH	ThinkSystem M.2 SATA 2-Bay RAID Enablement Kit	1
B919	ThinkSystem M.2 5300 480GB SATA 6Gbps Non-Hot Swap SSD	2
B88T	VMware ESXi 6.7 U3 (factory installed)	1
B5VV	ThinkSystem SR635/SR655 x16 PCIe Riser1	1
B5W0	ThinkSystem SR635/SR655 x16/x16 PCIe Riser2 (BF)	1
B5WZ	SR635/SR655 MS FH Riser BKT	1
B715	ThinkSystem SR635 MS LP Riser BKT Slot3	1
AUKP	ThinkSystem Broadcom 57416 10GBASE-T 2-Port PCIe Ethernet Adapter	1
B6XQ	ThinkSystem 1100W (230V/115V) Platinum Hot-Swap Power Supply	2
6400	2.8m, 13A/100-250V, C13 to C14 Jumper Cord	2
AXCA	ThinkSystem Toolless Slide Rail	1
5641PX3	XClarity Pro, Per Endpoint w/3 Yr SW S&S	1
1340	Lenovo XClarity Pro, Per Managed Endpoint w/3 Yr SW S&S	1

7Z01CTO1WW	ThinkSystem-SR655 : ThinkSystem SR655 - 3yr Warranty	1
B5VJ	ThinkSystem SR655 24x2.5" Chassis	1
B6TU	ThinkSystem AMD EPYC 7452 32C 155W 2.35 GHz Processor	1
B5XC	ThinkSystem 64GB TruDDR4 3200MHz (2Rx4 1.2V) RDIMM-A	8
AURA	ThinkSystem 2U/Twr 2.5" SATA/SAS 8-Bay Backplane	3
5977	Select Storage devices - no configured RAID required	1
AUNM	ThinkSystem 430-16i SAS/SATA 12Gb HBA	1
AUNL	ThinkSystem 430-8i SAS/SATA 12Gb HBA	1
AUM1	ThinkSystem 2.5" 1.2TB 10K SAS 12Gb Hot Swap 512n HDD	4
B5XH	ThinkSystem M.2 SATA 2-Bay RAID Enablement Kit	1
B919	ThinkSystem M.2 5300 480GB SATA 6Gbps Non-Hot Swap SSD	2
B88T	VMware ESXi 6.7 U3 (factory installed)	1
B5VW	ThinkSystem SR635/SR655 x8 PCIe Internal Riser	1
B69R	ThinkSystem SR655 x16 PCIe 2U Riser 1	1

B5VY	ThinkSystem SR655 x16/x16 PCIe Riser2	1
AUKP	ThinkSystem Broadcom 57416 10GBASE-T 2-Port PCIe Ethernet Adapter	1
B6XQ	ThinkSystem 1100W (230V/115V) Platinum Hot-Swap Power Supply	2
6400	2.8m, 13A/100-250V, C13 to C14 Jumper Cord	2
5641PX3	XClarity Pro, Per Endpoint w/3 Yr SW S&S	1
1340	Lenovo XClarity Pro, Per Managed Endpoint w/3 Yr SW S&S	1

## 5.3 BOM for hyper-converged compute and Management servers

This section contains the bill of materials for hyper-converged compute servers and management servers. Table 40 and Table 41 list the BOM for a vSAN server (a minimum of four servers per cluster is needed).

