VMware View® with FlexPod®

January 2013

DEPLOYMENT AND TECHNICAL CONSIDERATIONS GUIDE
Contents

Introduction ................................................................................................................................. 3

Infrastructure Components ........................................................................................................ 4
  VMware View .......................................................................................................................... 4
  FlexPod Architecture .............................................................................................................. 11
  Cisco Unified Computing System .......................................................................................... 12
  NetApp FAS Storage Systems ............................................................................................... 13

Planning and Design Guidance .................................................................................................. 14
  Types of Workers .................................................................................................................... 14
  VMware View ESX/ESXi Node ............................................................................................... 19
  Desktop Virtual Machine Configuration ............................................................................... 20
  vCenter and View Composer Virtual Machine Configuration ............................................ 21
  View Connection Server and Virtual Machine Configuration ............................................ 22
  View Transfer Server Virtual Machine Configuration and Storage .................................... 22
  vSphere Clusters ................................................................................................................... 23
  Network Bandwidth Considerations ...................................................................................... 24

Cisco Best Practices .................................................................................................................. 26
  UCS Manager with UCS Service Profiles ............................................................................. 26
  VLAN Configuration .............................................................................................................. 27
  QoS Implementation .............................................................................................................. 27

NetApp Best Practices .............................................................................................................. 28
  Deduplication ......................................................................................................................... 28
  Space Reclamation ................................................................................................................ 29
  Read and Write I/O Optimization ......................................................................................... 29
  Virtual Storage Tiering .......................................................................................................... 29

Conclusion .................................................................................................................................. 30

Resources .................................................................................................................................... 31

Acknowledgements .................................................................................................................. 34
Introduction

Today, many CIOs are challenged to deliver a solution that efficiently and securely delivers Windows desktops and applications to all types of end-workers and devices in a way that meets corporate objectives for cost, security, and compliance. In order to answer this challenge, VMware® View® with FlexPod™ provides the foundation for delivering efficient and scalable virtual desktop infrastructure using a pre-validated stack from industry-leading partners including VMware®, Cisco®, and NetApp®.

The VMware View with Flexpod architecture reduces processes, thereby minimizing operational challenges such as configuration errors, and it lowers administrative costs. Using Flexpod to scale VMware View desktops on demand provides increased service levels and resource optimization. This solution enables enterprises to rapidly deploy a pre-tested and validated virtual desktop infrastructure (VDI). It provides proactive, predictive, and centralized management that can be integrated with third-party tools.

This paper describes the Flexpod reference architecture for VMware View 5.1 that include the Cisco Unified Computing System™ (Cisco UCS) and the NetApp® Fabric-Attached Storage (FAS) Systems. It also provides detailed guidance on how to architect, implement, and manage a large, scalable VMware View solution with FlexPod. The paper details the best integration points for each of the key enabling Cisco and NetApp technologies and how each of these technologies plays a critical and complementary role in providing an integrated solution for VMware View deployments.

This paper is intended for IT decision makers, architects, administrators, and others who want to familiarize themselves with the components and capabilities of VMware View on FlexPod. This paper assumes that the reader has a general understanding of VMware View and the VMware vSphere platform, and the Flexpod, Cisco, and Netapp products described in this solution.
Infrastructure Components

The VMware View with Flexpod infrastructure components include:

- VMware View
- FlexPod Architecture
- Cisco Unified Computing System
- NetApp FAS Storage Systems

Each of these components is described in the sections below.

VMware View

VMware View desktop virtualization simplifies IT manageability and control, while delivering a highly dependable end user experience across devices and networks. VMware View helps IT organizations automate desktop and application management, reduce costs, and increase data security through the centralization of the desktop environment.

Virtual desktop centralization is enabled by using VMware vSphere® and server, storage, and networking resource virtualization. This centralization results in greater end user freedom and increased control for IT organizations.

By using VMware View, operating systems, applications, and user data can be encapsulated into isolated layers, allowing IT organizations to deliver a modern desktop. This solution can also deliver dynamic and elastic desktop cloud services such as applications, unified communications, and 3D graphics for real-world productivity and greater business agility. The benefits of managing enterprise desktops with VMware View include increased reliability, security, hardware independence, and convenience.

- **Reliability and security**: VMware View provides for reliability and security:
  - Access to data can be easily restricted. Sensitive data can be prevented from being copied onto a remote employee’s home computer.
  - Data backups can be scheduled anytime without considering when workers’ systems might be turned off.
  - Virtual desktops that are hosted in the datacenter experience little or no downtime. Virtual machines can reside on high-availability VMware server clusters.

Virtual desktops can also be connected to back-end physical systems and Windows® Terminal Services servers.

- **Convenience**: The unified management console is built for scalability on Adobe® Flex®, allowing even the largest VMware View deployments to be efficiently managed from a single VMware® View™ Manager interface. Wizards and dashboards enhance the workflow and facilitate drilling down to see details or change settings.
Manageability: Provisioning desktops for end users is a quick process. Applications are not installed one by one on each end user’s physical PC. Workers connect to a virtual desktop that is already complete with applications. They can access the same virtual desktop from various devices at various locations.

Using VMware vSphere to host virtual desktops provides these benefits:

- Administration tasks and management chores are reduced. Administrators can patch and upgrade applications and operating systems without direct intervention.
- Storage management is simplified. VMware vSphere supports volume and file system virtualization that avoids managing separate storage devices.

Hardware Independence: Virtual machines are hardware independent. A VMware View desktop runs in the datacenter and is only accessed from a client device. This means, a VMware View desktop can use operating systems that are incompatible with the client device hardware. For example, although Windows® 7 can run only on Windows 7-enabled PCs, Windows 7 can be installed in a virtual machine that runs on a PC that is not Windows 7 enabled. Virtual desktops can run on a PC, Mac®, tablet, thin client, and on a PC that has been repurposed as a thin client.

VMware View offers the following security features, among others:

- Uses RSA SecurID two-factor authentication or smart cards for log on.
- Uses Secure Sockets Layer (SSL) tunneling to ensure that all connections are completely encrypted.
- Uses VMware High Availability (HA) to host desktops and ensure automatic failover.

The centralized administration and management features include:

- Uses Microsoft® Active Directory® to manage access to virtual desktops and policies.
- Uses the Web-based administrative console to manage virtual desktops from any location.
- Uses a template or master image to quickly create and provision desktop pools.
- Sends updates and patches to virtual desktops without affecting user settings, data, or preferences.

Scalability features depend on the VMware virtualization platform to manage both desktops and servers including:

- Integrates with VMware vSphere to achieve cost-effective densities, high levels of availability, and advanced resource allocation control for virtual desktops.
- Configures View Connection Server to broker connections between workers and the virtual desktops each is authorized to access.
- Uses VMware® View™ Composer to quickly create desktop images that share virtual disks with a master image. Using linked-clones conserves disk space and simplifies the management of operating system patches and updates.
VMware View 5.1 Features

VMware View 5.1 features enhance the deployment of FlexPod in terms of usability, security, centralized control, and scalability. VMware View 5.1 is built upon the advancements in VMware View® 5. These include features that enhance the end user experience and improve IT control. Some of the more notable features include:

- **View Media Services for 3D Graphics**: Enable VMware View desktops to run basic 3D applications such as Aero, Microsoft® Office® 2010, or those requiring OpenGL or DirectX®, without specialized graphics cards or client devices.

- **View Media Services for Integrated Unified Communications**: Integrate Voice over IP (VoIP) and the VMware View desktop for the end user through an architecture that optimizes performance for both the desktop and unified communications.

- **View Client for Android**: Enables workers with Google® Android™ based tablets to access VMware View virtual desktops. Support for VMware vSphere® 5 leverages the latest functionality of the leading cloud infrastructure platform to provide highly available, scalable, and reliable desktop services. VMware View for iOS® was released with VMware View® 4.5 and is available today for VMware View 5 and later versions.

