Cisco Virtual Workspace (VXI) Smart Solution 2.7 with VMware View 5.1

Cisco Validated Design

April 2, 2013

Contents

Corporate Headquarters:
Cisco Systems, Inc., 170 West Tasman Drive, San Jose, CA 95134-1706 USA

List of Figures

List of Tables

About Cisco Validated Design (CVD) Program

The CVD program consists of systems and solutions designed, tested, and documented to facilitate faster, more reliable, and more predictable customer deployments. For more information visit http://www.cisco.com/go/designzone.

Cisco Virtual Workspace (VXI) Smart Solution

Introduction

Executive Summary

This Cisco® Validated Design guide provides design considerations and guidelines for deploying an end-to-end Cisco Virtual Workspace (VXI) Smart Solution. This is a service-optimized desktop virtualization system that spans the Cisco Data Center, Cisco Borderless Networks, and Cisco Collaboration architectures to deliver a superior collaboration and superior quality multimedia user experience in a fully integrated, open, and validated desktop virtualization solution. The Cisco Virtual Workspace (VXI) Smart Solution's validated design enables organizations to accelerate the successful implementation of desktop and application virtualization and a unified collaborative workspace. The Cisco Virtual Workspace (VXI) Smart Solution also helps reduce risks and costs and delivers a superior user experience.

This guide discusses the solution as a whole, and describes subsystems in separate chapters. The chapters describe major functional groups such as the data center, network, and endpoints, as well as pervasive system functions such as management and security.

Cisco Virtual Workspace (VXI) Smart Solution Vision

Cisco Virtual Workspace (VXI) Smart Solution delivers the new virtual workspace that unifies virtual desktops, voice, and video, enabling IT to provide exceptionally flexible, secure workspace services with an uncompromised user experience.

Document Audience and Objectives

This guide is intended for use by IT engineers and architects considering the implementation of Cisco Virtual Workspace (VXI) Smart Solution, and anyone who wants to understand the design principles underlying the solution using VMware View. This document provides design considerations and guidelines for deploying an end-to-end solution. Cisco Virtual Workspace (VXI) Smart Solution places the user's computing environment in the data center, allowing it to be accessed through a variety of endpoints, integrating it with collaboration tools, and helping ensure a high-quality user experience.

Cisco Virtual Workspace (VXI) Smart Solution is based on many subsystems, including the virtualized data center, the virtualization-aware network, and the unified workspace; for many customers, many of these subsystems are already deployed. This document focuses on design guidance specifically needed to augment an existing infrastructure for desktop and application virtualization.

This document does not cover all the foundational technologies and reference designs for routing, switching, security, storage, and virtualization. Please see the Cisco Introduction to End to End Desktop Virtualization for more information. It refers to detailed documents that discuss those technologies and reference designs. The Cisco Virtual Workspace (VXI) Smart Solution As-Deployed Reference Guide presents detailed configurations for validated devices. It lists specific software and hardware versions and includes complete device configurations and diagrams showing the topology used in testing.

What is New in Cisco Virtual Workspace (VXI) Smart Solution 2.7

- Integration of Cisco UCS Storage Accelerator with 768GB of disk space for desktop storage optimization

- UCS Central Manager – A manager for multiple instances of UCS Manager
- VMware Hypervisor ESXi5.1 with View 5.1.2
- Cisco WAAS 5.1
- Cisco Virtual EXperience Media Engine (VXME) and Cisco Jabber within the HVD
- TrustSec 3.0 with Nexus1000v
- Cisco BYOD Smart Solution Alignment

Note For detailed information on VMware View 5.1 capabilities, please refer to the appropriate VMware product documentation at http://www.VMware.com/products/view/overview.html

Overview

The Cisco Virtual Workspace (VXI) Smart Solution is an optimized end-to-end infrastructure for desktop virtualization deployments. This system architecture consists of three fundamental building blocks: Cisco Virtualized Data Center, Virtualization-Aware Network, and Unified Collaborative Workspace (Figure 1).

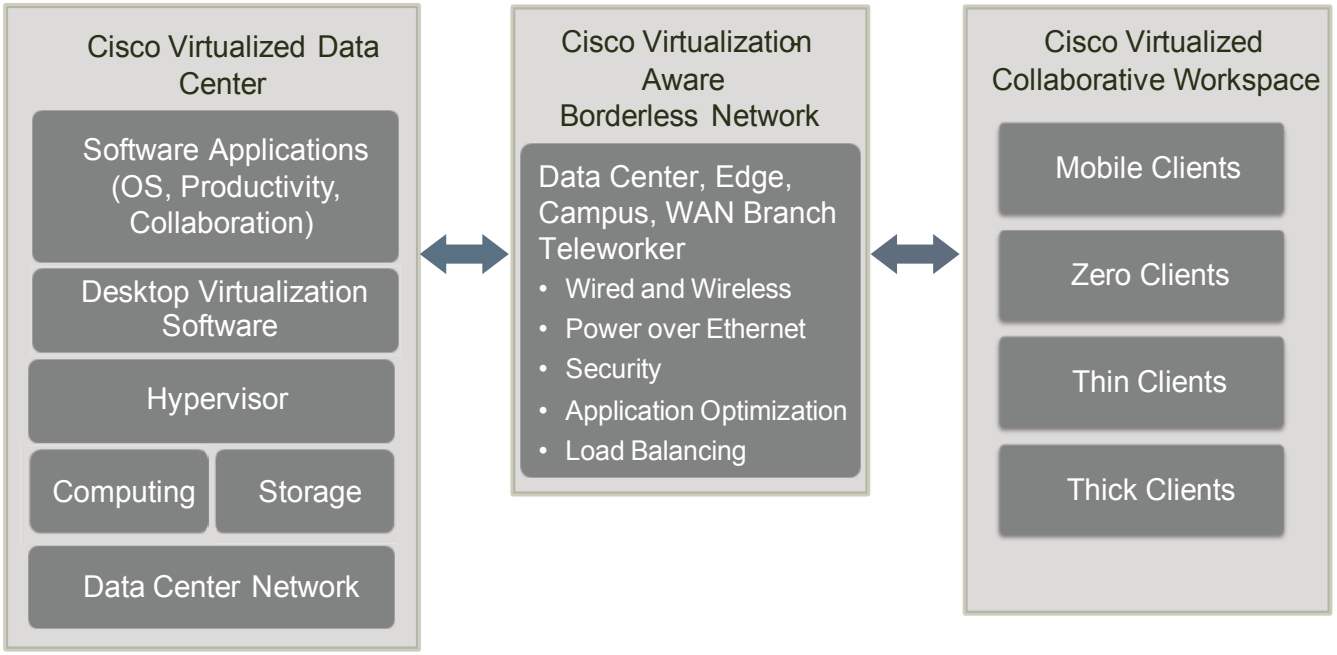
The Cisco Virtualized Data Center is based on the Cisco Unified Data Center architecture, which creates data centers that are efficient, agile, and transformative. Cisco's Virtualized Data Center provides the computing, switching, storage, and virtualization capabilities needed to support a hosted virtual desktop solution from VMware.

The Cisco Virtualization-Aware Network is based on the Cisco Borderless Networks architecture, which reduces operation complexity and provides the services needed to connect anyone, anywhere, on any device to his/her workspace. The Cisco Virtualization-Aware Network connects data centers, enterprise campuses, branch offices, and teleworkers to help ensure that traffic flowing between end users and their hosted desktops is transported securely, reliably, and efficiently. To achieve these goals, the network employs bandwidth optimization, load balancing, quality of service (QoS), security, and other technologies from Cisco's industry-leading portfolio.

The Cisco Unified Workspace builds on the Cisco Collaboration architecture, extending the reach of the virtual desktop to a wide range of end points while supporting critical collaboration capabilities hosted in the data center. End points can be zero clients, thin clients, mobile devices or thick clients, and can include USB-based print and storage capabilities. The Cisco Unified Workspace includes unique capabilities for integrating Cisco Unified Communications endpoints with hosted virtual desktops, including the control of Cisco IP Phones from virtual desktops.

The solution also supports management tools for both Cisco and ecosystem partner products, as well as a rich services portfolio that helps enterprises make the most of their virtualization investments.

Figure 1 Cisco Virtual Workspace (VXI) Smart Solution System Architecture



Cisco Virtual Workspace (VXI) Smart Solution Advantages

This Cisco validated design delivers the following critical advantages:

- **Unified Workspace**: Cisco Virtual Workspace (VXI) Smart Solution supports a comprehensive ecosystem of endpoints that include unified communications and multimedia capabilities. Endpoints supported by the solution include industry-leading capabilities such as Power over Ethernet (PoE), hardware and software form factors, mobility support, and native unified communication media engines. The integration with Cisco's BYOD Smart Solution further extends the reach of the Cisco Virtual Workspace Smart Solution into more mobile clients.
- **Integration with Cisco Unified Communications**: Users can connect to hosted-virtual desktops to make and receive voice or video calls from Cisco Jabber™ which controls the user's desk phone or Cisco VXME Software Client. The control plane is integrated into the user's desktop. The media plane remains outside the virtual desktop display protocol, which enables the network to perform QoS functions such as packet prioritization, call admission control, and path optimization.
- **Simplified Configuration**: The Cisco Unified Computing System™ (Cisco UCS®) integrates the computing, virtualization, hypervisor, fabric-interconnect, and storage functions in the Cisco Virtualized Data Center. Cisco UCS Manager simplifies configuration and improves manageability for all aspects of the Cisco UCS domain. The Cisco Nexus® 1000V Series Switches provide switching, traffic isolation, and policy-insertion capabilities for virtualized environments, extending full networking capabilities to the virtual machine level.
- **Network Optimization**: Cisco Wide Area Application Services (Cisco WAAS) technologies can improve application response times by optimizing bandwidth consumption. For instance, remote print operations can be launched from a user's virtual desktop within the data center to a network printer at a remote branch office. Cisco WAAS can automatically recognize and compress the printing traffic, and spool the resulting print file at the remote location. This capability provides a superior user experience while improving WAN efficiency.

- **Security**: Network connectivity can be controlled at the access layer, using industry-standard IEEE802.1x for port-level authentication. Cisco access switches can thus enforce a security policy at the physical device level and user level by interacting with the credentials-based access control integrated with directory services such as Microsoft Active Directory. Teleworker users, such as mobile users using laptop computers, as well as fixed users, such as home-based teleworkers, can use Cisco's award-wining VPN technology to connect to the enterprise network across the Internet. The user's virtual desktop data is fully protected as it traverses the Internet in an encrypted VPN tunnel. This technology can also be deployed for traffic traversing a managed WAN.
- **End to End Integration and Validation**: The Cisco Virtual Workspace (VXI) Smart Solution has been designed and tested as an integrated whole, and mitigates the system integration investment typically required when deploying desktop virtualization and related technologies. The design guidelines and best practices provided in this document reduce the risks associated with desktop virtualization deployments.
- **Services**: Complementing Cisco solutions, Cisco Desktop Virtualization Services deliver rich, expert-based services end-to-end that can help you rapidly realize a desktop virtualization solution of your choice anywhere, with any device, over any medium. These services also help provide the right fit with your existing investments and align your IT and business strategies. Our services can help you plan, build, and manage a secure desktop virtualization solution. These include:

  Plan

  - Desktop Virtualization Strategy Service: Develop a comprehensive business case and solution strategy for desktop virtualization. Assess operational and mobility services readiness. Create an architecture that may include desktop virtualization, collaboration, and innovation.
  - Desktop Virtualization Assessment Service: Conduct a comprehensive feasibility study and total cost of ownership (TCO) analysis for desktop virtualization.
  - Desktop Virtualization Planning and Design Service: Design a reliable desktop virtualization infrastructure that fits your IT strategy and user requirements.

  Build

  - Desktop Virtualization Pre-Production Pilot Service: Validate specific technical requirements for your proposed desktop virtualization design prior to full production.
  - Desktop Virtualization Implementation Service: Smoothly implement your desktop virtualization solution, including creating an implementation plan and migrating users.

  Manage

  - Desktop Virtualization Optimization Service: Understand the performance and utilization of your desktop environment and evolve your VDI or Cisco Virtual Workspace (VXI) Smart Solution to assure operational excellence as you expand.
  - Cisco Solution Support Service for Cisco Virtual Workspace (VXI) Smart Solution: Rapidly resolve operational issues with solution support that provides a single point of contact.