**Table 46: ThinkSystem SR630 (Xeon SP Gen 1)**

Code	Description	Quantity
7X02CTO1WW	ThinkSystem SR630 - 3yr Warranty	1
AUW3	Lenovo ThinkSystem Mainstream MB - 1U	1
AUW1	ThinkSystem MS 1U 10x2.5" Chassis	1
AWEN	Intel Xeon Gold 6130 16C 125W 2.1GHz Processor	2
AUND	ThinkSystem 32GB TruDDR4 2666 MHz (2Rx4 1.2V) RDIMM	12
AUW9	Lenovo ThinkSystem SR630 2.5" Anybay 10-Bay Backplane Kit	1
AUNM	ThinkSystem 430-16i SAS/SATA 12Gb HBA	1
AUMV	ThinkSystem M.2 with Mirroring Enablement Kit	1
AUUV	ThinkSystem M.2 CV3 128GB SATA 6Gbps Non-Hot Swap SSD	2
AUWC	ThinkSystem SR530/SR570/SR630 x8/x16 PCIe LP+LP Riser 1 Kit	1
AUWQ	Lenovo ThinkSystem 1U LP+LP BF Riser BKT	1
AUKJ	ThinkSystem 10Gb 2-port SFP+ LOM	1
AUPW	ThinkSystem XClarity Controller Standard to Enterprise Upgrade	1
AXCB	ThinkSystem Toolless Slide Rail Kit with 1U CMA	1
AVW9	ThinkSystem 750W (230V) Titanium Hot-Swap Power Supply	2
6311	2.8m, 10A/100-250V, C13 to IEC 320-C14 Rack Power Cable	2
5977	Select Storage devices - no configured RAID required	1
AXFT	VMware ESXi 6.5 (factory installed)	1
Select drive configuration for hyper-converged system (all flash or SDD/HDD combination)		
AUMG	ThinkSystem 2.5" HUSMM32 400GB Performance SAS 12Gb Hot Swap SSD	2
AUMH	ThinkSystem 2.5" HUSMM32 800GB Performance SAS 12Gb Hot Swap SSD	2
AUM1	ThinkSystem 2.5" 1.2TB 10K SAS 12Gb Hot Swap 512n HDD	8
AUMK	ThinkSystem 2.5" PM1633a 3.84TB Capacity SAS 12Gb Hot Swap SSD	6

**Table 47: ThinkSystem SR650 (Xeon SP Gen 1)**

Code	Description	Quantity
7X06CTO1WW	ThinkSystem SR650 - 3yr Warranty	1
AUQB	Lenovo ThinkSystem Mainstream MB - 2U	1
AUVV	ThinkSystem MS 2U 24x2.5" Chassis	1
AWEN	Intel Xeon Gold 6130 16C 125W 2.1GHz Processor	2
AUND	ThinkSystem 32GB TruDDR4 2666 MHz (2Rx4 1.2V) RDIMM	12
AURA	ThinkSystem 2U/Twr 2.5" SATA/SAS 8-Bay Backplane	2
AUNL	ThinkSystem 430-8i SAS/SATA 12Gb HBA	2
AUMV	ThinkSystem M.2 with Mirroring Enablement Kit	1
AUUV	ThinkSystem M.2 CV3 128GB SATA 6Gbps Non-Hot Swap SSD	2
AUKJ	ThinkSystem 10Gb 2-port SFP+ LOM	1
AUPW	ThinkSystem XClarity Controller Standard to Enterprise Upgrade	1
AXCH	ThinkSystem Toolless Slide Rail Kit with 2U CMA	1
AUS8	ThinkSystem SR550/SR590/SR650 EIA Latch w/ VGA Upgrade Kit	1
AVWF	ThinkSystem 1100W (230V/115V) Platinum Hot-Swap Power Supply	2
6311	2.8m, 10A/100-250V, C13 to IEC 320-C14 Rack Power Cable	2
5977	Select Storage devices - no configured RAID required	1
AXFT	VMware ESXi 6.5 (factory installed)	1
Select drive configuration for hyper-converged system (all flash or SDD/HDD combination)		
AUMG	ThinkSystem 2.5" HUSMM32 400GB Performance SAS 12Gb Hot Swap SSD	2
AUMH	ThinkSystem 2.5" HUSMM32 800GB Performance SAS 12Gb Hot Swap SSD	2
AUM1	ThinkSystem 2.5" 1.2TB 10K SAS 12Gb Hot Swap 512n HDD	8
AUMK	ThinkSystem 2.5" PM1633a 3.84TB Capacity SAS 12Gb Hot Swap SSD	6
Select optional GPU adapters (with both risers)		
AUR3	ThinkSystem SR550/SR590/SR650 x16/x8 PCIe FH Riser 1 Kit	1
AURC	ThinkSystem SR550/SR590/SR650 (x16/x8)/(x16/x16) PCIe FH Riser 2 Kit	1
B15U	ThinkSystem NVIDIA Tesla P40 24GB PCIe Passive GPU	2
B15V	ThinkSystem NVIDIA Tesla M10 32GB PCIe Passive GPU	2