View Storage Accelerator

View Storage Accelerator reduces the total cost of ownership (TCO) since IT administrators can size the shared FlexPod array for steady-state workloads, rather than for peak workloads. It reduces the cost of the VDI deployment by decreasing the amount of storage that customers need to address peak workloads. View Storage Accelerator leverages a VMware vSphere (version 5.0 or later) platform feature called Content Based Read Cache (CBRC) implemented in the VMware® ESX® and VMware® ESXi™ hypervisor.

When enabled for specific virtual machines, the host hypervisor scans the storage disk blocks to generate digests of block contents. When these blocks are read into the hypervisor, they are cached in the host-based CBRC. Subsequent reads of blocks with the same digest are served from the in-memory cache directly. This significantly improves desktop performance, especially during boot storms or antivirus scanning storms, when a large number of blocks with identical contents are read.

Simplified Desktop Management

Simplified Desktop Management preserves user profiles and dynamically synchronizes them with a remote profile repository. By preserving the user profile between sessions, View Persona Management enables IT administrators to provide a floating pool to workers who require a dedicated desktop, enabling Windows® XP to Windows® 7 migrations. The View Persona Management agent can be installed without the VMware View agent on physical desktops belonging to the same licensed VMware View desktop workers.

During a physical-to-virtual migration, an IT administrator can first install View Persona Management on the physical desktop. When the same user accesses a virtual desktop with View Persona Management enabled, user data and user settings are automatically synchronized.
WAN using PCoIP

VMware improved the performance of the PC-over-IP (PCoIP) protocol over higher latency and lossy networks by up to eight times to provide a better user experience. VMware View 5.1 provides better performance over wide area network (WAN) links with latency up to 300 ms round-trip time (RTT). Additionally, performance is more robust with the ability to gracefully tolerate packet loss, and curtail bandwidth at higher loss rates to avoid congestion problems.

For more information about the performance benefits of running PCoIP, go to Performance Does Matter: PCoIP vs. HDX at: http://blogs.vmware.com/euc/2012/05/performance-does-matter-pcoip-vs-hdx.html

For more information about VMware View details and features, go to the VMware View 5.1 Release Notes at: http://www.vmware.com/support/view51/doc/view-51-release-notes.html

VMware View Architecture

View Client enables users to log on to View Connection Server, as shown below. View Connection Server integrates with Windows® Active Directory® and provides access to a virtual desktop hosted in a VMware vSphere environment. View Client and View Connection Server are described in more detail in the sections below.

Figure 1: Major components of a VMware View deployment

Client Devices

VMware View desktops follow the end user regardless of device or location. Workers can access their personalized virtual desktop from a company laptop, home PC, thin client device, Mac, iPad®, or other tablet. This level of flexibility supports Bring Your Own Device (BYOD) functionality for the enterprise.
From tablets, Mac, Linux®, Windows® laptops, and PCs, workers can open View Client to display their VMware View desktop. Thin client devices can be configured to ensure that workers can only launch View Thin Client directly in the device. Repurposing a legacy PC into a thin client desktop can extend the life of the hardware by three to five years. For example, by using VMware View on a thin desktop, a new operating system such as Windows 7 can be used on older desktop hardware.

**View Connection Server**

View Connection Server acts as a broker for client connections. View Connection Server authenticates workers through Windows Active Directory and directs the request to the appropriate virtual machine, physical or blade PC, or Windows Terminal Services server. View Connection Server provides management capabilities including:

- Authenticates workers and entitles workers to specific desktops and pools.
- Assigns applications packaged with VMware® ThinApp® to specific desktops and pools.
- Manages local and remote desktop sessions.
- Establishes secure connections between workers and desktops.
- Enables single sign-on (SSO).
- Sets and applies policies.

Inside the corporate firewall, a group of two or more View Connection Server instances can be installed and configured. Their configuration data is stored in an embedded Lightweight Directory Access Protocol (LDAP) directory and is replicated among members of the group.

Outside the corporate firewall, in the DMZ, View Connection Server can be installed and configured as a security server. Security servers in the DMZ communicate with the View Connection Servers inside the corporate firewall. Security servers ensure that only remote desktop traffic having a strongly authenticated user can enter the corporate datacenter. Workers are only allowed to access the desktop resources to which they are authorized.

Security servers offer a subset of functionality and they are not required to be in an Active Directory domain. View Connection Server can be installed on Windows Server® 2008 R2, preferably on a VMware virtual machine.

**View Client**

View Client enables workers to access VMware View desktops on the various Client Devices, as described above. After logging on, workers can select from a list of virtual desktops that each is authorized to use. Authorization can require Active Directory credentials, a User Principal Name (UPN), a smart card personal identification number (PIN), or an RSA SecurID or other two-factor authentication token.

An administrator can configure View Client to allow workers to select a display protocol. Protocols include PCoIP and Microsoft Remote Desktop Protocol (RDP). The speed and quality of the PCoIP display protocol has equivalent performance to a physical PC.
VMware View® Client with Local Mode (formerly called Offline Desktop) is a version of View Client that has been extended to allow workers to download virtual machines and use them on their local Windows systems, regardless of whether or not they have a network connection.

The features of View Client differ according to which version of the client is in use. This paper focuses on View Client for Windows. The clients that are not described in detail include:

- Various third-party thin clients and zero clients, available only through certified partners.
- View Open Client, which supports the VMware partner certification program. View Open Client is not an official View Client and it is therefore not supported.

For more information about View Client, go to the VMware View Clients Documentation at: https://www.vmware.com/support/viewclients/doc/viewclients_pubs.html

View Portal

View Portal is available via any web browser regardless of PC or laptop operating system by entering the URL of the View Connection Server instance to open the View Portal page. View Portal provides links for to the VMware Download site for downloading the full View Client. The links on the View Portal page are configurable and can be set to point to an internal Web server, or limited to the client versions that are available on View Connection Server.

View Agent

View Agent service can be installed on all virtual machines, physical systems, and Terminal Service servers that are used as the source for VMware View desktops. On virtual machines, View Agent communicates with View Client to provide features such as connection monitoring, virtual printing, View Persona management, and access to locally-connected USB devices.

If the desktop source is a virtual machine, install the View Agent service on that virtual machine, and use the virtual machine as a template or as a parent for linked-clones. When a pool is created from the virtual machine, the agent is automatically installed on every virtual desktop.

Install the agent with an option for single sign-on. With single sign-on, workers are prompted to logon only when they connect to View Connection Server. They are not prompted to logon a second time to connect to a virtual desktop.

View Administrator

View Administrator is a web-based application that allows administrators to:

- Configure View Connection Server.
- Deploy and manage VMware View desktops.
- Control user authentication.
- Troubleshoot end user issues.

When a View Connection Server instance is installed, the View Administrator application is also installed. This application allows administrators to manage View Connection Server instances, without installing an application on their local computer.
View Composer

View Composer can be installed on a VMware® vCenter Server™ instance in order to manage virtual machines or it can be installed on a separate server. View Composer is used to create a pool of linked-clones from a specified parent virtual machine. This strategy reduces storage costs by up to 90 percent.

Each linked-clone acts as an independent desktop with a unique host name and IP address. However, the linked-clone requires significantly less storage because it shares a base image with the parent. Since linked-clone desktop pools share a base image, updates and patches can be quickly deployed just by updating the parent virtual machine. User settings, user data, and applications are not affected. Linked-clone technology for VMware View desktops can also be downloaded and checked out for use on local systems.

Although VMware View 5.1 enables View Composer to be installed on its own server host, a View Composer service can operate with only one vCenter Server instance. Similarly, a vCenter Server instance can be associated with only one View Composer service. vCenter Server is described in more detail below.

View Transfer Server

View Transfer Server manages and streamlines data transfers between the datacenter and VMware View desktops that are checked out for use on the worker’s local system. View Transfer Server is required to support desktops that run View Client with Local Mode.