Market Trends

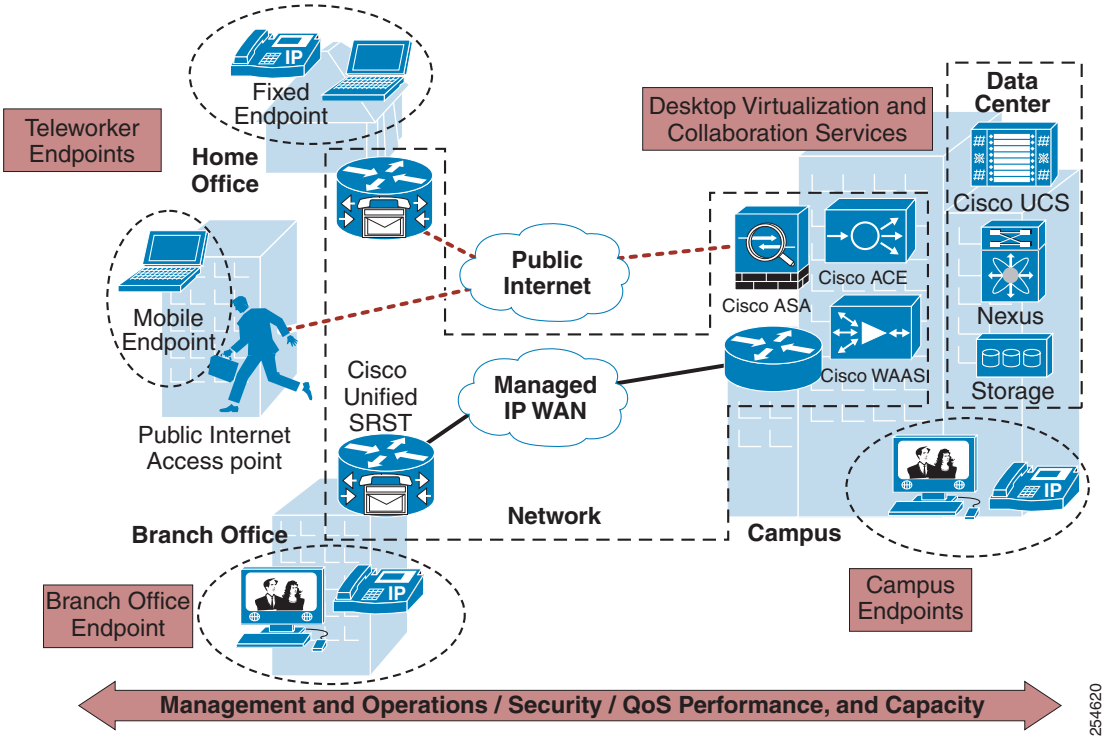
- **Multiple trends are influencing the deployment of desktop virtualization**: IT is seeking more virtualization and savings in the data center through cloud computing, IT is seeking better security and control in remote locations, and end users are seeking multimedia access from multiple devices and locations.

- **Cloud computing and software as a service (SaaS)**: As businesses are moving towards cloud computing and SaaS-based applications, desktop virtualization offers a better way to use computing resources and enable a single end-point to access multiple clouds, while providing the required security and domain isolation and network integrity.
- **Mobility**: Mobile devices such as smartphones and tablets are increasingly being used to access work applications. Desktop virtualization enables users to access work applications without taxing computing resources at endpoints. High-speed connectivity available on smartphones allows a high-quality experience for end-users who access their desktops on smartphones with keyboard and mouse attachments.
- **Globalization**: Off-shoring and expansion into new geographic locations is a prime desktop virtualization deployment scenario. Benefits include data security that facilitates compliance with privacy regulations for data crossing other countries, cost savings through centralized storage and computing resources, and increased productivity through access to data from anywhere at anytime.

Cisco Virtual Workspace (VXI) Smart Solution Deployment Models

A Cisco Virtual Workspace (VXI) Smart Solution system is fundamentally based on the centralization of the user's computing infrastructure in the data center. The computing infrastructure is virtualized and hosts the desktop virtualization and collaboration subsystems. Together, these subsystems provide the user community with desktop and telephony services delivered through the network to the various endpoints. These endpoints can be located on the enterprise campus, in branch, or regional offices, or home offices. Teleworkers are supported with both fixed and mobile clients. Figure 2 shows a typical deployment in which a wide range of clients access virtual desktops across the enterprise network. The term "enterprise network" refers to both the enterprise's own networking infrastructure and its virtual private network (VPN) based extensions into the Internet, and other various forms of externally-managed commercial networks.

Figure 2 Cisco Virtual Workspace (VXI) Smart Solution Top-Down View



Cisco Virtual Workspace (VXI) Smart Solution Validation Goals and Summary

The collective sum of all components and sub-systems that make up the larger Cisco Virtual Workspace (VXI) Smart Solution is designed, deployed and tested in an integrated, end-to-end fashion that is reflective of a customer deployment. Design guidance provided throughout this document is based on this validation. Enterprises can leverage this guidance that includes best practices, caveats, scale, performance and other characterization data to plan and guide their own deployment. This information is primarily in this document but also in the following documents that serve as addendums to this document:

Cisco Virtual Workspace (VXI) Smart Solution Performance and Capacity Validation Results Guide 2.7 for VMware

http://www.cisco.com/en/US/docs/solutions/Enterprise/Data_Center/VXI/CVD/VXI_PCV_V.pdf

Cisco Virtual Workspace (VXI) Smart Solution 2.7 As-Deployed Reference Guide

http://www.cisco.com/en/US/docs/solutions/Enterprise/Data_Center/VXI/configuration/VXI_Config_Guide.pdf

Cisco Virtual Workspace (VXI) Smart Solution 2.7 Release Notes

http://www.cisco.com/en/US/customer/docs/solutions/Enterprise/Data_Center/VXI/VXI_RN_CPE.pdf

The following tables include a high level summary of scalability, performance and other characterization testing performed in Cisco's labs. The testing provides key data needed for sizing a deployment and spans important aspects of the end to end system, specifically Compute, Storage, Collaboration applications, Rich Media and Network. Please refer to the above mentioned Results Guide for the actual results and data.

Compute and Storage Sizing

A series of scalability and performance tests are performed in the Cisco Virtual Workspace (VXI) Smart Solution to provide sizing data for common deployment profiles on different models of Cisco UCS servers. This testing includes both computing and storage aspects of desktop virtualization. Table 1 summarized the validation results.

Note All testing presented here is based on what is called the Cisco Knowledge Worker (KW+) workload. See the Scaling and High Availability chapter for a definition of this workload.

Table 1 Cisco Virtual Workspace (VXI) - Compute and Storage Sizing

Objective	Server Model	Storage	Desktop Virtualization Profile	HVD Profile
Scalability and performance characterization of Cisco UCS B200M3 server with VMware View (Vblock)	Cisco UCS B200 M3 with 384G of memory	VSPEX (EMC VNX 5500) - Fibre Channel	VMware View 5.1 on VMware ESXi5.0U1	Microsoft Windows 7 32-bit with 2 GB of memory and 20 GB disk; Persistent
Scalability and performance characterization of Cisco UCS B230M2 server (Vblock)	Cisco UCS B230 M2 with 256G of memory	VSPEX (EMC VNX 5500) - Fibre Channel	VMware View 5.0 on VMware ESXi 5.0	Microsoft Windows 7 32b with 1.5G of memory and 20G disk; Persistent
Scalability and performance characterization of Cisco UCS B250M2 server (Vblock)	Cisco UCS B250 M2 with 192G of memory	VSPEX (EMC VNX 5500) - Fibre Channel	VMware View 4.6 on VMware ESXi 4.1	Microsoft Windows 7 32b with 1.5G of memory and 20G disk; Persistent
Scalability and performance characterization on Cisco UCS B250 M2- Impact of success criteria (CPU utilization counter changed) and vSphere 5.1 changes (Vblock)	Cisco UCS B-250 M2 with 192G of memory	VSPEX (EMC VNX 5500) - Fibre Channel	VMware View 5.0 on VMware ESXi 5.0 RDP & PCoIP	Microsoft Windows 7 32b with 1.5G of memory and 20G disk; Persistent
Storage Optimization with VMware's View Storage Accelerator	Cisco UCS B200 M3 with 384G of memory	VSPEX (EMC VNX 5500) - Fibre Channel	VMware View 5.1 on VMware ESXi5.0U1	Microsoft Windows 7 32b with 1.5G of memory and 20G disk; Persistent

Application Characterization

The goal of application characterization is to characterize the performance of the application as a standalone application in the workload, running across all user desktops deployed on a server. An Enterprise looking to deploy a Cisco Collaboration application for their virtual desktop users can use the performance data to understand the incremental impact of that application as more and more users start using them concurrently. The impact is to the server resources which can change the number of users that can be deployed on that same server once the new application is rolled out. Table 2 summarizes the application characterization results.

Table 2 Cisco Virtual Workspace (VXI) - Application Characterization

Objective	Server Model	Storage	Desktop Virtualization Profile	HVD Profile
Scale and Performance characterization of Cisco Jabber for Windows with VMware View	Cisco UCS B200 M3 with 384 GB of memory	VSPEX (EMC VNX Series)	VMware View 5.1 on ESXi 5.1	Microsoft Windows 7 32-bit with 2 GB of memory
Scale and Performance characterization of Cisco Contact Center - CTIOS Agent	Cisco UCS B230 M2 with 256 GB of memory	NFS on NetApp FAS 3170	N/A - See test profile for more detail.	Microsoft Windows 7 32-bit with 2 GB of memory

Network Characterization

The goal of network characterization is to characterize different network aspects of desktop virtualization, including network services, optimizations, and other data needed for a successful deployment. A summary of these results, tested across a Cisco Virtual Workspace (VXI) Smart Solution, is shown in Table 3

Table 3 Cisco Virtual Workspace (VXI) - Network Characterization

Objective	Server Model	Wan Link	Desktop Virtualization Profile	HVD Profile
Understanding the bandwidth (BW) characteristics of a Cisco KW+ workload	Cisco UCS B200 M2 with 96G memory	T1 with 80ms Latency	VMware View 4.5 on VMware ESXi 4.1; PCoIP and RDP	Microsoft Windows 7 32b with 1.5G of memory and 20G disk; Persistent
Understanding the bandwidth characteristics of a video-only workload	Cisco UCS B200 M2 with 96G memory	T1 with 80ms Latency	VMware View 4.5 on VMware ESXi 4.1; PCoIP and RDP	Microsoft Windows 7 32b with 1.5G of memory and 20G disk; Persistent

Objective	Server Model	Wan Link	Desktop Virtualization Profile	HVD Profile
Impact of display protocol adaptiveness on server/compute performance at scale	Cisco UCS B200 M2 with 96G memory	T1 with 80ms Latency	VMware View 4.5 on VMware ESXi 4.1; PCoIP and RDP	Microsoft Windows 7 32b with 1.5G of memory and 20G disk; Persistent
Impact of Cisco WAAS Optimization on WAN deployments with VMware View RDP	Cisco UCS B200 M2 with 96G memory	T1 with 80ms Latency	VMware View 4.5 on VMware ESXi 4.1; PCoIP and RDP	Microsoft Windows 7 32b with 1.5G of memory and 20G disk; Persistent

Virtualized Data Center

Overview

The Cisco® Data Center unifies processor and storage resources, hosts virtual machines and desktops, and provides a network switching fabric that interconnects these resources to the rest of the enterprise network. The Cisco Virtualized Data Center is based on Cisco Unified Data Center architecture, which creates data centers that are efficient, agile, and transformative. It helps enable enterprises to consolidate data center infrastructure, reduce energy costs, improve workforce productivity, and ensure business continuity. The Cisco data center is based on three pillars of innovation:

- **Cisco Unified Fabric**: Cisco uses an innovative fabric-based architecture to unify computing, networking, storage, virtualization, and management into a single data center platform. This approach helps to ensure consistent delivery of highly-available and secure IT services.
- **Cisco Unified Computing**: The Cisco Unified Computing System™ (Cisco UCS®) combines industry-standard blade and rack servers, networking, and management into a single, centrally-managed, and massively scalable system. Infrastructure can be automatically provisioned to speed deployment of enterprise applications.
- **Cisco Unified Management**: To simplify data center management and support delivery of virtual desktops, Cisco offers centralized management of physical, virtual, and cloud-based resources. Cisco Unified Management solutions enable automatic provisioning, policy-based management, creation of an IT service catalog and self-service user portal, and pay-per-use tracking.