**Table 48: ThinkSystem SD530 (Xeon SP Gen 1)**

Code	Description	Quantity
7X21CTO1WW	SD530 : ThinkSystem SD530 - 3yr Warranty	1
AUXN	ThinkSystem SD530 Computing Node	1
AX6D	Intel Xeon Gold 6130 16C 125W 2.1GHz Processor	2
AUNE	ThinkSystem 64GB TruDDR4 2666 MHz (4Rx4 1.2V) LRDIMM	12
AUYG	ThinkSystem SD530 3x2 SAS/SATA BP	1
B0SS	ThinkSystem 430-8i SAS/SATA 12Gb Dense HBA	1
AUMV	ThinkSystem M.2 with Mirroring Enablement Kit	1
AUPW	ThinkSystem XClarity Controller Standard to Enterprise Upgrade	1
AUUL	ThinkSystem M.2 CV1 32GB SATA 6Gbps Non-Hot Swap SSD	2
5977	Select Storage devices - no configured RAID required	1
AXFT	VMware ESXi 6.5 (factory installed)	1
Select drive configuration for hyper-converged system (all flash or SDD/HDD combination)		
AUMG	ThinkSystem 2.5" HUSMM32 400GB Performance SAS 12Gb Hot Swap SSD	2
AUMH	ThinkSystem 2.5" HUSMM32 800GB Performance SAS 12Gb Hot Swap SSD	2
AUM1	ThinkSystem 2.5" 1.2TB 10K SAS 12Gb Hot Swap 512n HDD	4
AUMK	ThinkSystem 2.5" PM1633a 3.84TB Capacity SAS 12Gb Hot Swap SSD	4

Note that 64GB LRDIMMs are slower than 32GB RDIMMs and this should be taken into account when sizing a solution with the ThinkSystem SD530 that is limited to 12 DIMMs and may require the larger 64GB LRDIMMs for the same amount of system memory as the other servers that can support 24 DIMMs.

**Table 49: ThinkSystem D2 Enclosure (for SD530)**

Code	Description	Quantity
7X20CTO1WW	Chassis1 : ThinkSystem D2 Enclosure -3yr Warranty	1
AUXM	ThinkSystem D2 Enclosure	1
AUY9	ThinkSystem D2 10Gb 8 port E10M SFP+	1
AUYC	ThinkSystem D2 Slide Rail	1
AUYD	ThinkSystem D2 CMA (Cable Management Arm)	1
AUZ2	ThinkSystem D2 2000W Platinum PSU	2
6201	1.5m, 10A/100-250V, C13 to IEC 320-C14 Rack Power Cable	2
AVR1	ThinkSystem Single Ethernet Port SMM	1
AUY7	ThinkSystem D2 8-slot x8 Shuttle ASM	1



**Table 50: ThinkAgile VX3320 Appliance (Xeon SP Gen 2)**

Part number	Product Description	Qty
7Y93CTO1WW	Server : Lenovo ThinkAgile VX3320 Appliance	1
B1DK	ThinkAgile VX 1U 2.5" 10 Bay Chassis	1
B4HJ	Intel Xeon Gold 6230 20C 125W 2.1GHz Processor	2
AUND	ThinkSystem 32GB TruDDR4 2666 MHz (2Rx4 1.2V) RDIMM	12
AUKJ	ThinkSystem 10Gb 2-port SFP+ LOM	1
AVWB	ThinkSystem 1100W (230V/115V) Platinum Hot-Swap Power Supply	2
6400	2.8m, 13A/100-250V, C13 to C14 Jumper Cord	2
B88T	VMware ESXi 6.7 U3 (factory installed)	1
AUWC	ThinkSystem SR530/SR570/SR630 x8/x16 PCIe LP+LP Riser 1 Kit	1
AUWQ	Lenovo ThinkSystem 1U LP+LP BF Riser Bracket	1
AXCA	ThinkSystem Toolless Slide Rail	1
AUW9	ThinkSystem 1U 2.5" 4 AnyBay 10-Bay Backplane	1
5977	Select Storage devices - no configured RAID required	1
AUNM	ThinkSystem 430-16i SAS/SATA 12Gb HBA	1
AUMV	ThinkSystem M.2 with Mirroring Enablement Kit	1
AUUV	ThinkSystem M.2 128GB SATA 6Gbps Non-Hot Swap SSD	2
B7Y0	Enable IPMI-over-LAN	1
AVKG	ThinkSystem SR630 MB to 10x2.5" HDD BP NVME cable	1
AUWV	10x2.5"Cable Kit (1U)	1
Select drive configuration for hyper-converged system (all flash or SDD/HDD combination)		
B5MD	vSAN Hybrid Config	1
B4Y5	ThinkSystem 2.5" SS530 800GB Performance SAS 12Gb Hot Swap SSD	2
AUM2	ThinkSystem 2.5" 1.8TB 10K SAS 12Gb Hot Swap 512e HDD	6
B5MC	vSAN All Flash Config	1
B4Y7	ThinkSystem 2.5" SS530 3.2TB Performance SAS 12Gb Hot Swap SSD	6