Several operations use View Transfer Server to send data between the VMware View desktop in vCenter Server and the corresponding local desktop on the client system:

- When a user checks a desktop in or out, View Manager authorizes and manages the operation. View Transfer Server transfers the files between the datacenter and the local desktop.
- View Transfer Server synchronizes local desktops with the corresponding desktops in the datacenter by replicating user-generated changes to the datacenter. Replication occurs at intervals that are specified in local mode policies. Replication can be initiated in View Administrator. A policy can also be set to allow workers to initiate replications from their local desktops.
- View Transfer Server distributes common system data from the datacenter to local clients. View Transfer Server downloads View Composer base images from the Transfer Server repository to local desktops.

VMware vCenter Server

VMware vCenter Server acts as a central administrator for VMware ESX and ESXi servers that are connected on a network. vCenter Server, provides the central point for configuring, provisioning, and managing virtual machines in the datacenter.

In addition to using these virtual machines as sources for VMware View desktop pools, virtual machines can be used to host the server components of VMware View, including Connection Server instances, Active Directory servers, and vCenter Server instances.
View Composer can be installed on the same server as vCenter Server to create linked-clone desktop pools. vCenter Server then manages the assignment of the virtual machines to FlexPod and manages the assignment of CPU and memory resources to virtual machines.

Note that vCenter Server is installed on a Windows Server 2008 R2 server, preferably on a VMware virtual machine.

**FlexPod Architecture**

The FlexPod architecture is a modular or “pod-like” architecture that can vary in its exact configuration, and it can be scaled easily as requirements change.

The benefits of the Flexpod architecture include:

- **Operational efficiency**: The converged Flexpod architecture reduces processes, minimizing operational challenges such as configuration errors and administrative costs. FlexPod also reduces risk by leveraging a pre-validated infrastructure stack that includes cooperative support from an ecosystem of industry-leading partners including VMware®, Cisco®, and NetApp®.

- **Performance**: The Flexpod architecture allows VMware View desktops to be scaled on demand, providing increased service levels and resource optimization. For example, a base FlexPod configuration can securely support 1,500 VMware View workers, and have the capacity to scale-up or scale-out. This configuration can also be used to accommodate other workloads simultaneously, such as Microsoft® Exchange™, SQL Server®, and SharePoint®.

- **Agility**: Deploying VMware View with Flexpod enables enterprises to rapidly deploy a pre-tested and validated VDI to increase the time-to-market for workers.

- **Flexibility**: The Flexpod architecture offers seamless integration to existing technology, and it protects and maximizes current investments by incorporating existing infrastructure. Both FlexPod and VMware View provide proactive, predictive, and centralized management that can be integrated with third-party tools.

FlexPod is comprised of a hardware and software configuration that serves as an integrated foundation for all virtualization solutions, as shown below. The Flexpod architecture is flexible and it enables the networking, computing, and storage to fit in a single datacenter rack, or it can be deployed according to a customer’s datacenter design. Port density enables the networking components to accommodate multiple configurations.

The FlexPod architecture can be customized or “flex” the VMware View environment to suit customer requirements. The reference architecture detailed in this paper describes the resiliency, cost benefit, and ease of deployment of the storage solution. A storage system capable of serving multiple protocols across a single interface allows for customer choice and investment protection because it truly is “wired-once” architecture.
Figure 2 provides an example configuration of a VMware View 5.1 deployment built on FlexPod. This example configuration uses the Cisco Unified Computing System™ (Cisco UCS) B-Series, an Access Layer consisting of Cisco Nexus® 5548UP networking, and the NetApp® Fabric-Attached Storage (FAS) Systems. Each of these components is described in the sections below.

This example configuration also uses Cisco® B200-M2, Cisco® 6240 Fabric Interconnect and Cisco® 5548 Access Layer Switch to build a resilient server platform that balances simplicity and performance for virtual desktop deployments.

**Cisco Unified Computing System**

Cisco Unified Computing System is the first converged datacenter platform that combines industry-standard, x86-architecture servers with networking and storage access to form a single converged system. Cisco UCS is entirely programmable using unified model-based management to simplify and speed the deployment of VMware View.

The system’s x86-architecture rack-mount and blade servers are powered by Intel® Xeon® processors. These industry-standard servers deliver the performance needed to power business-critical workloads. Cisco® servers, combined with a simplified and converged architecture, drive IT productivity and price/performance, thereby ensuring a lower TCO.
Cisco UCS is integrated with a standards-based, high-bandwidth, low-latency, virtualization-aware unified fabric. The system is wired once to support the desired bandwidth and it carries all traffic, including IP, storage, management, virtual, and others, with security isolation, visibility, and control that is equivalent to physical networks. The system meets the bandwidth demands of today’s multi-core processors, eliminates costly redundancy, and increases workload agility, reliability, and performance.

Cisco UCS is designed from the ground-up to be programmable and self integrating. The entire hardware stack of a server, ranging from the server firmware and settings to network profiles, is configured through model-based management. With Cisco UCS Virtual Interface Cards (VICs), even the number and type of I/O interfaces is programmed dynamically, making every server ready to power any workload at any time.

With model-based management, administrators can manipulate a model of a desired system configuration, associate a model’s service profile with hardware resources, and the system configures itself to match the model. This automation speeds provisioning and workload migration with accurate and rapid scalability. The result is increased IT staff productivity, improved compliance, and reduced risk of failures due to inconsistent configurations.

Cisco Fabric Extender technology condenses three network layers into one. By simplifying the architecture, fewer system components are required to accomplish the same solution. This represents a radical simplification over traditional systems. This technology reduces capital and operating costs while increasing business agility, simplifying and speeding deployment, and improving performance.

Cisco UCS enables organizations to become more effective by using technologies that breed simplicity, rather than complexity. The result is flexible, agile, high-performance, self-integrating information technology, reduced staff costs with increased uptime through automation, and more rapid return on investment (ROI).

**NetApp FAS Storage Systems**

When implementing storage for the VMware View infrastructure, virtual desktops introduce a new requirement for datacenter storage. NetApp FAS Storage Systems, including the FAS 22xx, 32xx, and 62xx series, ensure optimal performance at the lowest cost through highly efficient de-duplication, cloning, data protection, and replication technologies. Netapp FAS Storage Systems provide for reduced costs and simplify desktop data management throughout the information lifecycle. The cascading effect of NetApp storage efficiency across all tiers of data management means that storage costs drop as more virtual desktop workers are added to the system.

The NetApp® Unified Storage Architecture offers customers an agile and scalable storage platform. NetApp FAS Storage Systems use the Data ONTAP® operating system to provide a single unified platform that enables all workloads to be hosted on the same storage array. NetApp FAS Storage Systems support Storage Area Network (SAN), Fibre Channel over Ethernet (FCoE), Fibre Channel (FC), Internet Small Computer System Interface (iSCSI), Networked-attached Storage (NAS) including Common Internet File System (CIFS) and Network File System (NFS), and primary and secondary storage.
NetApp FAS Storage Systems use a single process for activities such as installation, provisioning, mirroring, backup, and upgrading throughout the entire product line, from the entry level to enterprise-class controllers. Having a single set of software and processes simplifies even the most complex enterprise data management challenges. Unifying storage and data management software and processes streamlines data ownership, enables companies to adapt to changing business needs without interruption, and reduces the TCO.

With shared infrastructure, it is critical for the storage architecture to provide a high level of availability and performance because storage outages or performance issues can affect thousands of workers. For more detailed information, see the Resources section below for the NetApp set of best practice documents.

NetApp® VASA Provider 1.0 supports the VMware vStorage APIs for Storage Awareness (VASA). The VASA API permits storage arrays to integrate with VMware® vCenter™ for management functionality. This API serves as a link between NetApp FAS Storage Systems and VMware vCenter Server to collect data about storage capabilities, topology, and other information. The VASA API delivers this data to vCenter Server to enable policy enforcement, and to simplify the operational management of storage in a virtualized environment. In particular, the VASA API can monitor when thin provisioned datastores have reached a defined threshold, or if there have been any changes to a published storage capability (for a volume or LUN).