The Cisco Virtualized Data Center is highly scalable, and can be adopted in an incremental, granular fashion to help ensure a graceful evolution in response to enterprise needs. The virtualized data center is designed to provide optimum levels of performance, scalability, availability, and security. This section provides design guidance and best practices for achieving those goals in a Cisco deployment. The main topics include --

- Virtualized data center architecture
- Compute design and best practices
- Virtualization design and best practices

- Storage design and best practices
- Data center network design and best practices
- Management tools for the virtualized data center

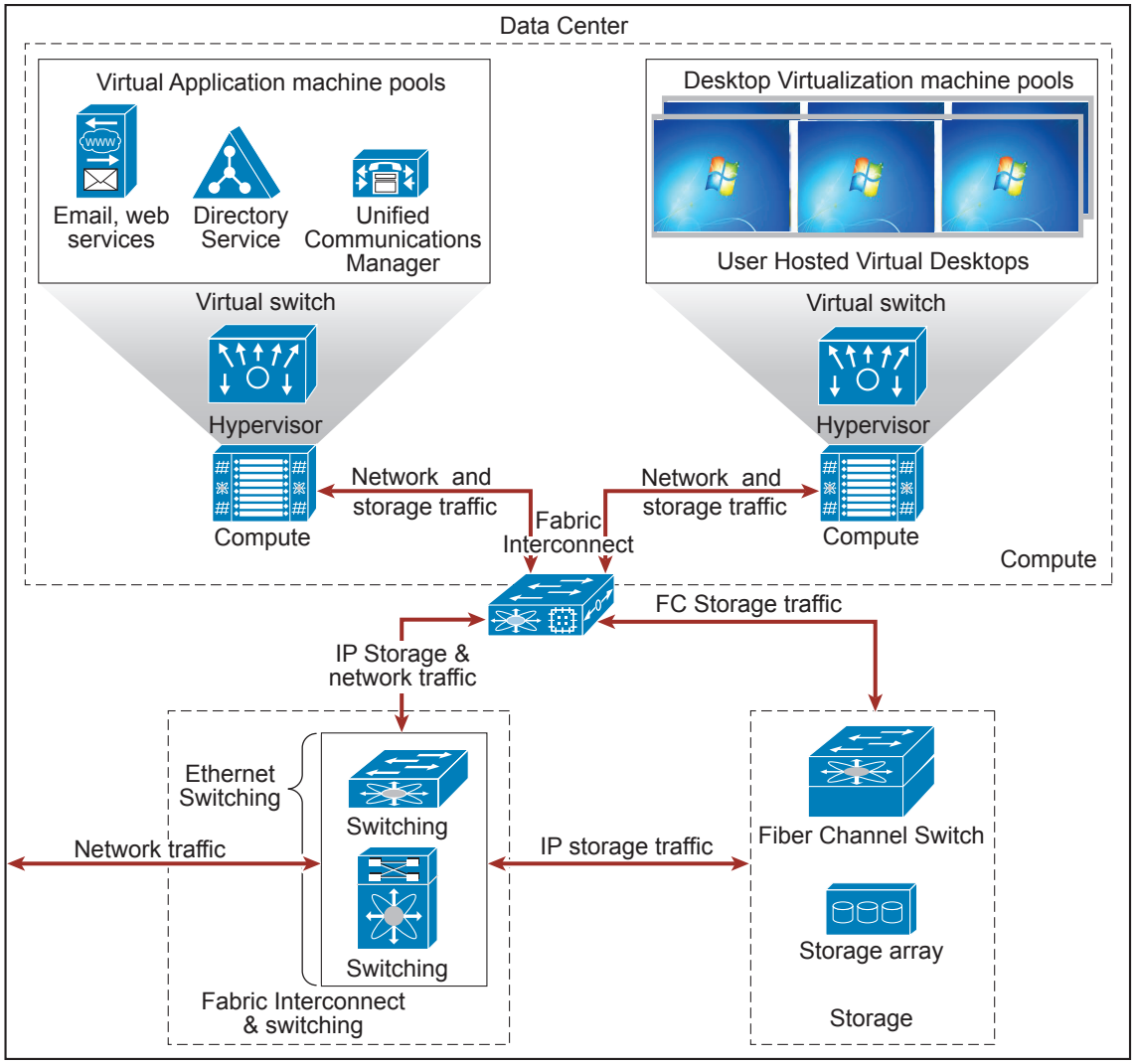
What is New in Release 2.7

Cisco UCS Storage Accelerator - a PCIe Flash-based caching solution that resides on the Cisco UCS B200 M3 blade server. It supports extremely high numbers of IO requests and enables organizations to significantly reduce storage costs.

Virtualized Data Center Architecture

The Cisco Virtualized Data Center architecture consists of the computing, virtualization, storage, and networking subsystems needed to deliver an effective desktop virtualization solution. This structured approach to data center design provides outstanding flexibility as needs change, while helping ensure world-class performance, security, and resilience. The architecture tightly integrates Cisco and partner products, services, and best practices to provide a simplified, secure, and scalable solution.

Figure 3 Architecture



Note The solution also supports the Vblock™, VSPEX™ and FlexPod™ prepackaged infrastructure platforms. See http://www.vce.com/solutions/ for more information on the Vblock architecture. See www.cisco.com/go/vspex for more information on VSPEX. See http://www.netapp.com/us/technology/flexpod/ and http://www.cisco.com/ for more information on FlexPod.

Computing Design and Best Practices

The Compute subsystem is based on Cisco UCS components. The system combines Cisco UCS B-Series Blade Servers and C-Series Rack Servers with networking and storage access in a single converged system that simplifies management and delivers greater cost efficiency and agility with increased visibility and control. The Cisco UCS B-Series and C-Series servers support Cisco Unified Fabric, which connects computing, LAN, and storage networks through a single medium. Cisco UCS servers are designed to reduce energy consumption, with highly efficient power supplies and Intel Xeon processors

that match power consumption with workloads. Each server contains the processor, RAM, and I/O resources needed to support a virtual desktop environment. Cisco UCS servers are managed by Cisco UCS Manager, which implements role-based and policy-based management using service profiles and templates. This section discusses the following important elements of the data center design:

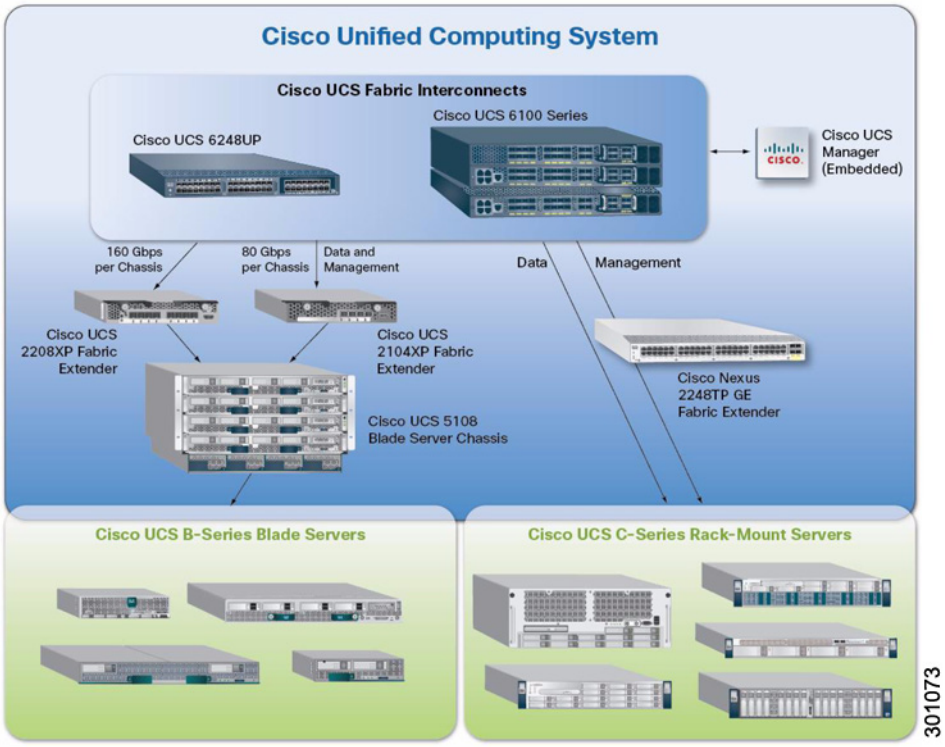
- Compute subsystem building blocks
- Server selection and validation
- Connectivity for Cisco UCS servers
- Fabric Interconnects
- Server Topology
- Server BIOS Settings

Computing Subsystem Building Blocks

- **Cisco UCS B-Series Blade Servers**: Cisco UCS B Series servers are based on Intel Xeon processors and offer exceptional memory capacity. Each blade server's front panel provides direct access to video connections, USB ports, and console connections. Cisco has validated several models of the Cisco UCS B-Series Blade Server, and in this phase it continues validating the Cisco UCS B200 M3. The blade servers connect to the chassis by means of converged network adapter (CNA) cards, such as the Cisco VIC 1240 and 1280. The UCS B200 M3 blade server now supports the Cisco UCS Storage Accelerator, an on-blade Flash-based caching solution that offloads IOPS processing and reduces storage costs. Refer to the Storage Design and Best Practices section of this chapter for more information.
- **Cisco UCS 5100 Series Blade Server Chassis**: The chassis provides an enclosure for Cisco UCS B-Series Blade Servers. It is six rack units (6RU) high, can mount in an industry-standard 19-inch rack, and uses standard front-to-back cooling. Each chassis can accommodate up to eight half-width or four full-width Cisco UCS B-Series Blade Servers. The chassis also supports up to four single-phase, hot-swappable power supplies.
- **Cisco UCS C-Series Rack Servers**: The Cisco UCS C-Series extends Cisco UCS innovations to an industry-standard rack-mount form factor. The Cisco UCS C-Series servers can operate both in standalone environments and as part of the Cisco Unified Computing System. The Cisco UCS C-Series servers can be deployed incrementally according to an organization's timing and budget. Cisco UCS C-Series servers interface with the Cisco Unified Communications System through network adapters such as the Cisco UCS P81E VIC. This card is a dual-port 10 Gigabit Ethernet PCI Express (PCIe) adapter that provides dynamically configurable virtual interfaces.
- **Cisco fabric extenders**: The Cisco UCS 2100 and 2200 Series Fabric Extenders reside in the Cisco UCS 5100 Series Blade Server Chassis and provide 10 Gigabit Ethernet connections between servers and fabric interconnects. The fabric extenders function as distributed line cards and are managed as extensions of the fabric interconnects. The Cisco Nexus® 2000 Series Fabric Extenders connect rack-mount servers to the fabric interconnects. Like the Cisco UCS fabric extenders, the Cisco Nexus fabric extenders function as line cards for the parent switch.
- **Cisco UCS 6100 and 6200 Series Fabric Interconnects**: Typically deployed in pairs to provide highly available network connectivity and management capabilities for Cisco UCS, the fabric interconnects offer line-rate, low-latency, lossless 10 Gigabit Ethernet, Fibre Channel over Ethernet (FCoE), and Fibre Channel functions. All chassis and servers attached to the fabric interconnects become part of a single, highly available management domain. The fabric interconnects provide connectivity between Cisco UCS and the rest of the data center network.

- **Cisco UCS Manager**: Cisco UCS Manager provides embedded management of all software and hardware components of Cisco UCS across multiple chassis and rack-mount servers and thousands of virtual machines. It manages the system as a single entity through an intuitive GUI, a command-line interface (CLI), or an XML API. Cisco UCS Manager is embedded on a pair of Cisco UCS 6100 or 6200 Series Fabric Interconnects using a clustered, active-standby configuration for high availability. The manager participates not only in server provisioning, but also in device discovery, inventory, configuration, diagnostics, monitoring, fault detection, auditing, and statistics collection.

Figure 4 Cisco Unified Computing System



Connectivity for Cisco UCS Servers

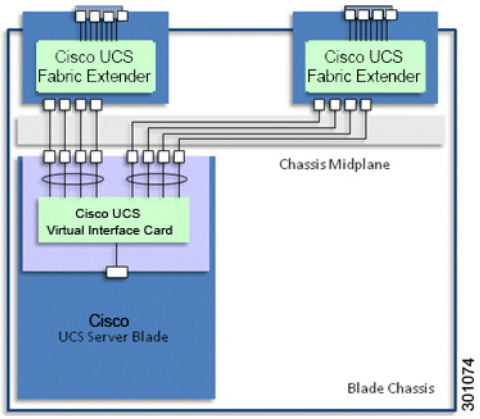
Cisco UCS B-Series blade servers connect to the chassis mid-plane by means of Virtual Interface Cards, installed as mezzanine cards. These cards provide virtual Ethernet NICs or Host Bus Adapters (HBAs) to the server operating system, and high speed FCoE ports for connecting to Cisco UCS 2x00 Series Fabric Extenders located on the back of each chassis. The Fabric Extenders aggregate traffic from the interface cards and pass this traffic to an upstream pair of Cisco UCS Series 6x00 Fabric Interconnects (not shown).

Cisco UCS C-Series rack mount servers provide PCIe slots for network adapters (such as the VIC). These network adapters connect to Cisco Nexus 2200 Series Fabric Extenders, which in turn connect to the Fabric Interconnects. Alternately, C-Series servers can be connected directly to a switch such as the Cisco Nexus 5500 Series.

Cisco UCS Virtual Interface Cards can support advanced technologies such as VM-FEX, which implements Cisco VN-Link in hardware. VM-FEX collapses physical and virtual switching layers, reduces the number of management points, and extends the network all the way to the virtual machine. VM-FEX is especially appropriate in environments where hardware-based performance is more critical than high virtual machine density or virtualized service availability.