**Table 51: ThinkAgile VX 1U Certified Node (Xeon SP Gen 2)**

Part number	Product Description	Qty
7Y93CTO4WW	Server : Lenovo ThinkAgile VX 1U Certified Node	1
B1DK	ThinkAgile VX 1U 2.5" 10 Bay Chassis	1
B4HJ	Intel Xeon Gold 6230 20C 125W 2.1GHz Processor	2
AUND	ThinkSystem 32GB TruDDR4 2666 MHz (2Rx4 1.2V) RDIMM	12
AUW9	ThinkSystem 1U 2.5" 4 AnyBay 10-Bay Backplane	1
5977	Select Storage devices - no configured RAID required	1
AUNM	ThinkSystem 430-16i SAS/SATA 12Gb HBA	1
AUMV	ThinkSystem M.2 with Mirroring Enablement Kit	1
AUUV	ThinkSystem M.2 128GB SATA 6Gbps Non-Hot Swap SSD	2
B88T	VMware ESXi 6.7 U3 (factory installed)	1
AUKJ	ThinkSystem 10Gb 2-port SFP+ LOM	1
AVWB	ThinkSystem 1100W (230V/115V) Platinum Hot-Swap Power Supply	2
6400	2.8m, 13A/100-250V, C13 to C14 Jumper Cord	2
Select drive configuration for hyper-converged system (all flash or SDD/HDD combination)		
B5MD	vSAN Hybrid Config	1
B4Y5	ThinkSystem 2.5" SS530 800GB Performance SAS 12Gb Hot Swap SSD	2
AUM2	ThinkSystem 2.5" 1.8TB 10K SAS 12Gb Hot Swap 512e HDD	6
B5MC	vSAN All Flash Config	1
B4Y7	ThinkSystem 2.5" SS530 3.2TB Performance SAS 12Gb Hot Swap SSD	6

**Table 52: ThinkAgile VX 2U Certified Node (Xeon SP Gen 2)**

Part number	Product Description	Qty
7Y94CTO5WW	Server : ThinkAgile VX 2U Node	1
B1DH	ThinkAgile VX 2U 2.5" 8/16/24 Bay Chassis	1
B4HJ	Intel Xeon Gold 6230 20C 125W 2.1GHz Processor	2
AUND	ThinkSystem 32GB TruDDR4 2666 MHz (2Rx4 1.2V) RDIMM	12
AUR5	ThinkSystem 2U/Twr 2.5" AnyBay 8-Bay Backplane	1
AURA	ThinkSystem 2U/Twr 2.5" SATA/SAS 8-Bay Backplane	2
5977	Select Storage devices - no configured RAID required	1
AUNM	ThinkSystem 430-16i SAS/SATA 12Gb HBA	1
AUMV	ThinkSystem M.2 with Mirroring Enablement Kit	1
AUUV	ThinkSystem M.2 128GB SATA 6Gbps Non-Hot Swap SSD	2
B88T	VMware ESXi 6.7 U3 (factory installed)	1
AURC	ThinkSystem SR550/SR590/SR650 (x16/x8)/(x16/x16) PCIe FH Riser 2 Kit	1
AUKK	ThinkSystem 10Gb 4-port SFP+ LOM	1
AVWF	ThinkSystem 1100W (230V/115V) Platinum Hot-Swap Power Supply	2
6400	2.8m, 13A/100-250V, C13 to C14 Jumper Cord	2
AUSM	MS 2nd 2U 8X2.5" Cable Kit	2
AUSQ	On Board to 2U 8x2.5" HDD BP NVME Cable	1
Select drive configuration for hyper-converged system (all flash or SDD/HDD combination)		
B5MD	vSAN Hybrid Config	1
B4Y5	ThinkSystem 2.5" SS530 800GB Performance SAS 12Gb Hot Swap SSD	2
AUM2	ThinkSystem 2.5" 1.8TB 10K SAS 12Gb Hot Swap 512e HDD	6
B5MC	vSAN All Flash Config	1
B4Y7	ThinkSystem 2.5" SS530 3.2TB Performance SAS 12Gb Hot Swap SSD	6