Planning and Design Guidance

VMware View 5.1 deployment built on FlexPod provides a standard, scalable architecture that can be adapted to the enterprise environment and any special requirements. The key details about the requirements for memory, CPU, storage capacity, and network components for deploying VMware View with FlexPod are described below. These planning and design guidelines describe the tested architecture used for this paper, and provide additional information based on validations from VMware, Cisco, and NetApp.

Types of Workers

Many configuration elements, including RAM, CPU, and storage sizing requirements depend largely on the type of worker who uses the virtual desktop, and on the applications that must be installed. For architecture planning, workers can be categorized into several types.

Table 1: Types of workers using virtual desktop

<table>
<thead>
<tr>
<th>Type of worker</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Task workers</td>
<td>Task workers and administrative workers perform repetitive tasks within a small set of applications, usually at a stationary computer. The applications are typically not as CPU- and memory-intensive as the applications used by knowledge workers. Task workers who work on specific shifts might all log on to their virtual desktops at the same time. Task workers include call center analysts, retail employees, warehouse workers, and so on.</td>
</tr>
<tr>
<td>Knowledge workers</td>
<td>Knowledge workers daily tasks include accessing the internet, using email, and creating complex documents, presentations, and spreadsheets. Knowledge workers include accountants, sales managers, marketing research analysts, and so on.</td>
</tr>
</tbody>
</table>
## Type of worker | Description
--- | ---
Power workers | Power workers include application developers and people who use graphics-intensive applications.
Employees who use desktops in Local Mode only | These workers download and run VMware View desktops only on their local systems. This reduces datacenter costs associated with bandwidth, memory, and CPU resources. Scheduled replication ensures that systems and data are backed up appropriately. Administrators configure how often workers' systems must contact View Manager to avoid being locked out.
Kiosk workers | These workers share a desktop in a public place. Examples of kiosk workers include students using a shared computer in a classroom, nurses at nursing stations, and computers used for job placement and recruiting. These desktops require automatic logon. Authentication can be done through certain applications if necessary.

### Estimating Memory Requirements for Virtual Desktops

The cost of RAM is greater for servers than for personal computers, and high as compared to the overall cost of server hardware and total storage capacity. For this reason, determining the correct memory allocation is crucial to planning a successful desktop deployment.

If RAM allocation is too low, storage I/O can be negatively affected because it results in too much memory swapping. If the RAM allocation is too high, storage capacity can be negatively affected because the paging file in the guest operating system, and the swap and suspend files for each virtual machine, grow too large.

Note that this section addresses memory allocation for remote access to VMware View desktops. If workers run View Client with Local Mode for desktops on their client systems, the amount of memory used is only a proportion of that available on the client device.

Memory is required to run the host operating system on the client computer. It is also required to run the VMware View desktop operating system and applications on the client computer and VMware View desktop. VMware recommends using 2 GB or greater RAM when running the Windows XP and Windows Vista® operating systems, and 3 GB or greater for Windows 7.

When a desktop configured in vCenter Server requires more memory than the local client system can accommodate, it cannot be checked out until changes are made to Windows registry setting. For more information, see the VMware View Administration link in the Resources section below.

### RAM Sizing Impact on Performance

When allocating RAM, avoid choosing conservative settings and consider the following:

- Insufficient RAM allocations can cause excessive guest swapping. This generates I/O that causes significant performance degradations and increases storage I/O load.
- VMware ESX/ESXi supports sophisticated memory resource management algorithms, such as transparent memory sharing and memory ballooning, that can significantly reduce the physical RAM required to support a given guest RAM allocation. For example, although 2 GB might be allocated to a virtual desktop, only a fraction of that is consumed in physical RAM.
• Virtual desktop performance is sensitive to response time. For an ESX/ESXi host, set non-zero values for RAM reservation settings. Reserving some RAM guarantees that idle but in-use desktops are never completely swapped out to disk. It can also reduce storage space consumed by ESX/ESXi swap files. However, higher reservation settings affect the ability to over-commit memory on an ESX/ESXi host and they might affect VMware vSphere® vMotion® maintenance operations.

**RAM Sizing Impact on Storage**

The amount of RAM allocated to a virtual machine is directly related to the size of certain files that the virtual machine utilizes. In order to access the files in the list shown in the table below, use the Windows guest operating system to locate the Windows page and hibernate files, and use the ESX/ESXi host file system to locate the ESX/ESXi swap and suspend files.

**Table 2: RAM sizing impact on storage**

<table>
<thead>
<tr>
<th>File type</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Windows page file</td>
<td>By default, the Windows page file is sized at 150 percent of guest RAM. This file, located by default at C:\pagefile.sys, causes thin provisioned storage to grow because it is accessed frequently. On linked-clone virtual machines, the page file and temporary files can be redirected to a separate virtual disk that is deleted when the virtual machines are powered off. Disposable page-file redirection saves storage, thereby slowing the growth of linked-clones and improving performance. Note that the size can also be adjusted from Windows, although it might have a negative effect on application performance.</td>
</tr>
<tr>
<td>Windows hibernate file for laptops</td>
<td>This file can equal 100 percent of guest RAM. This file can be safely deleted because it is not required in VMware View deployments, even with View Client in Local Mode.</td>
</tr>
<tr>
<td>ESX/ESXi swap file</td>
<td>This file (.vswp extension) is created if less than 100 percent of a virtual machine’s RAM is reserved. The size of the swap file is equal to the unreserved portion of guest RAM. For example, if 50 percent of guest RAM is reserved and guest RAM is 2 GB, the ESX/ESXi swap file is 1 GB. This file can be stored on the local datastore on the ESX/ESXi host or cluster.</td>
</tr>
<tr>
<td>ESX/ESXi suspend file</td>
<td>This file (.vmss extension) is created when the desktop pool logoff policy is set to suspend the virtual desktop when the end user logs off. The size of this file is equal to the size of guest RAM.</td>
</tr>
</tbody>
</table>

**RAM Sizing for Monitor Configurations When Using PCoIP**

For the display protocol from VMware, the amount of extra RAM that the ESX/ESXi host requires depends in part on the number of monitors configured for workers, and on the display resolution. The table below lists the amount of overhead RAM required for the various display configurations. The amount of memory listed in the columns is in addition to the amount of memory required to support other PCoIP functionality.
Table 1: PCoIP client display overhead

<table>
<thead>
<tr>
<th>Display resolution standard</th>
<th>Width, in pixels</th>
<th>Width, in pixels</th>
<th>1-Monitor overhead</th>
<th>2-Monitor overhead</th>
</tr>
</thead>
<tbody>
<tr>
<td>VGA</td>
<td>640</td>
<td>480</td>
<td>2.34 MB</td>
<td>4.69 MB</td>
</tr>
<tr>
<td>SVGA</td>
<td>800</td>
<td>600</td>
<td>3.66 MB</td>
<td>7.32 MB</td>
</tr>
<tr>
<td>720p</td>
<td>1280</td>
<td>720</td>
<td>7.03 MB</td>
<td>14.65 MB</td>
</tr>
<tr>
<td>UXGA</td>
<td>1600</td>
<td>1200</td>
<td>14.65 MB</td>
<td>29.30 MB</td>
</tr>
<tr>
<td>1080p</td>
<td>1920</td>
<td>1080</td>
<td>15.82 MB</td>
<td>31.64 MB</td>
</tr>
<tr>
<td>WUXGA</td>
<td>1920</td>
<td>1200</td>
<td>17.58 MB</td>
<td>35.16 MB</td>
</tr>
<tr>
<td>QXGA</td>
<td>2048</td>
<td>1536</td>
<td>24.00 MB</td>
<td>48.00 MB</td>
</tr>
<tr>
<td>WQXGA</td>
<td>2560</td>
<td>1600</td>
<td>31.25 MB</td>
<td>62.50 MB</td>
</tr>
</tbody>
</table>

RAM Sizing for Specific Workloads and Operating Systems

As the amount of RAM required can vary widely depending on the type of user, many companies first determine the correct setting for various pools of workers in their enterprise. A good starting point is to allocate 1 GB for the Windows® XP, 32-bit Windows Vista®, and 32-bit Windows 7 desktops, and 2 GB for 64-bit Windows 7 desktops. The best practice is to monitor the performance and disk space used with various types of workers and make adjustments to determine the optimal setting for each pool of workers.