The solution validated the VIC as a standard non-VM-FEX adapter, using Cisco's software-based implementation of VN-Link, the Cisco Nexus 1000V. The Cisco Nexus distributed virtual switch scales to very high densities, and supports virtualized services such as the Cisco Virtual Security Gateway, as well as virtual firewalling and load-balancing. The Cisco Nexus 1000V is discussed in greater detail in the Data Center Networking portion of this chapter.

Figure 5 Cisco UCS B-Series Blade Server Connectivity



Deploying Cisco UCS Fabric Interconnects in Cisco Virtual Workspace (VXI) Smart Solution

The Cisco UCS 6100 and 6200 Series Fabric Interconnects are top-of-rack (ToR) controllers that provide network connectivity and management for Cisco UCS servers (Figure 6). Cisco fabric interconnects also function as parent switches for the fabric extenders, which act as distributed line cards. These fabric interconnects offer line-rate, low-latency, lossless 10 Gigabit Ethernet, Fibre Channel, and FCoE capabilities. The Cisco UCS 6100 and 6200 Series deliver high-performance unified fabric, centralized unified management with Cisco UCS Manager, and virtual machine–optimized services with support for Cisco VN-Link technology.

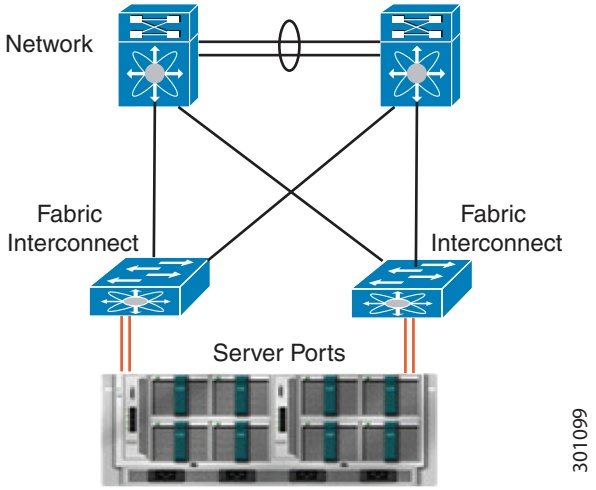
Fabric interconnects terminate FCoE traffic flows from the fabric extenders. Ethernet traffic is separated and forwarded to the data center switch fabric (composed of Cisco Nexus 5000 and 7000 Series Switches). Fibre Channel traffic is forwarded using Fibre Channel uplinks to the Fibre Channel SAN. The latest generation of fabric interconnects offers support for unified ports, which can be configured as either Ethernet or Fibre Channel ports.

Embedded Cisco UCS Manager facilitates management of the entire Cisco UCS domain. Using Cisco UCS Manager in combination with desktop virtualization management software, administrators can deploy virtual machines, perform software upgrades, migrate virtual machines between physical servers, and extend computing resource control over thousands of virtual desktops. For virtual environments, fabric interconnect design considerations include:

- **Paired deployment**: By design the fabric interconnects are deployed in redundant pairs to provide uniform reliable access to both network and storage resources. A virtual IP address is created to link the two switches and provide a single point of management.

- **Operating mode**: The fabric interconnect operates in either of two modes: switch mode or end-host mode. In switch mode, it acts as a normal Layer 2 switch with spanning tree enabled (which results in the blocking of upstream links). In end-host mode, server virtual NICs (vNICs) are pinned to a specific uplink port with no uplink switching (which obviates the need for spanning tree, and enables active-active uplinks). End-host mode, which is the default mode, provides the most predictable results for desktop virtualization environments, and therefore is preferred for deployments.
- **Static or dynamic pinning**: Server ports can be dynamically pinned to a given uplink or Port Channel, or they can be statically pinned. With static pinning, specific pin groups are created and associated with an adapter. Static pinning provides deterministic traffic flows and enables traffic management if desired, and is the method used in testing.

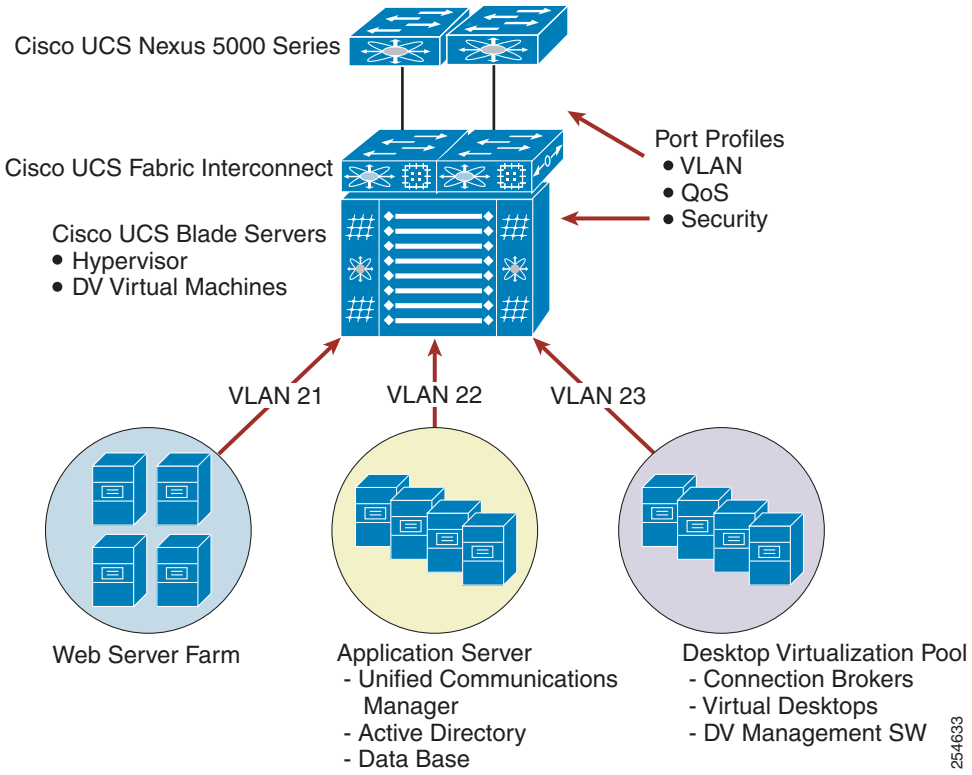
Figure 6 Cisco UCS 6x00 Fabric Interconnects



Server Topology Recommendations

The validation process segregated infrastructure-oriented servers from servers that host virtual desktops. Infrastructure servers are those that provide critical services to a broad user population, such as application servers, license servers, unified communications, and Microsoft Active Directory. Grouping physical servers according to function and load helps ensure that resources are used more efficiently and protects infrastructure services from potential aggregate load effects and sudden surges. This approach also may insulate desktop users from demand peaks that might be associated with infrastructure servers. Figure 7 shows a deployment in which servers are grouped by function.

Figure 7 Grouping Servers by Function



Determining Cisco UCS BIOS Settings for Desktop Virtualization

Cisco UCS BIOS settings define basic server behavior during bootup. Some of these BIOS settings can directly affect system performance, especially in virtualized environments. Many of the most important parameters can be configured through Cisco UCS Manager. Table 4 shows the main BIOS parameters and the settings used for validation.

Table 4 Cisco UCS BIOS Settings

BIOS Parameter	Setting
CPU Performance	Enterprise
Direct Cache Access	Enabled
Intel Speedstep	Enabled
Hyper-Threading	Enabled
Turbo-Boost	Enabled
Processor C3 Report	Enabled
Processor C6 Report	Enabled
Memory reliability, availability, and serviceability (RAS)	Maximum Performance

BIOS Parameter	Setting
Low-voltage double-data-rate (LV DDR) mode	Power Saving Mode
Non-uniform memory access (NUMA) optimization	Enabled

Note Testing was conducted with Intel SpeedStep enabled. This feature saves power when a host is running a reduced load. Some organizations disable Intel SpeedStep for virtual machines that experience poor performance. This condition typically occurs on lightly loaded servers running Intel EX processors, when CPU use is lower. Rather than disabling Intel SpeedStep, the recommended solution is to disable the C1E halt state in the BIOS.

Virtualization Design and Best Practices

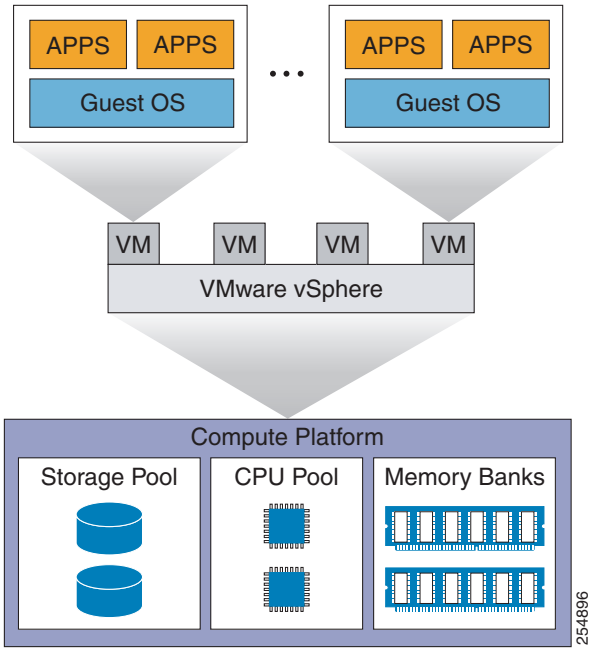
The Virtualization subsystem includes the various solutions required to virtualize servers, desktops, users, and applications in the system. The Virtualization subsystem includes the following elements:

- **Hypervisor**: The hypervisor abstracts physical processor, memory, storage, and networking resources into virtual machines. Hypervisors enable virtual machines to be created, managed, optimized, and migrated. This Cisco Validated Design is based on the VMware vSphere environment.
- **Desktop Virtualization**: Desktop virtualization enables creation of virtual desktops on virtual machines. Applications and operating systems are hosted on these desktops and accessed by remote users. This design guide is based on the VMware View solution for desktop virtualization.
- **Application Virtualization**: Application virtualization permits packaging and on-demand delivery of applications to desktop users from a central repository. This design validates user virtualization solutions from VMware.

Hypervisor

VMware vSphere is a suite of products that provides virtualization, management, resource optimization, availability, and optimization capabilities. The VMware vSphere (ESXi) hypervisor is the primary virtualization component of vSphere. The hypervisor abstracts the processor, memory, storage, and networking resources of its physical server into multiple virtual machines, and ensures that each virtual machine receives its appropriate share of these resources. The hypervisor is installed on each Cisco UCS server to allow virtualized desktops and servers to run as independent virtual machines (VMs).

Figure 8 VMware vSphere Hypervisor



The hypervisor bridges virtual and physical realms by interfacing virtual desktops with Virtual Interface Cards on the blade servers. The Virtual Interface Cards provide the server (and thus the hypervisor) access to Ethernet NICs and Fibre Channel host bus adapters (HBAs). Each card combines the Ethernet and Fibre Channel traffic onto a unified fabric in the form of Fibre Channel over Ethernet (FCoE) traffic. The hypervisor maps the physical network interface cards (NICs) to virtual NICs as part of the virtual switch. Physical Host Bus Adapters (HBAs) for Fibre Channel traffic are bridged directly to virtual HBAs.

Hypervisors include advanced tools for managing servers and associated virtual machines, in high density production environments. These tools enable IT administrators to identify over-committed host machines, move VMs among pooled hosts, manage power up sequences, consolidate active VMs on the fewest possible hosts, and more.

Many factors influence the successful deployment of a hypervisor, and a detailed discussion is beyond the scope of this document. Key questions to consider when designing a hypervisor implementation environment are addressed in the Table 5 below.

Table 5 Main Questions When Designing a Hypervisor Installation for Cisco Virtual Workspace (VXI) Smart Solution

Design Issue	Description	Cisco Approach
Virtual machine topology definition	Physical and virtual resources can be grouped according to functions.	Management and infrastructure server applications are installed on dedicated Cisco UCS blades. Dedicated database servers host the vCenter and VMware View databases. The vSphere Enterprise Plus License is deployed on the vCenter server.
Virtual resource provisioning	Virtualization enables a certain degree of over-provisioning, but it's possible to reach a point of diminishing returns. Ex: provisioning a virtual machine with more vCPUs than it can use will increase resource consumption and may actually reduce performance under some circumstances.	Virtualization enables a certain degree of over-provisioning, but too much can have diminishing returns. For example, provisioning a virtual machine with more virtual CPUs (vCPUs) than it can use will increase resource consumption and may reduce performance under certain circumstances.
Storage Types	Storage can be physically located on the Cisco UCS blade server or network attached. To support virtual machine migration, clustering, distributed resource scheduling and other advanced virtualization features, shared storage is typically deployed. User preferences vary widely, and some deployments may use hybrid shared/local solutions.	The system includes support for shared storage solutions (NAS and SAN) from multiple ecosystem partners. Cisco Virtual Workspace also has validated hybrid storage solutions that blend both local and centralized share resources.
Virtual Machine Boot Source	Booting from centralized shared storage allows for off-the-shelf server replacements as well as centralized configuration management of the ESX/VMware ESXi™ software.	The Cisco Validated Design is configured to Boot from SAN, as recommended by the hypervisor vendor.