## 5.4 BOM for enterprise and SMB management servers

Table 28 on page 30 lists the number of management servers that are needed for the different numbers of users. To help with redundancy, the bill of materials for management servers must be the same as compute servers. For more information, see “BOM for enterprise and SMB compute servers” on page 38.

Because the Windows storage servers use a bare-metal operating system (OS) installation, they require much less memory and can have a reduced configuration as listed below.

**Table 53: ThinkSystem SR630 (Xeon SP Gen 1)**

Code	Description	Quantity
7X02CTO1WW	ThinkSystem SR630 - 3yr Warranty	1
AUW3	Lenovo ThinkSystem Mainstream MB - 1U	1
AUW0	ThinkSystem MS 1U 8x2.5" Chassis	1
AWEN	Intel Xeon Gold 6130 16C 125W 2.1GHz Processor	2
AUU1	ThinkSystem 8GB TruDDR4 2666 MHz (1Rx8 1.2V) RDIMM	12
AUWB	Lenovo ThinkSystem SR630 8x2.5" SATA/SAS Backplane	1
AUNL	ThinkSystem 430-8i SAS/SATA 12Gb HBA	1
AUMV	ThinkSystem M.2 with Mirroring Enablement Kit	1
AUUV	ThinkSystem M.2 CV3 128GB SATA 6Gbps Non-Hot Swap SSD	2
AUWC	ThinkSystem SR530/SR570/SR630 x8/x16 PCIe LP+LP Riser 1 Kit	1
AUWQ	Lenovo ThinkSystem 1U LP+LP BF Riser BKT	1
AUKK	ThinkSystem 10Gb 4-port SFP+ LOM	1
AUPW	ThinkSystem XClarity Controller Standard to Enterprise Upgrade	1
AXCB	ThinkSystem Toolless Slide Rail Kit with 1U CMA	1
AVW9	ThinkSystem 750W (230V) Titanium Hot-Swap Power Supply	2
6311	2.8m, 10A/100-250V, C13 to IEC 320-C14 Rack Power Cable	2
5977	Select Storage devices - no configured RAID required	1

**Table 54: ThinkSystem SR650 (Xeon SP Gen 1)**

Code	Description	Quantity
7X06CTO1WW	ThinkSystem SR650 - 3yr Warranty	1
AUQB	Lenovo ThinkSystem Mainstream MB - 2U	1
AUVX	ThinkSystem MS 2U 2.5" Chassis, up to 16	1
AWEN	Intel Xeon Gold 6130 16C 125W 2.1GHz Processor	2
AUU1	ThinkSystem 8GB TruDDR4 2666 MHz (1Rx8 1.2V) RDIMM	12
AURA	ThinkSystem 2U/Twr 2.5" SATA/SAS 8-Bay Backplane	1
AUNL	ThinkSystem 430-8i SAS/SATA 12Gb HBA	1
AUMV	ThinkSystem M.2 with Mirroring Enablement Kit	1
AUUV	ThinkSystem M.2 CV3 128GB SATA 6Gbps Non-Hot Swap SSD	2
AUKK	ThinkSystem 10Gb 4-port SFP+ LOM	1
AUPW	ThinkSystem XClarity Controller Standard to Enterprise Upgrade	1
AXCH	ThinkSystem Toolless Slide Rail Kit with 2U CMA	1
AUS8	ThinkSystem SR550/SR590/SR650 EIA Latch w/ VGA Upgrade Kit	1
AVWF	ThinkSystem 1100W (230V/115V) Platinum Hot-Swap Power Supply	2
6311	2.8m, 10A/100-250V, C13 to IEC 320-C14 Rack Power Cable	2
5977	Select Storage devices - no configured RAID required	1