Estimating CPU Requirements for Virtual Desktops

When estimating CPU requirements, first gather information to determine the average CPU utilization for various types of workers in the enterprise. In addition, calculate another 10 to 25 percent of processing power for virtualization overhead and peak periods of usage.

Note that this section addresses issues regarding CPU requirements when accessing VMware View desktops remotely. If workers are running a View Client with Local Mode on their desktop client systems, the VMware View desktop uses the available CPUs on the client device to a maximum of 2 CPUs.

CPU requirements vary according to the worker type. In this case, use a performance monitoring tool, such as Perfmon in the virtual machine, esxtop in ESX/ESXi, or vCenter performance monitoring tools, to understand both the average and peak CPU usage levels for each user group. Adhere to these guidelines:

- Software developers and other power workers with high performance needs might have higher CPU requirements than knowledge and task workers. Dual virtual CPUs are recommended for compute-intensive tasks, such as to play 720p video using the PCoIP display protocol, and for Windows 7 desktops.

- Single virtual CPUs are generally recommended for other cases.
Because many virtual machines run on one server, CPU can spike if agents, such as antivirus agents, all check for updates at exactly the same time. Determine which agents and how many agents could cause performance issues and adopt a strategy for addressing these issues. For example, strategies might be helpful in the enterprise include:

- Use View Composer to update images, rather than having software management agents download software updates to each individual virtual desktop.
- Schedule antivirus and software updates to run at non-peak hours, when fewer workers are logged on.
- Stagger or randomize when updates occur.

As an informal initial sizing approach, assume that each virtual machine requires 1/8 to 1/10 of a CPU core as the minimum guaranteed compute power. That is, use 8 to 10 virtual machines per core. For example, suppose the sizing assumes that there are 8 virtual machines per core and there is a 2-socket, 8-core ESX/ESXi host. In this case, 128 virtual machines can be hosted on the server. Monitor the overall CPU usage on the host during this period and ensure that it rarely exceeds the given safety margin such as 80 percent to ensure that there is enough headroom for spikes.

**Choosing the Appropriate System Disk Size**

When allocating disk space, provide enough space for the operating system, applications, and additional content that workers might decide to install or generate. Usually, the amount of disk space is smaller than the size of the disk that is included on a physical PC.

Because datacenter disk space usually costs more per gigabyte than desktop or laptop disk space in a traditional PC deployment, optimize the operating system image size. Suggestions for optimizing the image size include:

- Remove unnecessary files. For example, reduce the quotas on temporary Internet files. Choose a virtual disk size that is sufficient to allow for future growth, but not one that is unrealistically large.
- Use centralized file shares or a View Composer persistent disk for user-generated content and user-installed applications.

For each virtual desktop, the amount of storage space required must take these files into account:

- ESX/ESXi suspend file is equivalent to the amount of RAM allocated to the virtual machine.
- Windows page file is equivalent to 150 percent of RAM.
- Log files use approximately 100 MB for each virtual machine.
- The virtual disk, also called the .vmdk file, must accommodate the operating system, existing applications, future applications, and software updates. The virtual disk must also accommodate local user data and user-installed applications if they are located on the virtual desktop, rather than on file shares.
If you use View Composer, the .vmdk files grow over time. However, the amount of growth can be controlled by scheduling View Composer refresh operations, setting a storage over-commit policy for VMware View desktop pools, and redirecting Windows page and temporary files to a separate non-persistent disk.

In addition, add 15 percent to this estimate to ensure that workers do not run out of disk space.

**VMware View ESX/ESXi Node**

A node is a single VMware ESX/ESXi host that hosts virtual machine desktops in a VMware View deployment.

VMware View is most cost-effective when the consolidation ratio is maximized. The consolidation ratio is the number of desktops hosted on an ESX/ESXi host. Although many factors affect server selection, when optimizing strictly for acquisition price, the server configuration requires an appropriate balance of processing power and memory.

There is no substitute for measuring performance under actual, real-world scenarios such as in a pilot to determine an appropriate consolidation ratio for the environment and hardware configuration. Consolidation ratios can vary significantly based on usage patterns and environmental factors. In general, follow these guidelines:

- As a general framework, consider compute capacity in terms of 8 to 10 virtual desktops per CPU core. For information about calculating CPU requirements for each virtual machine, see the [Estimating CPU Requirements for Virtual Desktops](#) section above.

- Determine memory capacity in terms of virtual desktop RAM, host RAM, and over-commit ratio. Although there are between 8 and 10 virtual desktops per CPU core, if virtual desktops use 1 GB or more of RAM, carefully consider the physical RAM requirements. For information about calculating the amount of RAM required per virtual machine, see the [Estimating Memory Requirements for Virtual Desktops](#) section above.

- Consider cluster requirements and any failover requirements. For more information, see the [Determining Requirements for High Availability](#) section below.

**Additional Recommendations Targeting Minimal Capital Expenditure**

When increasing the number of virtual machines per ESX/ESXi host, reduce the number of ESX/ESXi hosts required for the View Client with Local Mode pool. An ESX/ESXi 4.1 host can accommodate up to 500 virtual machines if most are not powered on at the same time, as is frequently the case for Local Mode pools.

Reduce the amount of bandwidth and I/O operations required by each virtual machine and maximize the number of virtual machines on an ESX/ESXi host using these recommendations:

- Set a VMware View policy that enables workers to use their VMware View desktops in Local Mode only. With this setting, the virtual machines in the datacenter remain locked and powered off.
- Set Local Mode policies such that workers cannot initiate desktop rollbacks, data backups, or check-ins to the datacenter.

- Do not schedule automatic backups.

- Do not turn on SSL for provisioning or downloading Local Mode desktops.

- If the performance of View Connection Server is affected by the number of local desktops, set the heartbeat interval to be less frequent. The heartbeat informs the View Connection Server that the local desktop has a network connection. The default interval is five minutes.

**Desktop Virtual Machine Configuration**

The amount of RAM, CPU, and disk space that virtual desktops require depends on the guest operating system. For this reason, separate configuration examples are provided for Windows XP, Windows Vista, and Windows 7 virtual desktops in the tables below. Note that the example settings for virtual machines such as memory, number of virtual processors, and disk space are VMware View-specific.

The guidelines listed in Table 4 are for a standard Windows XP virtual desktop running in remote mode.

**Table 2: Windows XP desktop virtual machine example**

<table>
<thead>
<tr>
<th>Item</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>Operating system</td>
<td>32-bit Windows XP (with the latest service pack)</td>
</tr>
<tr>
<td>RAM</td>
<td>1 GB (512 MB low-end, 2 GB high-end)</td>
</tr>
<tr>
<td>Virtual CPU</td>
<td>1</td>
</tr>
<tr>
<td>System disk capacity</td>
<td>16 GB (8 GB low-end, 40 GB high-end)</td>
</tr>
<tr>
<td>User data capacity</td>
<td>5 GB (starting point)</td>
</tr>
</tbody>
</table>

The amount of system disk space required depends on the number of applications required in the base image. VMware has validated a setup that includes 8 GB of disk space. Applications include Microsoft Word®, Excel®, PowerPoint®, Adobe® Reader®, Internet Explorer®, McAfee® Antivirus, and PKZIP®.

The amount of disk space required for user data depends on the role of the end user and organizational policies for data storage. For View Composer, this data is kept on a persistent disk.

The guidelines listed in Table 5 are for a standard Windows Vista virtual desktop running in remote mode.