Design Issue	Description	Cisco Approach
Disable unused physical devices	Hypervisor vendors typically recommend that unused devices such as COM, USB, and LPT1 ports, optical drives, and others be disabled. These devices consume resources even when not used. Some devices are polled even though inactive, or reserve blocks of memory that won't be used.	Cisco recommends disabling unused ports and drives to ensure that these do not consume resources.
Advanced hypervisor configuration parameters (VMware)	Hypervisors may provide the ability to fine-tune CPU and memory resources in virtual machines, which can increase virtual machine density per server. In early versions of vSphere, for example, tuning CPU fairness algorithms or memory reclamation features was a manual process, to be done only with the guidance of technical support. Newer (VMware vSphere 5.0 and later) versions have default values optimized for large deployments.	Cisco recommends deploying the most recent versions of these hypervisors to leverage optimized default settings.
System synchronization	System-wide synchronization is critical to ensuring the accuracy of all logged information.	Cisco recommends synchronizing the hypervisor clock to an external NTP server, then synchronizing the virtual desktop OS clocks to the hypervisor clock.
High availability	Hosts can be clustered to provide fault tolerance and high availability.	A best practice is to distribute hosts across multiple Cisco UCS chassis to provide better availability in the event of a chassis failure; at the same time, all hosts are grouped within a single High Availability cluster for Distributed Resource Scheduling and vMotion.

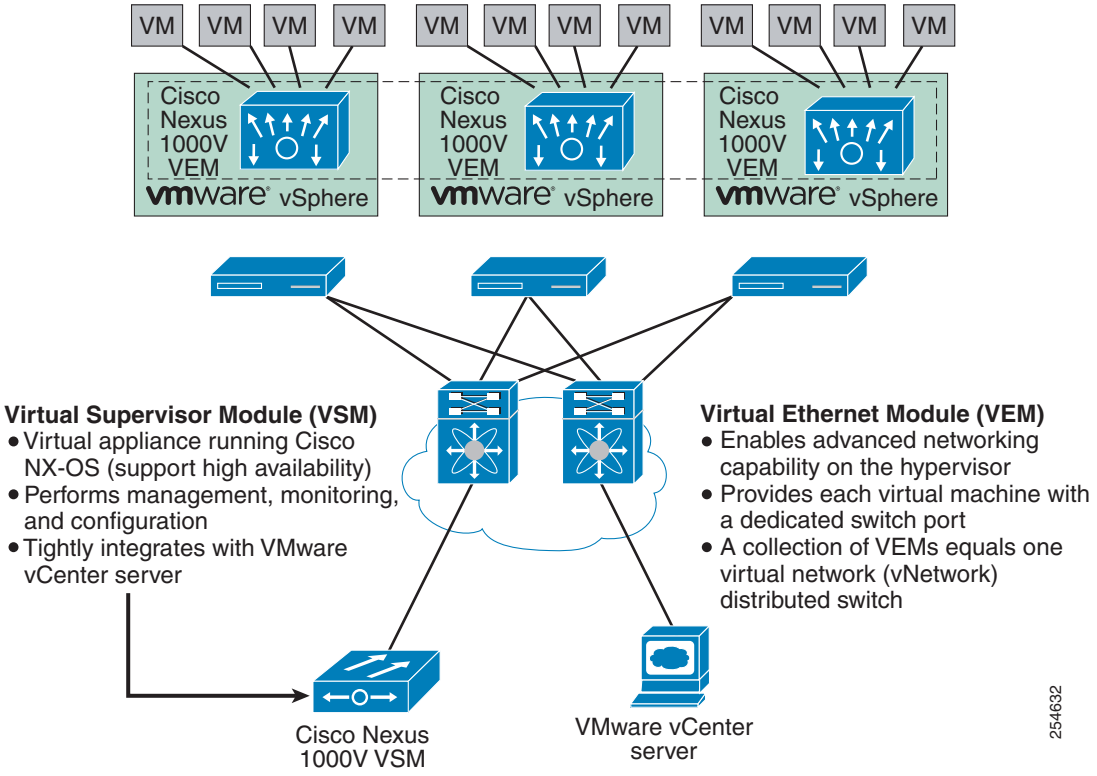
Virtual Switching: The Cisco Nexus 1000v Switch

The Cisco Nexus 1000V Switch is a virtual machine access switch that runs inside a hypervisor. The switch uses Cisco VN-Link server virtualization technology to deliver policy-based virtual machine connectivity, mobile virtual machine security and network policy, and a nondisruptive operation model. The Cisco Nexus 1000V provides administrators with a consistent networking feature set and provisioning process from the virtual machine access layer to the data center network core. Virtual servers can use the same network configuration, security policy, diagnostic tools, and operation models as their physical server counterparts.

A single Cisco Nexus 1000V Switch can encompass several physical servers. Targeted port profiles can be created for the specific requirements associated with each type of user and virtual desktop. Cisco Nexus 1000V profiles contain information such as VLAN assignment, quality-of-service (QoS) policies, and security access control lists (ACLs). The port profile is linked to the virtual machine profile, so that if the hypervisor migrates to a particular virtual desktop, the associated profile also migrates. Refer to the Securing Cisco Virtual Workspace chapter of this guide for more information on creating and deploying security zones.

Troubleshooting of connectivity problems is enhanced through the built-in Cisco Switched Port Analyzer (SPAN). Increased security is implemented by the use of several additional features such as VLANs, private VLANs, port security, and security ACLs. The Cisco Nexus 1000V also provides a foundation for other virtual networking solutions such as the Cisco VSG and Cisco Virtual Wide Area Application Services (Cisco vWAAS) . The Cisco Nexus 1000V is currently supported on VMware vSphere hypervisors with Enterprise Plus licenses (Figure 9).

Figure 9 VMware ESX/ESXi and Cisco Nexus 1000V Series Integration



Virtual Machine Based Security Zones: The Cisco Virtual Security Gateway

Cisco VSG for Cisco Nexus 1000V Series Switches is a virtual appliance that controls and monitors access to trust zones in enterprise and cloud provider environments (Figure 10). Cisco VSG provides secure segmentation of virtualized data center virtual machines using detailed, zone-based control and monitoring with context-aware security policies. Controls are applied across organizational zones, lines of business, or multitenant environments. Context-based access logs are generated with network and virtual machine activity levels. Trust zones and security templates can be provisioned on demand as virtual machines are created.

Cisco VSG employs Cisco Virtual Network Service Data Path (vPath) technology embedded in the Cisco Nexus 1000V Series Virtual Ethernet Module (VEM). Cisco vPath steers traffic to the designated Cisco VSG for initial policy evaluation and enforcement. Subsequent policy enforcement is off-loaded directly to Cisco vPath. Cisco VSG can provide protection across multiple physical servers. It can be transparently inserted in one-arm mode, and it offers an active-standby mode for high availability.

Cisco VSG can be deployed across multiple virtual machine zones and virtualized applications. It requires a virtual 1.5-GHz CPU, 2 GB of RAM, a 3-GB hard drive, and three network interfaces (data, management, and high availability). It also requires VMware vSphere and vCenter Release 4.0 or later, Cisco Nexus 1000V VEMs and VSMs, and Cisco Virtual Network Management Center (VNMC).

Virtual Machine Based Network Optimization: Cisco vWAAS

Cisco vWAAS is a virtual appliance that accelerates business applications delivered from private and virtual private clouds. Cisco vWAAS runs on Cisco UCS servers and the VMware vSphere hypervisor, using policy-based configuration in the Cisco Nexus 1000V Switch. Cisco vWAAS can be associated with application server virtual machines as these are instantiated or moved. With Cisco vWAAS, cloud providers can rapidly provision WAN optimization services with little to no configuration or disruption.

For more information on Cisco vWAAS, see the Virtualization Aware Network and Scaling and High Availability chapters, in this guide.

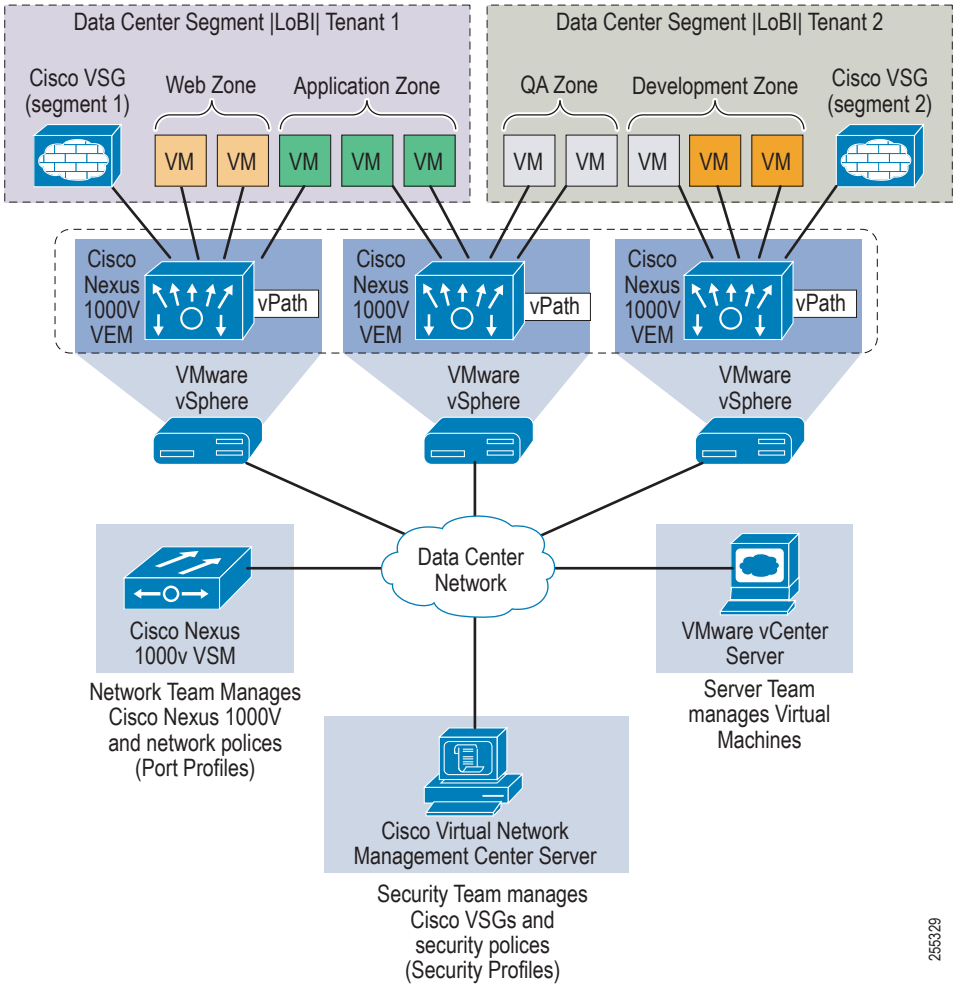
Desktop Virtualization

This Cisco Validated Design guide describes the implementation of the solution with VMware's View solution for desktop virtualization. With VMware View, organizations can create and run virtual desktops in the data center, and deliver these desktops to employees using a wide range of endpoints. Users access their virtual desktops from zero, thin, or thick clients by means of a remote display protocol (PCoIP or RDP). Figure 11 displays a typical desktop virtualization implementation based on VMware View. Key components of the VMware View solution include:

- View Agent
- View Client
- View Connection Server
- View Composer
- View Transfer Server
- View Agent

VMware View Agent on each HVD in the pool is required to create the connection between the client and HVD. The features and policies on VMware View Agent can be controlled via Active Directory and/or View Connection Server settings. The agent also provides features such as connection monitoring, virtual printing, and access to locally connected USB devices. To install the agent automatically on all HVDs in a pool, install the VMware View Agent service on a virtual machine and then use the virtual machine as a template or as a parent of linked clones. The capability set of the agent is tied to the operating system. Please consult the appropriate VMware documentation for compatibility information.

Figure 10 Cisco Virtual Security Gateway



View Client

VMware View Client is installed on each endpoint that needs to access its HVD. View Client supports PC over IP® and Microsoft Remote Desktop Protocol (RDP). VMware View Client supports Local Mode operations, in which users can download a View desktop to a local system and operate with or without a network connection. View desktops in local mode function the same way as equivalent remote desktops, but are also able to take advantage of local system resources such as memory and CPU. This capability may reduce latency and enhance performance for desktop users. If a network connection is available, the local mode desktop continues to communicate with the View Connection Server for updates. Users can access their checked-out desktop directly, and can log off and on without going through the Connection Server. The local desktop can be checked back in at the user's convenience. Local mode requires View Transfer Server in the data center.