**Table 55: ThinkSystem SN550 (Xeon SP Gen 1)**

Code	Description	Quantity
7X16CTO1WW	ThinkSystem SN550 - 3yr Warranty	1
5977	Select Storage devices - no configured RAID required	1
AXHY	Intel Xeon Gold 6130 16C 125W 2.1GHz Processor	2
AUU1	ThinkSystem 8GB TruDDR4 2666 MHz (1Rx8 1.2V) RDIMM	12
AUXP	Lenovo ThinkSystem SN550 Server	1
ATBS	Flex System CN4052S 4-port 10Gb Virtual Fabric Adapter	1
AUMV	ThinkSystem M.2 with Mirroring Enablement Kit	1
AUUV	ThinkSystem M.2 CV3 128GB SATA 6Gbps Non-Hot Swap SSD	1

**Table 56: ThinkSystem SD530 (Xeon SP Gen 1)**

Code	Description	Quantity
7X21CTO1WW	SD530 : ThinkSystem SD530 - 3yr Warranty	1
AUXN	ThinkSystem SD530 Computing Node	1
AX6D	Intel Xeon Gold 6130 16C 125W 2.1GHz Processor	2
AUU1	ThinkSystem 8GB TruDDR4 2666 MHz (1Rx8 1.2V) RDIMM	12
AUYG	ThinkSystem SD530 3x2 SAS/SATA BP	1
B0SS	ThinkSystem 430-8i SAS/SATA 12Gb Dense HBA	1
AUMV	ThinkSystem M.2 with Mirroring Enablement Kit	1
AUPW	ThinkSystem XClarity Controller Standard to Enterprise Upgrade	1
AUUL	ThinkSystem M.2 CV1 32GB SATA 6Gbps Non-Hot Swap SSD	2
5977	Select Storage devices - no configured RAID required	1
AXFT	VMware ESXi 6.5 (factory installed)	1

## 5.5 BOM for shared storage

The following three tables list the BOMs for Lenovo ThinkSystem DS6200 storage array.

**Table 57: Lenovo ThinkSystem DS6200 control enclosure**

Code	Description	Quantity
4619HC2	Lenovo ThinkSystem DS6200 SFF Chassis	1
AU32	Lenovo ThinkSystem DS6200 SFF Chassis	1
AUE6	DS6200 Product Label	1
AT1F	Lenovo USB A Male-to-Mini-B 1.5m cable	1
AU0W	12Gb SAN Rack Mount Kit-Rails 25"-36"	1
AU34	DS6200 FC/iSCSI Controller Module	2
AXD2	DS Series ShipKit	1
6201	1.5m, 10A/100-250V, C13 to IEC 320-C14 Rack Power Cable	2
A1PJ	3m Passive DAC SFP+ Cable	4
Select drive configuration		
AU1U	Lenovo Storage 800GB 3DWD 2.5" SAS SSD	24
AVPA	Lenovo Storage 3.84TB 1DWD 2.5" SAS SSD (PM1633a)	24

**Table 58: Lenovo ThinkSystem DS6200 expansion enclosure**

Code	Description	Quantity
4588HC2	Lenovo ThinkSystem DS Series SFF Exp	1
AU26	Lenovo ThinkSystem DS Series SFF Exp	1
AU2K	DS Series 12G SAN Expansion Module	2
6201	1.5m, 10A/100-250V, C13 to IEC 320-C14 Rack Power Cable	2
AXD2	DS Series ShipKit	1
AU16	External MiniSAS HD 8644/MiniSAS HD 8644 0.5M Cable	2
AU0W	12Gb SAN Rack Mount Kit-Rails 25"-36"	1
AUE5	EXP V2 Product Labeling	1
Select drive configuration		
AU1U	Lenovo Storage 800GB 3DWD 2.5" SAS SSD	24
AVPA	Lenovo Storage 3.84TB 1DWD 2.5" SAS SSD (PM1633a)	24