**Table 5: Windows Vista desktop virtual machine example**

<table>
<thead>
<tr>
<th>Item</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>Operating system</td>
<td>32-bit Windows Vista (with the latest service pack)</td>
</tr>
<tr>
<td>RAM</td>
<td>1 GB</td>
</tr>
</tbody>
</table>
VMware View® with FlexPod®

DEPLOYMENT AND TECHNICAL CONSIDERATIONS GUIDE /21

<table>
<thead>
<tr>
<th>Item</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>Virtual CPU</td>
<td>1</td>
</tr>
<tr>
<td>System disk capacity</td>
<td>20 GB (standard)</td>
</tr>
<tr>
<td>User data capacity</td>
<td>5 GB (starting point)</td>
</tr>
</tbody>
</table>

The guidelines listed in Table 6 are for a standard Windows 7 virtual desktop running in remote mode.

Table 6: Windows 7 desktop virtual machine example

<table>
<thead>
<tr>
<th>Item</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>Operating system</td>
<td>32-bit Windows 7 (with the latest service pack)</td>
</tr>
<tr>
<td>RAM</td>
<td>1 GB</td>
</tr>
<tr>
<td>Virtual CPU</td>
<td>1</td>
</tr>
<tr>
<td>System disk capacity</td>
<td>20 GB (slightly less than standard)</td>
</tr>
<tr>
<td>User data capacity</td>
<td>5 GB (starting point)</td>
</tr>
</tbody>
</table>

vCenter and View Composer Virtual Machine Configuration

vCenter Server and View Composer can be installed on the same virtual machine or on separate hosts. vCenter Server requires much more memory and processing power than a desktop virtual machine.

View Composer can create and provision up to 1,000 desktops per pool with vSphere® 4.1 or later. Beginning in VMware View 5.1, View Composer can be installed on a separate virtual machine, allowing the vCenter appliance and View Composer to be used in a VMware View deployment.

View Composer can also perform a recompose operation on up to 1,000 desktops at a time. The desktop pool size is limited by these factors:

- Each desktop pool can contain only one ESX/ESXi cluster.
- With VMware View 5.1 and later and vSphere 5.0 and later, an ESXi cluster can contain more than 8 ESXi hosts.
- Up to 32 hosts can be included in a View Composer cluster when using NFS-based datastores.
- Each CPU core has compute capacity of 8 to 10 virtual desktops.
- The number of IP addresses available for the subnet limits the number of desktops in the pool. For example, if the network subnet for the pool contains only 256 usable IP addresses, the pool size is limited to 256 desktops.
vCenter Server and View Composer can be installed on a physical machine or a virtual machine, although a virtual machine is recommended. The ESX/ESXi host for this virtual machine can be part of a VMware vSphere® High Availability (HA) cluster to guard against physical server failures.

View Connection Server and Virtual Machine Configuration

When View Connection Server is installed, the View Administrator user interface is also installed. This server requires more memory and processing resources than a vCenter Server instance.

View Connection Server Configuration

View Connection Server can be installed on a physical machine or a virtual machine, although a virtual machine is recommended. The ESX/ESXi host for the virtual machine can be part of a VMware HA cluster to guard against physical server failures.

View Connection Server Cluster Design Considerations

Multiple replicated View Connection Server instances can be deployed in a group to support load balancing and high availability. Groups of replicated instances are designed to support clustering within a LAN-connected, single datacenter environment. VMware does not recommend using a group of replicated View Connection Server instances across a WAN due to the communication traffic required between the grouped instances. In scenarios where a VMware View deployment is required to span multiple datacenters, create a separate VMware View deployment for each datacenter.

Maximum Connections for View Connection Server

View Desktop Connections provides information about the tested limits regarding the number of simultaneous connections that a VMware View deployment can accommodate.

This example assumes VMware View is being used with vSphere 4.1 and later, and vCenter Server 4.1 and later. It also assumes that View Connection Server is running on a 64-bit Windows Server® 2008 R2 Enterprise operating system.

PCoIP Secure Gateway connections are required when using security servers for PCoIP connections from outside the corporate network. Tunneled connections are required when using security servers for RDP connections from outside the corporate network, and for USB and multimedia redirection (MMR) acceleration with a PCoIP Secure Gateway connection. Multiple security servers can be paired to a single connection server.

View Transfer Server Virtual Machine Configuration and Storage

View Transfer Server is required to support desktops that run View Client with Local Mode. This server requires less memory than View Connection Server.
View Transfer Server Configuration

Install the View Transfer Server on a virtual, rather than a physical machine. The virtual machine must be managed by the same vCenter Server instance as the local desktops that manage it.

Storage and Bandwidth Requirements for View Transfer Server

Several operations use View Transfer Server to send data between the VMware View desktop in vCenter Server and the corresponding local desktop on the client system. When a user checks a desktop in or out, View Transfer Server transfers the files between the datacenter and the local desktop. View Transfer Server also synchronizes local desktops with the corresponding desktops in the datacenter by replicating user-generated changes to the datacenter.

When using View Composer with linked-clones for local desktops, the disk drive on which the Transfer Server repository is configured requires enough space to store the static image files. Image files are View Composer base images. Faster network storage disks provide better performance. For information about determining the size of the base image files, see VMware View Administration in the Resources section below.

Each Transfer Server instance can accommodate approximately 60 concurrent disk operations, although network bandwidth will most likely be saturated at a lower number. VMware tested 20 concurrent disk operations, such as 20 clients downloading a local desktop at the same time using a 1 GB per second network connection.

vSphere Clusters

VMware View deployments can use VMware HA clusters to provide for high availability and guard against physical server failures. With VMware View 5.1 and later, and vSphere 5 and later, the cluster can contain up to 32 servers or nodes when using View Composer and storing replica disks on NFS datastores.

VMware vSphere and VMware vCenter provide a rich set of features for managing clusters of servers that host VMware View desktops. The cluster configuration is also important because each VMware View desktop pool must be associated with a vCenter resource pool. Therefore, the maximum number of desktops per pool is related to the number of servers and virtual machines that are run per cluster.

In very large VMware View deployments, vCenter performance and responsiveness can be improved by having only one cluster object per datacenter object, which is not the default behavior. By default, VMware vCenter creates new clusters within the same datacenter object.

Determining Requirements for High Availability

VMware vSphere, through its efficiency and resource management, can achieve industry-leading levels of virtual machines per server. But achieving a higher density of virtual machines per server means that more workers are affected if a server fails.
Requirements for high availability can differ substantially based on the purpose of the desktop pool. For example, a stateless desktop image (floating-assignment) pool might have different recovery point objectives (RPOs) than a stateful desktop image (dedicated-assignment) pool. For a floating-assignment pool, an acceptable solution might be to have workers log on to a different desktop in the event that their desktop becomes unavailable.

In cases where availability requirements are high, the proper configuration of VMware HA is essential. If VMware HA is used, and a fixed number of desktops per server are planned, each server is required to run at a reduced capacity. If the server fails, the capacity of desktops per server is not exceeded when the desktops are restarted on a different host.

For example, in an 8-host cluster, where each host is capable of running 128 desktops, and the goal is to tolerate a single server failure, make sure that no more than 128 * (8 - 1) = 896 desktops are running on that cluster. In addition, use VMware vSphere® Distributed Resource Scheduler™ (DRS) to help balance the desktops among all eight hosts. This provides for full use of the extra server capacity without letting any hot-spare resources sit idle. Additionally, DRS can help rebalance the cluster after a failed server is restored to service.

The requirements must also ensure that storage is properly configured to support the I/O load that results when many virtual machines restart at once in response to a server failure. Storage (Input/Output Operations Per Second (IOPS) have the most effect on how quickly desktops recover from a server failure.

Network Bandwidth Considerations

For display traffic, many elements can affect network bandwidth, such as the protocol used, monitor resolution and configuration, and the amount of multimedia content in the workload. Concurrent launches of streamed applications can also cause usage spikes.

As the effects of these issues can vary widely, many companies monitor bandwidth consumption initially. As a starting point, plan for 150 to 200 Kbps of capacity for a typical knowledge user.