View Connection Server

This software service acts as a broker for client connections. VMware View Connection Server authenticates users via Active Directory and post-authentication directs the user to an appropriate HVD. Some other important functions performed by View Connection Server are desktop entitlement for users, desktop session management, establishing secure connections between users and desktops, policy application, and single sign-on. View Administrator is a user interface that comes prepackaged with the

View Connection Server and provides an administrative interface for management. The sizing guidelines for View Connection Server can vary based on display protocol type, virtual machine resources, use of encryption, and choice of tunneled or non tunneled mode. Implementing load balancing for View Connection Servers with solutions such as the Cisco ACE Application Control Engine Module is highly recommended.

VMware recommends that administrators use dedicated View Connection Servers to handle clients in kiosk mode, and use dedicated groups in Active Directory for their accounts. To configure kiosk mode, refer to the VMware View Administrator's Guide.

View Composer

View Composer is an important VMware View component that allows for storage optimization. In virtual desktop environments, data redundancy per HVD is very high since typically the same OS and application sets are replicated across the virtual desktop pool. To deal with this, View Composer creates a pool of linked clones from a specified parent virtual machine. Each linked clone acts like an independent desktop, with a unique host name and IP address, yet the linked clone requires significantly less storage because it shares a base image with the parent.

View Composer can create images that share the base OS image while still keeping the user profile data separate. It is highly recommended to separate the OS files from user profiles in the storage array. Disk space requirements when using View Composer can be reduced by more than 50 percent.

View Composer is often set up on the same virtual machine as vCenter to allow control of the VMware ESX® hosts. However, View Composer also can be installed on a separate VM to support larger deployments. Each View Composer server in a cluster can handle up to 1000 VMs per pool; in a large deployment, clustering multiple View Composer instances may be required.

Note If the company policy allows users to install custom applications, the full benefits of View Composer can't be realized. It is highly recommended to separate such user profiles and place these HVDs on a backed up data storage system.

View Transfer Server

This server manages transfers between the data center and View desktops that have been checked out for use on local systems. It is required for desktops that run View Client with Local Mode. The View Transfer Server also ensures that changes generated on the local desktop are propagated to the appropriate desktop in the data center. It keeps local desktops current by distributing common system data from the data center to the local clients. If a local computer is corrupted or lost, View Transfer Server can provision the local desktop and recover user data by downloading the data and system image.

Deployment

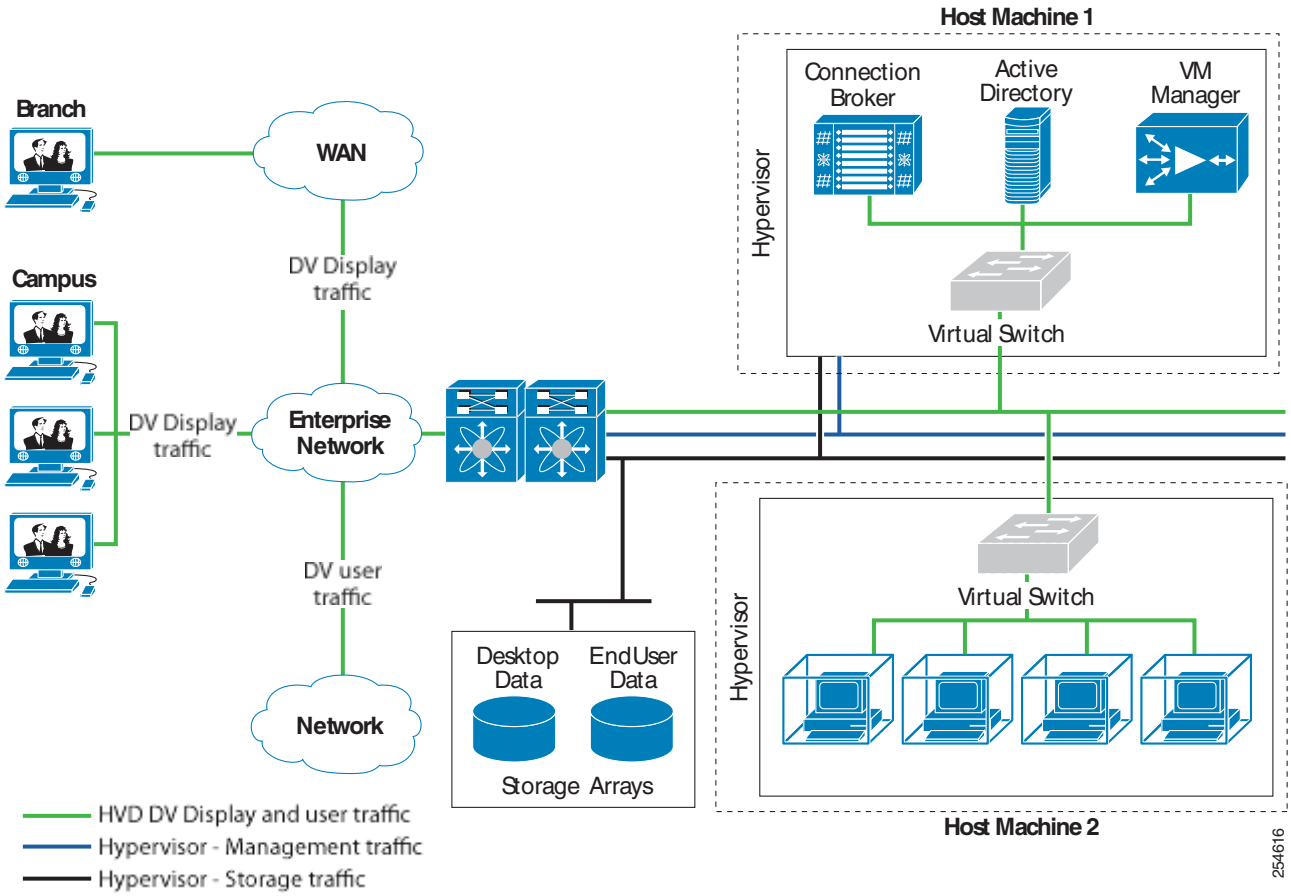
In a typical VMware View deployment the data center houses, at a minimum, an HVD pool (Host Machine 2 in Figure 11), the View Connection Server, Active Directory, and vCenter Server. Each HVD in the pool is a virtual machine and the pool itself is on a physically separate host machine. Another host houses the VMs on which the connection broker, Active Directory, and the vCenter Server are installed (Host Machine 1 in Figure 11).

Note It is highly recommended that HVD pools and enterprise services be housed on different host machines.

The hypervisors in each host connect to separate storage and management networks marked black and blue, respectively. Note here that the endpoints in the campus or the branch connect to the already existing access network infrastructure, and the HVDs in the pool attach to the virtual switch installed

inside the hypervisor. All the endpoints, HVDs, and Desktop Virtualization server components are on the "green" network. Devices in the green network can reach each other using Layer 2 or Layer 3 connectivity. All the display protocol traffic originates and terminates between the endpoints and HVDs.

Figure 11 Connection Paths in VMware View Deployments



User Virtualization

User virtualization solutions link a user's unique desktop configuration (or persona) with a generic, pooled desktop (persistent or non-persistent). Cisco has validated the following solutions for user virtualization:

VMware View 5 Persona Management: separates user and application settings from the operating system, and stores personas in a per-user repository located on a network-based storage system. The repository is managed by an agent residing on a virtual desktop. At login, the agent assembles the persona information needed to launch the operating system.

Application Virtualization

Application virtualization "containerizes" and insulates applications from the guest operating system. This permits a single instance of a virtualized application to run across different Microsoft Windows OS versions, and helps reduce the risk that any malware infecting a virtualized application will not escape the container to infect the OS or other applications. Virtualized applications are delivered to these

"sand-boxed" containers from a central repository. VMware ThinApp can be used for application virtualization in View environments. ThinApp packages applications into single executable files, which can be run in isolation from the OS and other applications. This functionality enables applications to be run on different Microsoft Windows platforms and/or user profiles. Refer to the vendor documentation link for information on installing and configuring this solution.

Storage Design and Best Practices

Storage is one of the most critical elements of the virtualized data center.Analysts estimate that storage represents anywhere from 40% to 80% of the cost of a desktop virtualization investment, so designing an effective storage solution is essential to achieving Return-On-Investment goals.The storage system also has a major impact on desktop performance and user experience.With traditional desktop PCs, users have enjoyed fast, direct access to dedicated disks.With desktop virtualization, these storage resources become centralized and designers must consider the effects of latency and resource contention on the user population.Storage system design also influences the ability to scale to accommodate new users, as pilot programs expand to full-scale deployments.More users generate greater demand for storage resources, higher levels of I/Os per Second (IOPS), and higher latency.In short, the storage architecture is a significant factor in determining the success or failure of a virtualized desktop implementation.

Designing storage for desktop virtualization presents significant challenges in terms of cost and performance.It can be difficult to balance costs with capacity and IO access requirements.Organizations often over-provision their storage arrays, adding more disks than are needed to ensure sufficient levels of IO access.This tactic may preserve the user experience, but at great cost.Storage design is also complicated by the episodic nature of user access.Steady state operation, for example, may require only eight or ten IOPS per desktop.During a login storm, when large numbers of users login at about the same time, IOPS may burst to very high levels.Deploying sufficient storage to handle peak loads results in excess capacity during steady state operations.On the other hand, insufficient IOPS during peak load situations can cause significant delays for end users, as they are unable to access their desktops.

In virtualized environments, IO also tends to be highly random.The Windows operating system is designed to optimize IO operations so that blocks are written to (or read from) disk sequentially.This sequential operation minimizes the time required for disk reads/writes and improves performance.In the virtual world, hypervisors tend to produce smaller random blocks of IO (also known as the blender effect).These random IO operations have a significant impact on disk performance, and may introduce unacceptable levels of latency.

The goal of the storage system designer, then, is to implement a solution that is cost-effective, predictable, and responsive to sudden changes in workload.To that end, the storage architecture is designed to integrate a wide range of storage technologies and optimization techniques. Design goals include:

- **Flexibility**: validate multiple approaches so that customers can confidently deploy solutions tailored to their environments.Support NAS- and SAN-based central, shared storage, as well as a range of local storage options (server memory, local HDDs, SSDs, and UCS Storage Accelerator).
- **Efficiency**: protect customer Return-On-Investment (ROI) by ensuring that storage resources are optimized.Blend Cisco and ecosystem partner capabilities such as local storage, thin provisioning, caching, non-persistent pooled desktops, and other approaches that improve the efficiency of the overall storage system.
- **Performance**: ensure a superior end user experience by validating Cisco and ecosystem partner solutions such as SSDs, UCS Storage Accelerator, IOPS offloading, and other approaches that can help provide predictable performance.

Storage Deployment Models

The solution has validated three major approaches to implementing storage in a desktop virtualization deployment.These models include: (1) central shared storage, (2) VMware View Storage Accelerator server-based caching with central shared storage, and (3) local server-based Cisco UCS Storage Accelerator with central shared storage.

Central Shared Storage

Traditionally, virtualization vendors recommended centralized, shared storage systems for virtual desktop deployments.In this deployment model, the shared storage system contains the user's machine environment, including the operating system, user profiles, and user data (see Figure 12).These systems, based on high capacity storage arrays, offer centralized security, high performance, and resource sharing.Centralized shared storage also facilitates the use of advanced hypervisor features and aids in desktop migration.

The solution supports both Network-Attached Storage (NAS) and Storage Area Networks (SAN).Shared storage systems are provided by ecosystem partners EMC and NetApp.Each partner offers solutions that support both NAS and SAN. In the test environment, the EMC VNX system was deployed for SAN-based block storage, and the NetApp FAS system was configured for file-based NAS.

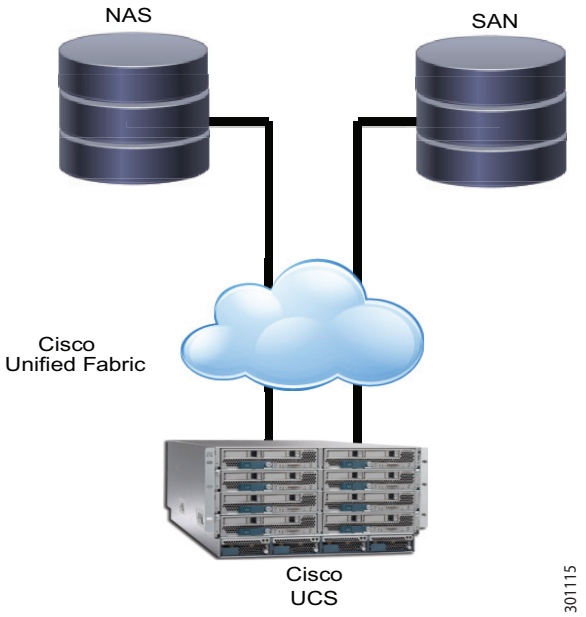
The EMC Unified Storage solution provides advanced failover and fully automated storage tiering to virtual desktop environments. EMC Unified Storage solutions can connect to multiple storage networks through NAS, Small Computer System Interface over IP (iSCSI), and Fibre Channel SANs. system testing validated the EMC Unified Storage solutions for SAN deployments. Hosts were configured to boot from the SAN.