**Table 59: Lenovo ThinkSystem DS6200 optional software licenses**

Code	Description
01GV559	512 Snapshot Upgrade License
01GV560	1024 Snapshot Upgrade License
01GV561	SSD Data Tiering License (only needed for hybrid SSD/HDD arrays)
01GV562	Asynchronous Replication License

## 5.6 BOM for networking

For more information about the number and type of TOR network switches that are needed for different user counts, see the “Networking” section on page 35.

### RackSwitch G8052

Code	Description	Quantity
7159G52	Lenovo System Networking RackSwitch G8052 (Rear to Front)	1
6201	1.5m, 10A/100-250V, C13 to IEC 320-C14 Rack Power Cable	2
3802	1.5m Blue Cat5e Cable	3
A3KP	Lenovo System Networking Adjustable 19" 4 Post Rail Kit	1

### ThinkSystem NE1032

Code	Description	Quantity
7159A1X	Lenovo ThinkSystem NE1032 RackSwitch (Rear to Front)(Standard)	2
39Y7938	2.8m, 10A/100-250V, C13 to IEC 320-C20 Rack Power Cable	4

### RackSwitch G8272

Code	Description	Quantity
7159CRW	Lenovo System Networking RackSwitch G8272 (Rear to Front)	1
6201	1.5m, 10A/100-250V, C13 to IEC 320-C14 Rack Power Cable	2
A3KP	Lenovo System Networking Adjustable 19" 4 Post Rail Kit	1
5053	SFP+ SR Transceiver	2
A1DP	1m QSFP+-to-QSFP+ cable	1
A1DM	3m QSFP+ DAC Break Out Cable	0

## 5.7 BOM for Flex System chassis

Table 54 lists the BOM for a Lenovo Flex System chassis.

**Table 60: Chassis BOM**

Code	Description	Quantity
8721HC2	Lenovo Flex System Enterprise Chassis w/ CMM2	1
ASUS	Lenovo Flex System Enterprise Chassis w/ CMM2	1
ASUV	Lenovo Flex System Fabric SI4093 System Interconnect Module	2
6292	2m, 16A/100-250V, C19 to IEC 320-C20 Rack Power Cable	6
A0UC	Flex System Enterprise Chassis 2500W Power Module Standard	2
A0UD	Flex System Enterprise Chassis 2500W Power Module	4
A0UA	Flex System Enterprise Chassis 80mm Fan Module	4

## 5.8 BOM for rack

Table 55 lists the BOM for a rack.

**Table 61: Rack BOM**

Code	Description	Quantity
93634PX	42U 1100mm Enterprise V2 Dynamic Rack	1
39Y8941	DPI Single Phase C19 Enterprise PDU (without line cord)	6
40K9614	DPI 30a Line Cord (NEMA L6-30P)	6



# Resources

---

For more information, see the following resources:

- Lenovo Client Virtualization reference architecture  
[lenovopress.com/lp0756](http://lenovopress.com/lp0756)
- Citrix Virtual Apps and Desktops  
[citrix.com/products/Virtual Apps and Desktops](http://citrix.com/products/Virtual Apps and Desktops)
- VMware vSphere  
[vmware.com/products/datacenter-virtualization/vsphere](http://vmware.com/products/datacenter-virtualization/vsphere)
- VMware vSAN  
[vmware.com/products/vsan.html](http://vmware.com/products/vsan.html)
- VMware vSAN 6.6 design and sizing guide  
[storagehub.vmware.com/#!/vmware-vsan/vmware-r-vsan-tm-design-and-sizing-guide](http://storagehub.vmware.com/#!/vmware-vsan/vmware-r-vsan-tm-design-and-sizing-guide)
- Flex System Interoperability Guide  
[lenovopress.com/redpfsig-flex-system-interoperability-guide](http://lenovopress.com/redpfsig-flex-system-interoperability-guide)
- Lenovo Storage Interoperability Links  
[lenovopress.com/lp0584-lenovo-storage-interoperability-links](http://lenovopress.com/lp0584-lenovo-storage-interoperability-links)