When the PCoIP display protocol has an enterprise LAN with 100 Mb or a 1 Gb switched network, workers can expect excellent performance under these conditions:

- Two monitors (1920 x 1080)
- Heavy use of Microsoft® Office® applications
- Heavy use of Adobe® Flash® embedded web browsing
- Frequent use of multimedia with limited use of full-screen mode
- Frequent use of USB-based peripherals
- Network-based printing
Optimization Controls Available with PCoIP

By using the PCoIP display protocol from VMware, several elements that affect bandwidth usage can be adjusted:

- Adjust the size of the image cache on Windows and Linux client systems, from 50 MB to 300 MB. Image caching reduces the amount of display data retransmitted.

- Configure the image quality level and frame rate used during periods of network congestion. The quality level setting can be used to limit the initial quality of the changed regions of the display image. Unchanged regions of the image progressively build to a lossless (perfect) quality. Adjust the frame rate from 1 to 120 frames per second.

  This control works well for static screen content that does not need to be updated or in situations where only a portion needs to be refreshed.

- Turn off the Build-to-Lossless feature if instead of progressively building to perfect quality (lossless), the requirement is to build to perceptual lossless.

- Control which encryption algorithms are advertised by the PCoIP endpoint during session negotiation. By default, both Salsa20-256round12 and AES-128-GCM algorithms are available.

- With regards to session bandwidth, configure the maximum bandwidth in kilobits per second to correspond to the type of network connection, such as a 4 Mbits Internet connection. The bandwidth includes all imaging, audio, virtual channel, USB, and control PCoIP traffic.

- Configure a lower limit in kilobits per second for bandwidth that is reserved for the session to ensure that a user does not have to wait for bandwidth to become available. Specify the Maximum Transmission Unit (MTU) size for User Datagram Protocol (UDP) packets for a PCoIP session, from 500 to 1500 bytes.

- Specify the maximum bandwidth that can be used for audio (sound playback) in a PCoIP session.

WAN Support and PCoIP

For WANs, consider bandwidth constraints and latency issues. The PCoIP display protocol provided by VMware adapts to varying latency and bandwidth conditions.

When using the RDP display protocol, use a WAN optimization product to accelerate applications for workers in branch offices or small offices.

With PCoIP, many WAN optimization techniques are built into the base protocol:

- WAN optimization is valuable for TCP-based protocols such as RDP because these protocols require many handshakes between client and server. The latency of these handshakes can be quite large. WAN accelerators spoof replies to handshakes enabling the latency of the network to be hidden from the protocol. Because PCoIP is UDP-based, this form of WAN acceleration is unnecessary.
• WAN accelerators can compress network traffic between client and server, but this compression is usually limited to 2:1 compression ratios. PCoIP is able to provide compression ratios of up to 100:1 for images and audio.

Bandwidth Requirements for Various Types of Users
While determining minimum bandwidth requirements for PCoIP, plan with these estimates:

• 100 to 150 Kbps average bandwidth for basic office productivity desktop, typical Office applications with no video, no 3D graphics, and default Windows and VMware View settings.

• 50 to 100 Kbps average bandwidth for an optimized Office productivity desktop, typical Office applications with no video, no 3D graphics, with Windows desktop settings optimized and VMware View optimized.

• 400 to 600 Kbps average bandwidth for virtual desktops utilizing multiple monitors, 3D, Aero, and Office 2010.

• 500 Kbps to 1 Mbps minimum peak bandwidth to provide headroom for bursts of display changes. In general, size your network using the average bandwidth, but consider peak bandwidth to accommodate bursts of imaging traffic associated with large screen changes.

• 2 Mbps per simultaneous user running 480p video, depending on the configured frame rate limit and the video type.

The estimate of 50 to 150 Kbps per typical user is based on the assumption that all workers are operating continuously and performing similar tasks over an 8- to 10-hour day. 50 Kbps bandwidth usage figure is from View Planner testing on a LAN with the Build-to-Lossless feature disabled. Situations may vary where some workers may be fairly inactive and consume almost no bandwidth, allowing more workers per link. Therefore, these guidelines are intended to provide a starting point for more detailed bandwidth planning and testing.

Cisco Best Practices
The Cisco features to consider when deploying VMware View with FlexPod include UCS Manager with UCS service profile templates, Virtual Local Area Network (VLAN) configuration, and Quality of Service (QoS) implementation. Each of these features is described in the sections below.

UCS Manager with UCS Service Profiles
The Cisco UCS Manager contains UCS service profile templates that allow IT administrators to create a custom hardware baseline for the type of virtual desktop workload, and for as many servers as required. Service profiles can be abstracted from the specifics of a given server to create a service profile template. This template defines policies that can be applied any number of times to provision any number of servers.
By using service profiles, Cisco UCS Manager provides logical grouping capabilities for both physical servers, and service profiles and the associated templates. This pooling or grouping, combined with fine-grained, role-based access, allows businesses to treat a farm of compute blades as flexible resource pools. These pools can be reallocated in real time to meet the needs of virtual desktop user pools, while maintaining any organizational overlay on the environment.

From a VMware View perspective, service profile templates can be used to define firmware policies, enable BIOS settings, configure the local disk array, configure memory speed, and select a boot sequence. Once created, these templates can be deployed to as many blades as required.

**VLAN Configuration**

In Cisco UCS, a named VLAN creates a connection to a specific external LAN. This VLAN isolates traffic to that external LAN including any broadcast traffic.

In VMware View instances, a VLAN is reserved for the kernel traffic (connection to vCenter) and is placed into a separate network and assigned a unique IP address. A second VLAN is reserved for VMware vMotion and is placed in a separate network and assigned a second unique IP address. Neither of these interfaces is seen by the virtual desktops.

For security reasons, the virtual desktops do not have access to the VLANs. One final reserved “vETH” may be created if the VMware vSphere hypervisor datastore is connected via Ethernet-based storage (NFS). The Ethernet-based storage traffic is isolated into a separate VLAN.

Additional VLANs are created for the virtual desktops. The IP addresses can be statically defined or acquired via Dynamic Host Configuration Protocol (DHCP). The addresses are assigned and subnet mask are created according to the size of the virtual desktop pool. Virtual desktops can migrate only within the same pool (subnet, VLAN, and so on).

**QoS Implementation**

Quality of service (QoS) is implemented where there is contention for limited network resources, such as bandwidth and queues inside switches, in key network switches and access points. The network maintains awareness of the virtual machine user session, permitting QoS and security policies to be applied at the virtual desktop level, and then prioritizing media and applications across VDI sessions. Service policies that constrain the amount of bandwidth that is dedicated to a given protocol are defined and applied at this point. These same queuing and bandwidth configurations can be placed anywhere there is a rate transition from high-bandwidth to low-bandwidth connectivity.

Specific QoS recommendations for VMware View 5.1 include:

- Ensure guaranteed network bandwidth for PCoIP traffic during congestion. In general, set PCoIP traffic to have 80 percent of the remaining bandwidth after the higher priority traffic is allocated. For example, consider a network that guarantees 20 percent of a link bandwidth for critical traffic such as VoIP. Set PCoIP to receive 80 percent of the remaining bandwidth, or 64 percent. This enables other protocols, such as file transfers or web traffic, to transfer traffic without starving the PCoIP sessions.
- Ensure proper delivery of PCoIP by tagging it in QoS to compete fairly across the network with other real time protocols. In addition, prioritize PCoIP above other non-critical and latency tolerant protocols, such as file transfers or print jobs. Failure to tag PCoIP properly in a congested network environment leads to PCoIP packet loss and a poor user experience, as PCoIP adapts down in response.

- Tag and classify PCoIP as interactive real time traffic. Generally, classify PCoIP just below VoIP, but above all other TCP-based traffic. While this recommendation is likely to have a far larger effect in a WAN scenario, consider it a best practice for LAN environments as well.

**NetApp Best Practices**

The Netapp features to consider when deploying VMware View with Flexpod include deduplication, space reclamation, read and write I/O optimization, and virtual storage tiering. Each of these items is described in the sections below.