The NetApp FAS Series is a unified storage solution that supports both SAN and NAS deployments. NetApp storage arrays are highly optimized for the intensive read and write workload typical of virtual desktop environments (see the product documentation for more information). NetApp storage runs the Data ONTAP operating system, which provides SAN (FCoE, Fibre Channel, and iSCSI), NAS (Common Internet File System [CIFS] and Network File System [NFS]), primary storage, and secondary storage on a single platform so that all virtual desktop components (virtual machine OS, user persona, user data, applications, data, etc.) can be hosted on the same unified storage array.

Deployment Notes

- In this model, the VMs are configured to Boot from the SAN system in accordance with hypervisor recommendations.
- Virtual machines, desktop images, and user data are stored centrally in the shared arrays.
- Advanced storage features integral to the partner systems such as compression, de-duplication, caching tiering, and other optimizations are enabled.
- Storage optimization appliances that offload IOPS processing, perform in-line de-duplication, and deliver sequential IO can provide substantial benefits in this model.Refer to Cisco Design Zone for documentation on optimization products validated.

Figure 12 Centralized, Shared Storage



VMware View Storage Accelerator and Central Shared Storage

VMware View Storage Accelerator uses UCS server RAM and vSphere's Content-Based Read Cache capability to build an on-board read cache.The read cache stores VMDK blocks frequently accessed by virtual desktops.Caching these blocks offloads read IO from the shared storage system by serving read requests from server memory.This model works best for managing read-intensive operations such as boot and login cycles.However, it also improves read IO during steady state operations, and thus improves application response times.

Deployment Notes

- This feature is used with VMware View linked clones.
- First deploy through vCenter for all hosts in a domain; deploy for virtual desktops through View Administrator when creating pools.
- The cache is emptied upon reboot, and start rebuilding when the first desktop starts up.

Cisco UCS Storage Accelerator and Central Shared Storage

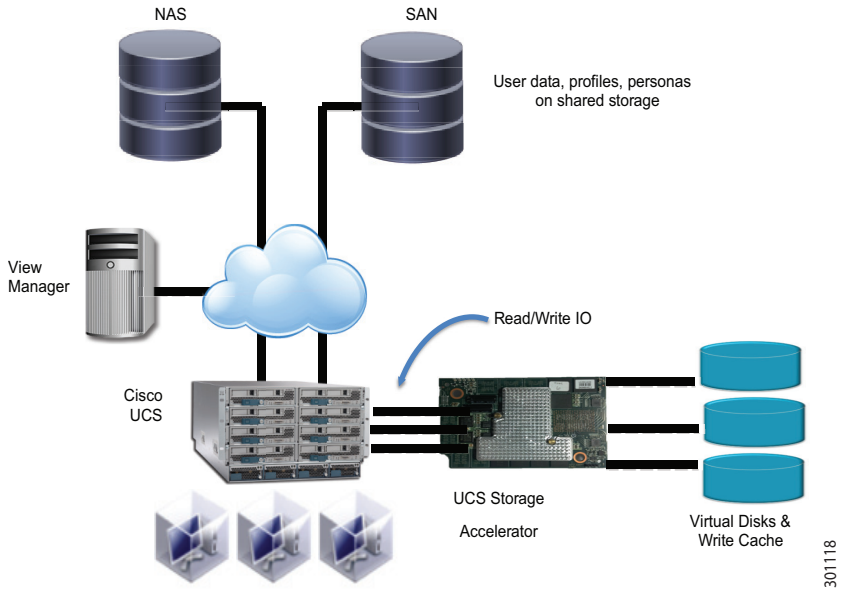
The UCS Storage Accelerator is a Flash-based caching solution that resides on a Cisco UCS B200 M3 blade server.The accelerator is a 785 GB mezzanine card that delivers much higher levels of IOPS than a typical shared storage system.IO requests for a particular server are contained within that server, which reduces latency compared to traditional approaches.This enables network administrators to support a guaranteed number of users at a lower cost, and with predictable performance.

In this model, a golden master image and all associated clone images are hosted on the Cisco UCS Storage Accelerator, which is installed on the blade server. A central copy of the golden master image can be maintained on the shared storage

Deployment Notes

- The hypervisor should be configured to store linked clone desktop images on the server-based accelerator.
- Store persistent information such as user data on the shared storage system.
- If a server fails, users simply re-connect to another desktop through their connection broker.
- VMware vSphere 5.1 supports live migration of virtual desktops without requiring shared storage.

Figure 13 Cisco UCS Storage Accelerator and Central Shared Storage



Storage Capacity Best Practices

Storage capacity is the amount of disk space needed to accommodate the virtual machines used for desktop virtualization. Capacity requirements can be roughly estimated by multiplying the number of virtual machines by the size of the hard-disk space allocated per virtual machine. For example, with 100 virtual machines each having assigned a disk space of 30GB, 100 x 30GB= 3TB of data capacity would be required. Refer to the Scaling and High Availability chapter of this guide for information on storage capacity planning. Cisco also recommends working closely with both storage and virtualization vendors to ensure that storage resources are efficiently utilized.

The storage arrays validated support various forms of thin provisioning. Thin provisioning is a form of oversubscription that makes storage available as needed. This capability enables higher storage use and reduces the need to overprovision and preallocate physical disk capacity.

Storage arrays also typically provide the capability to identify and eliminate duplicate data from disks. With deduplication enabled, only unique data blocks are stored. Duplicate blocks are discarded, and their associated disk space is reclaimed. The deduplication process can be scheduled to run at specific intervals. A best practice is to run deduplication at off-peak intervals to reduce its impact on application performance.

Storage requirements also can be minimized with features provided by the desktop virtualization solution. For example, capabilities such as View Storage Accelerator, which caches common image blocks to decrease storage consumption during boot storms, can further reduce the amount of disk space consumed.

IOPS Best Practices

IOPS loads for virtualized environments can vary greatly. For example, when virtual machines boot up, much data needs to be read from disk. Other operations (such as logins and logoffs) generate a lot of disk write operations. IOPS loads are subject to large peaks, such as when a lot of users start using their virtual desktops at roughly the same time. Scheduled tasks such as virus scans and file index searching also can cause peaks in IOPS load.

To optimize read IOPS, master images that will be read by many virtual machines should be placed in a fast-access data store. Both EMC and NetApp storage solutions have been validated and provide technologies to cache the master image in the first and fastest level of cache on the storage array.

The solution also validates the use of View Storage Accelerator, which leverages VMware vSphere's Content-based Read Caching to minimize the impact of boot and login storms on shared storage systems.

Latency Best Practices

Desktop operating systems originally were developed for platforms in which storage and CPU were in very close proximity (for example, a personal computer with a local hard drive). Virtualized environments, on the other hand, may utilize multiple types of local and shared storage, each having different access times. The latency associated with various storage technologies can impact the overall user experience. Cisco has validated two best practices for minimizing storage-associated latency.

One approach is to deploy reference building blocks (such as FlexPod and Vblock) that combine compute and storage. With the building block approach, the compute and storage resources are interconnected by a single layer of very low latency switches. These building blocks are also an effective method for scaling deployments, because capacity can be added in a linear manner as demands increase.

Latency can be further reduced by deploying local SSDs or server RAM for storage purposes.

Availability Best Practices

High availability is essential with user desktops and data now centralized. Best practices include deploying a redundant data center architecture, and creating network redundancy where practical. With redundant active/active data centers, each can accommodate 50% of the user workload. Each user will have ample CPU and storage resources. Should one data center fail, those users can fail-over to the other data center. The additional load may impact overall performance somewhat, but this is generally an acceptable alternative to major over-provisioning. Data availability is also important, and user data is usually separated from the operating system image and maintained and a centralized shared storage solution. Cisco recommends that organizations work with their storage vendors to develop the appropriate availability and backup strategies.

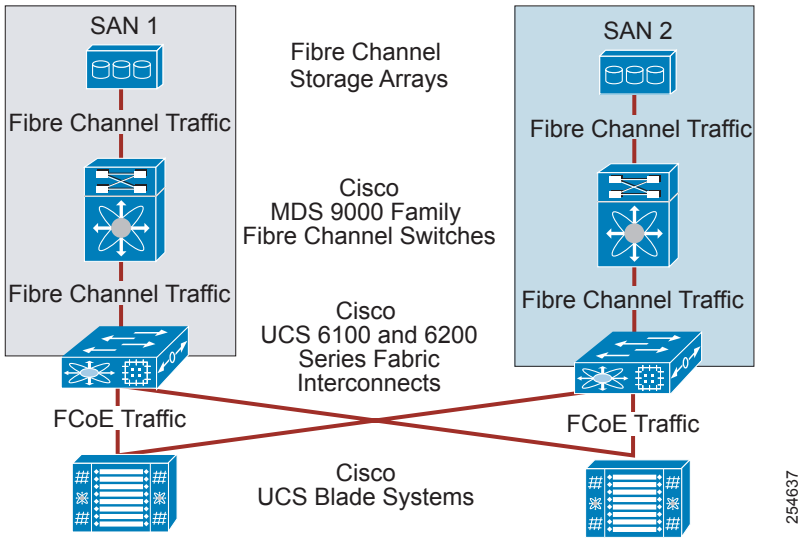
Storage Networking (SAN)

Fibre Channel storage arrays are attached to members of the Cisco MDS family of switches. The Cisco MDS switches allow multiple arrays to talk to multiple hosts (servers) in much the same manner as an Ethernet switch would. The Cisco MDS switch connects to Fibre Channel uplinks on the Cisco UCS

6x00 Fabric Interconnects. Traffic is encapsulated by the Cisco Fabric Interconnects into FCoE frames and passed to the server's Virtual Interface Cards, where it is converted back into Fibre Channel packets and presented on the vHBA.

Fibre Channel connections typically operate in an active or standby mode. If SAN A is the primary array, SAN B is a mirror copy. In some cases, depending on the storage array vendor, the uplinks from Cisco MDS switch A and Cisco MDS switch B may point to the same array. However, a particular logical unit number (LUN) will show up on only one of the interfaces at a time. For traffic load balancing, the user should mix the LUN assignments across the two SANs.

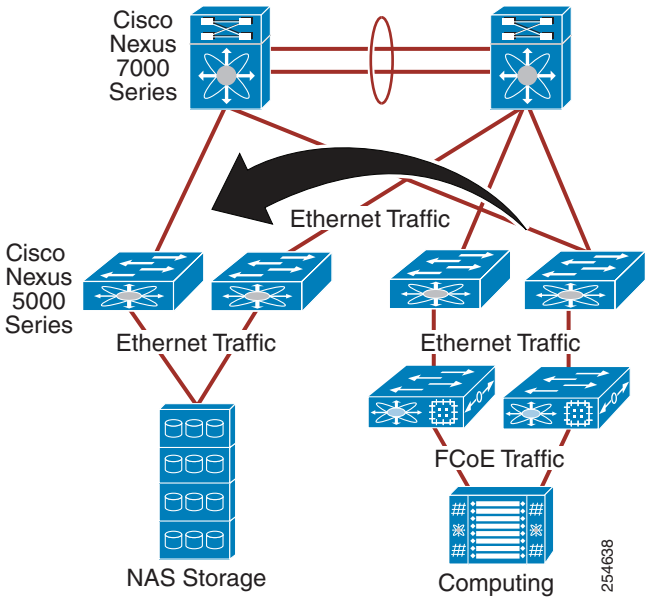
Figure 14 Fibre Channel Storage Network



Network-Attached Storage

With network-attached storage, traffic originating from the NFS server is switched as Ethernet traffic to the storage array via the Cisco Nexus 5000 or 7000 Series Switches. The Cisco Nexus switches provide a universal platform for diverse data center needs, facilitate network consolidation, and enable a seamless transition from traditional to advanced technologies. Like Fibre Channel attached storage, the remote array is mapped to a local hypervisor storage pool, and the hypervisor provides virtual desktops with simulated local storage. Since this is Ethernet-based, storage is assigned to a separate storage VLAN. In most cases, the storage array is placed in the same VLAN. This alleviates the need for the hypervisor or Desktop Virtualization virtual machines to have multiple default routes, which can cause network issues.