# Document History

---

16 October 2017	<ul style="list-style-type: none"><li>• Initial version</li></ul>
8 January 2018	<ul style="list-style-type: none"><li>• Added section on GPU graphics acceleration</li><li>• Replaced G8124E RackSwitch with ThinkSystem NE1032 switch</li><li>• Added ThinkSystem SD530 server</li></ul>
22 January 2018	<ul style="list-style-type: none"><li>• Added XenServer hypervisor</li></ul>
21 June 2018	<ul style="list-style-type: none"><li>• Added NVidia M10 and P40 GPUs</li></ul>
5 November 2018	<ul style="list-style-type: none"><li>• Added results for ESXi 6.7 using vSAN 6.7</li></ul>
2 December 2019	<ul style="list-style-type: none"><li>• Added results for ThinkAgile VX Appliances with Xeon SP Gen 2 processors using ESXi 6.7U3</li><li>• Changed terminology Citrix XenDesktop to Citrix Virtual Apps and Desktops; and Citrix XenServer to Citrix Hypervisor</li></ul>
6 February 2020	<ul style="list-style-type: none"><li>• Added results for ThinkServer SR635 and SR655 AMD EPYC Processor Gen 2 (Rome) based servers using ESXi 6.7U3</li><li>• Added BOMs for the SR635 and SR655 servers</li></ul>

# Trademarks and special notices

---

© Copyright Lenovo 2020.

References in this document to Lenovo products or services do not imply that Lenovo intends to make them available in every country.

Lenovo, the Lenovo logo, ThinkSystem, ThinkCentre, ThinkVision, ThinkVantage, ThinkPlus and Rescue and Recovery are trademarks of Lenovo.

IBM, the IBM logo, and ibm.com are trademarks or registered trademarks of International Business Machines Corporation in the United States, other countries, or both.

Microsoft, Windows, Windows NT, and the Windows logo are trademarks of Microsoft Corporation in the United States, other countries, or both.

Intel, Intel Inside (logos), MMX, and Pentium are trademarks of Intel Corporation in the United States, other countries, or both.

Other company, product, or service names may be trademarks or service marks of others.

Information is provided "AS IS" without warranty of any kind.

All customer examples described are presented as illustrations of how those customers have used Lenovo products and the results they may have achieved. Actual environmental costs and performance characteristics may vary by customer.

Information concerning non-Lenovo products was obtained from a supplier of these products, published announcement material, or other publicly available sources and does not constitute an endorsement of such products by Lenovo. Sources for non-Lenovo list prices and performance numbers are taken from publicly available information, including vendor announcements and vendor worldwide homepages. Lenovo has not tested these products and cannot confirm the accuracy of performance, capability, or any other claims related to non-Lenovo products. Questions on the capability of non-Lenovo products should be addressed to the supplier of those products.

All statements regarding Lenovo future direction and intent are subject to change or withdrawal without notice, and represent goals and objectives only. Contact your local Lenovo office or Lenovo authorized reseller for the full text of the specific Statement of Direction.

Some information addresses anticipated future capabilities. Such information is not intended as a definitive statement of a commitment to specific levels of performance, function or delivery schedules with respect to any future products. Such commitments are only made in Lenovo product announcements. The information is presented here to communicate Lenovo's current investment and development activities as a good faith effort to help with our customers' future planning.

Performance is based on measurements and projections using standard Lenovo benchmarks in a controlled environment. The actual throughput or performance that any user will experience will vary depending upon considerations such as the amount of multiprogramming in the user's job stream, the I/O configuration, the storage configuration, and the workload processed. Therefore, no assurance can be given that an individual user will achieve throughput or performance improvements equivalent to the ratios stated here.

Photographs shown are of engineering prototypes. Changes may be incorporated in production models.

Any references in this information to non-Lenovo websites are provided for convenience only and do not in any manner serve as an endorsement of those websites. The materials at those websites are not part of the materials for this Lenovo product and use of those websites is at your own risk.