**Deduplication**

Data deduplication is a means of reducing storage space. It works by eliminating redundant data and ensures that only one unique instance of the data is actually retained on storage media, such as disk or tape. Redundant data is replaced with a pointer to the unique data copy. NetApp deduplication saves space on primary storage by removing redundant copies of blocks within a volume hosting hundreds of virtual desktops. This process is transparent to the application and user, and it can be enabled and disabled on the fly.

In a VMware View environment, data deduplication provides significant space savings, given that each virtual machine enables an identical copy of the operating system, applications, and patches. Deduplication is also ideal for user and persona (profile) data stored in CIFS home directories. Note that not all data within a VDI environment is ideal for deduplication. Data such as swap and other transient data should not be deduplicated.

Deduplication guidelines include:

- Deduplication is configured and operates on flexible volumes only.
- Data can be deduplicated up to 255:1 without consuming additional space.
- Each storage platform has different deduplication limits.
- Each volume has dense and non-dense size limits.
- Deduplication is configured using the command line.
- Data ONTAP 7.2.5.1, 7.3P1, or later is required.
- Both A-SIS and NearStore® must be licensed for deduplication to work.
- Deduplication must be run before Snapshot copies are created or SnapMirror or SnapVault updates are run.
Space Reclamation

A virtual machine can be thinly or thickly provisioned. When a virtual machine is thin provisioned, storage for that virtual machine is not pre-allocated on the storage controller. This allows for oversubscription of the storage controller in order to increase overall utilization. The drawback to thin provisioning of virtual machines is that they are storage efficient only initially. Once thin provisioned, the virtual machine is written and storage is then allocated in the virtual machine, and in the shared storage controller. Even if the data is deleted within the guest, the storage controller storage continues to be allocated.

NetApp Virtual Storage Console 2.1.1 (VSC) plug-in introduces a technology called Space Reclamation that can be used on any Windows virtual machine that uses New Technology File System (NTFS) and resides on an NFS datastore. Space Reclamation allows the storage controller to reclaim storage space that would otherwise be wasted.

When a virtual machine is thin provisioned, the amount of storage used within the guest equals the amount of storage on the storage controller. When data is added to the virtual machine, it is consumed on the storage controller. When some of the data is deleted in the virtual machine, nothing happens to the storage. Space Reclamation brings storage efficiency to thin provisioning by returning wasted space back to the storage controller.

Read and Write I/O Optimization

VMware View desktops can be both read- and write-intensive at different times during the lifecycle of the desktop, depending on the user activity and desktop maintenance cycle. Performance-intensive activities called storm activities are experienced by most large-scale deployments, such as boot storms, log-on storms, and virus scan or definition update storms.

A boot storm is an event in which some or all virtual desktops boot simultaneously, creating a large spike in I/O. This can happen as a result of rolling out mass operating system updates and having to reboot, desktop redeploy operations, new application installations, maintenance windows, server failures, or any number of practical issues or interruptions. Daily log-on storms and virus scan storms also create similar I/O spikes.

With virtual desktops using a shared infrastructure, these peaks in I/O affect the entire desktop environment. The environment must be able to handle both the read- and write-intensive scenarios in the desktop lifecycle. The typical methods for addressing these peaks are increasing cache for ESX servers, storage devices, spindle count, and the number of storage arrays.

Virtual Storage Tiering

Virtual Storage Tiering (VST) is performed within the NetApp® Data ONTAP™ operating system. It can be extended with the use of Flash Cache. Flash Cache is a Peripheral Component Interconnect (PCI) Express card that can be installed on many of the current NetApp storage controller systems. Flash Cache is the hardware component and the software component is called FlexScale.
VST can be extended with the use of Flash Cache. As long as that block has not been evicted from both caches, all subsequent reads are performed from the main memory or Flash Cache, thereby improving performance by not having to go to disk. Again, the more heavily the data is deduplicated and the more frequently accessed, the longer it stays in cache. Transparent storage array caching combined with NetApp disk deduplication provides for cost savings on many levels.

By using Flash Cache, in addition to Data ONTAP, VST is based on the amount of deduplicated data and the percentage of reads within the environment. As workers of the VMware View environment create more data, the amount of deduplicated data changes, thus affecting the cache hit rate. In this case, more cache might be required if the data becomes more unique, even after running regular deduplication operations on the new data.

The net result of VST is that customers can buy less storage because of read cache and allow the disk to be used for write I/O. Using deduplication and VST can greatly enhance the end user experience.

Conclusion

VMware View enables organizations to increase corporate IT control, manageability, and flexibility, without increasing cost, and provide end users with a familiar desktop experience. VMware® View® 5.1 on FlexPod™ provides the foundation for delivering efficient and scalable virtual desktop infrastructure using a pre-validated infrastructure stack that includes industry-leading partners including VMware®, Cisco®, and NetApp®.

This paper describes the Flexpod reference architecture for VMware View 5.1 that include the Cisco Unified Computing System™ and NetApp® FAS Storage Systems. It also provides detailed guidance on how to architect, implement, and manage a large, scalable VMware View solution on FlexPod. The paper details the best integration points for each of the key Cisco and NetApp enabling technologies. It describes how each of these technologies plays a critical and complementary role in providing an integrated solution for VMware View deployments.

This paper is not intended to be a definitive implementation or solutions guide. Additional expertise might be required to solve issues for specific deployments. For more information, contact your local VMware representative and make an appointment to speak with one of our solutions experts.
Resources

Customers can find more information about the products and technologies described in this paper using the links listed below.

**VMware**

**General Information**

- Featured VMware Documentation Sets:
  http://www.vmware.com/support/pubs/
- VMware Licensing Help Center:
  http://www.vmware.com/support/licensing/
- VMware Product Podcasts:
  http://www.vmware.com/technical-resources/podcasts/
- Community, VMware Knowledge Base:
  http://communities.vmware.com/community/vmtn/resources/knowledgebase
- VMware Support Insider:
  http://blogs.vmware.com/kbtv/
- VMware TV:
  http://www.youtube.com/user/vmwaretv
- VMworld TV:
  http://www.youtube.com/user/VMworldTV
- VMware KB TV (external):
  http://www.youtube.com/user/VMwareKB

**VMware View**

- VMware View Architecture Planning Guide (5.1):
- VMware View Installation:
- VMware View Administration:
- VMware View Security:
- VMware View Upgrades:
• VMware View User Profile Migration:

• VMware View Integration:

• Configuration Maximums, VMware vSphere 5.0:

• VMware View 5 with PCoIP, Network Optimization Guide:

• VMware View Persona Management, Deployment Guide:

• Mobile Secure Desktop, Validated Design Guide:

Cisco

• Cisco Unified Computing System:

• Cisco Unified Computing System C-Series Rack Servers:

• Cisco Unified Computing System B-Series Blade Servers:

• Cisco Virtualization Experience Infrastructure CVD for VMware View:

• Cisco UCS Manager Configuration Common Practices and Quick-Start Guide:

• Understanding Cisco Unified Computing System Service Profiles:

NetApp

• NetApp Data Storage Systems:
  www.netapp.com/us/products/storage-systems/

• NetApp FAS3200 Series:
- NetApp TR-3437: Storage Subsystem Resiliency Guide:
- NetApp TR-3749: NetApp Storage Best Practices for VMware vSphere:
- NetApp TR-3884: FlexPod Solutions Guide:
- NetApp FAS/V-Series VASA Provider 1.0 Installation and Administration Guide:
  https://library.netapp.com/ecm/ecm_get_file/ECMP1117738
- NetApp TR-3450: High Availability Pair Controller Configuration Overview and Best Practices:
Acknowledgements

The following individuals contributed to the creation of this deployment guide:

- **Sean Gilbert**, Senior Alliances Technology Manager, VMware
- **Randy Keener**, Staff Technical Operations Architect, VMware
- **Alexander Kobuzatsky**, Desktop Virtualization Technical Marketing Manager, VMware
- **Wen Yu**, Senior Solutions Architect, VMware
- **Jon Catanzano**, Senior Technical Writer/Editor, Consultant