Figure 15 Ethernet-based Storage Network



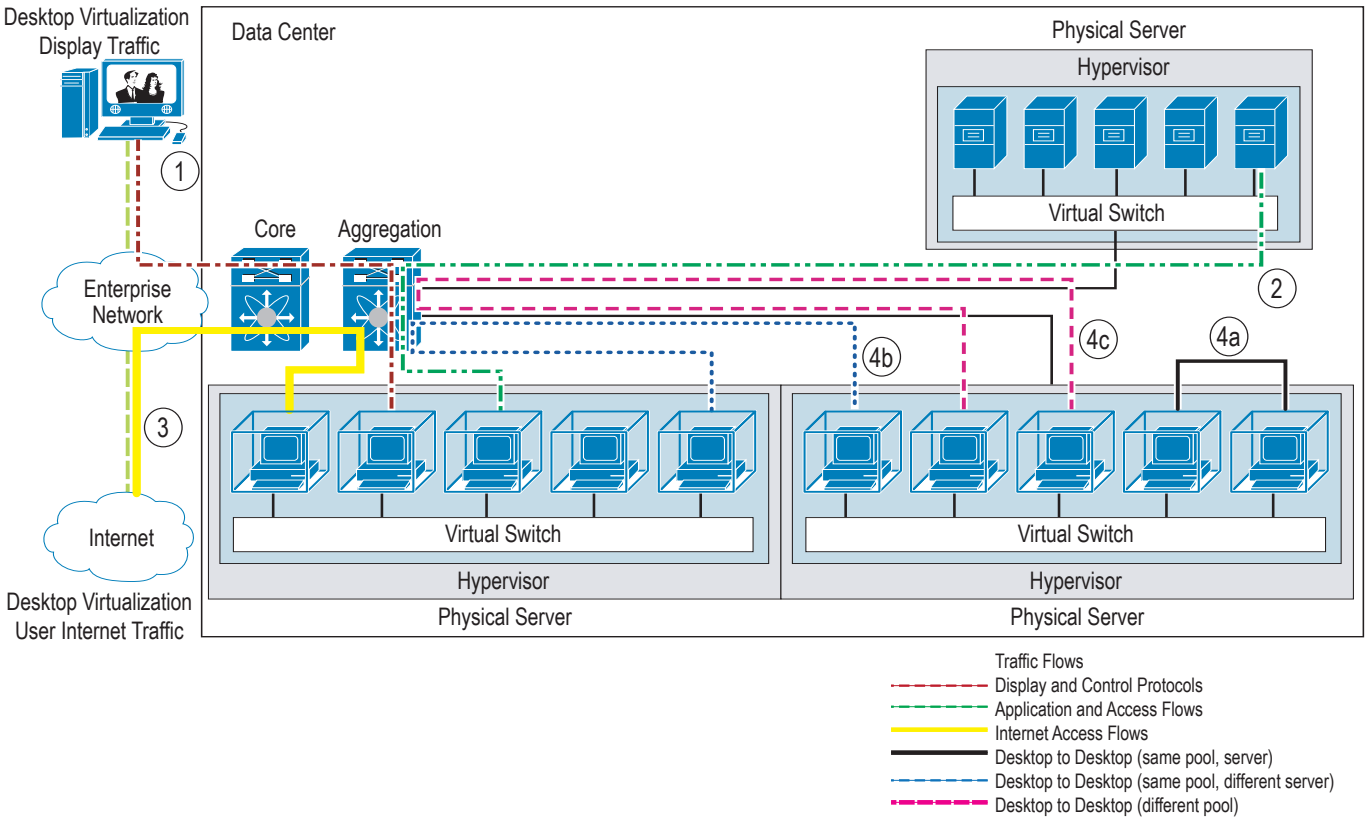
Data Center Networking

The data center network design follows the Cisco standard layered data center design, with access, aggregation, and core layers, with the physical core and aggregation layers collapsed within the Cisco Nexus 7000 switches by the use of virtual device instances. This approach provides high levels of performance, scalability, and availability for the validation environment. To ensure maximum efficiency and security, the Cisco Virtual Workspace system makes extensive use of VLANs to separate traffic flows in the data center network.

For virtual desktop traffic, as shown in Figure 16, the traffic flows include:

- Display and control protocols that will be sourced from the virtual desktop and terminated on the remote endpoint.
- Flows associated with access to applications existing within the data center which will travel across the L2 segment until it reaches either the router in the Aggregation Layer or a Firewall and then routed to another L2 segment containing the application.
- Internet access flows sourced from virtual desktops will travel up the data center network stack, cross into the Enterprise Network, and exit the corporate network via the Internet Router. This traffic may cross several firewalls along the way.
- Traffic may flow directly between virtual desktops. If traffic flows between desktops in the same pool and on the same server, traffic will never exit the server. If this needs to be prevented, use private VLANs or the Cisco Virtual Security Gateway to establish zones.
- If flows are between desktops on different servers but within the same pool, traffic travels up to the aggregation layer switch than back down to the other server.
- If flows are between desktops in different server pools, traffic travels up the network stack to the aggregation layer, where it is routed to the appropriate subnet and forwarded to destination server/desktop.

Figure 16 Virtual and Physical Traffic Flows in the Data Center Network



VLANs for the Virtualized Data Center

Figure 15 shows a server with a pair of interface cards installed and the logical connections made within the server. Each card presents to the server Fibre Channel Host Bus Adapters and Ethernet NICs. When the hypervisor (vSphere in this example) is installed, the Fibre Channel HBAs are mapped to virtual HBAs, which can be used for attaching remote storage. The Ethernet interfaces are mapped to the Cisco Nexus 1000V virtual switch. The Cisco Nexus 1000V Series treats these as uplink ports. These ports are usually configured as trunk ports and connected through EtherChannel for redundancy and load balancing.

Within the Cisco Nexus 1000V Series, several virtual interfaces (vETHs) are created. One is reserved for the kernel traffic (connection to vCenter) and is put into a separate VLAN and assigned a unique IP address. A second one is reserved for vMotion, and again placed in a separate VLAN and assigned a second unique IP address. Neither interface is seen by the virtual desktops. For security reasons, the virtual desktops do not have access to these VLANs. One final reserved vETH may be created if the vSphere hypervisor datastore is connected via Ethernet-based storage (NFS). The Ethernet-based storage traffic should also be isolated into its own VLAN.

Additional VLANs are created for the virtual desktops. IP addresses may be statically defined or acquired via Dynamic Host Configuration Protocol (DHCP). Address assignments and subnet mask need to be created according to the size of the virtual desktop pool. Virtual desktops can migrate only within the same pool (subnet, VLAN, and so on). For each virtual desktop, only a single vNIC needs to be created. It is possible to have virtual desktops from different pools running on the same physical server. The configuration for the trunk ports needs to include all possible VLANs, including those for the

allowed virtual desktops. The Cisco Nexus 1000V Series creates an EtherChannel across the two CNA (NIC) interfaces, so loads on each interface should be monitored to ensure the appropriate load balancing algorithm was selected.

The Cisco Integrated Management Controller is forwarded internally to the Cisco UCS Fabric Interconnects. The Cisco UCS Manager assigns and manages this interface. The IP address assigned to the Cisco Integrated Management Controller must be on the same subnet as the Cisco UCS Manager. This address is used for remote KVM and server maintenance.

The VMware vKernel interface is used to communicate with vCenter. The IP address should be assigned to a separate VLAN from the Cisco Integrated Management Controller interface. This VLAN should be routable to the vCenter server's VLAN.

The vMotion interface is used for virtual desktop mobility. It also should be assigned to a separate VLAN that is common to all the servers within the same virtual desktop pool. This network does not need to be extended beyond the access layer, since vMotion is a not routable protocol. vCenter will use the vMotion interfaces on the servers (and the associated VLAN) for the traffic flow needed to migrate the virtual desktop. Because the virtual desktop's storage is maintained remotely, only the contents of RAM (the CPU state) of the desktop are transmitted.

Figure 17 Data Center Connectivity (vSphere Example)

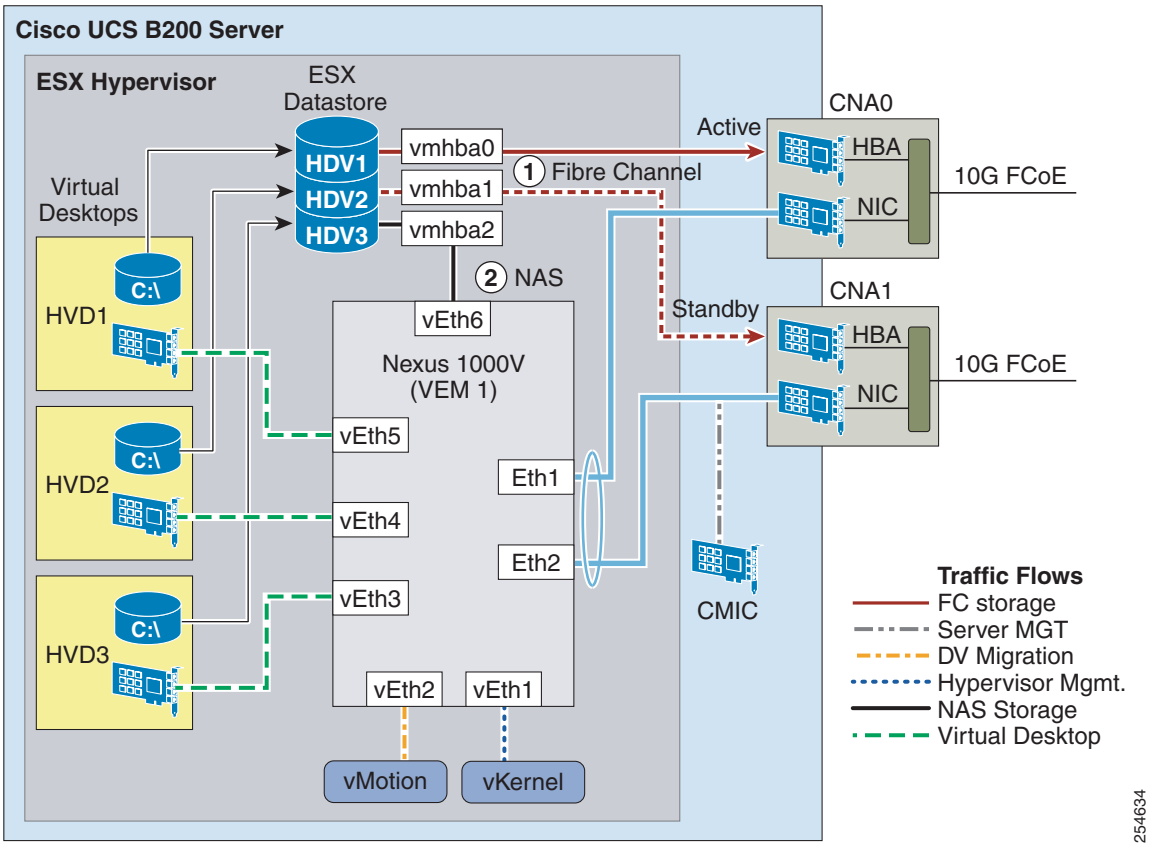


Table 6 List of Recommended VLANs

Name	Devices	Usage	Notes
Data Center			
Cisco UCS Manager	Cisco UCS blade server chassis, fabric interconnects, and blade server integrated management controller	Out-of-band management of all Cisco UCS components	May be shared with other management VLANs.
vMgmt	Hypervisor	Management of all hypervisors	Specific to a particular Desktop Virtualization pool; therefore, there may be separate VLANs for each pool.
vMigration	Hypervisor	Migration of DV VMs within the DV pool	As with vMgmt, there may be several separate VLANs for each pool
Cisco Nexus 1000vControl & Cisco Nexus 1000vData	Cisco Nexus 1000V Series VEM and VSM	Cisco Nexus 1000V Series management	Common to all physical servers that provide the distributed switch functionality. There may be more than one Cisco Nexus 1000V installed in a data center
VMData	DV virtual machines	User's network accessDV display protocol	DV machines should only need one interface One VLAN for each DV pool.
VMStorage	Hypervisor	Provides isolation for storage traffic	For Fibre Channel or IP-based storage
Endpoint Location			
Data	DV endpoint	Connection to the DV virtual machine	The DV data VLAN may be segregated from the legacy compute endpoint VLANs

Name	Devices	Usage	Notes
Data Center			
Cisco UCS Manager	Cisco UCS blade server chassis, fabric interconnects, and blade server integrated management controller	Out-of-band management of all Cisco UCS components	May be shared with other management VLANs.
vMgmt	Hypervisor	Management of all hypervisors	Specific to a particular Desktop Virtualization pool; therefore, there may be separate VLANs for each pool.
Voice*	Unified communications endpoints	Provides unified communications access	Traditional voice VLANs

*The endpoint data and voice VLANs are discussed in the section Modular Data Center Blocks.

Virtualized Data Center Management

The design creates a separate IP network for management traffic with a dedicated IP subnet and VLAN. This approach enables remote-access, SNMP, syslog, and FTP traffic to be managed out of band. Separating traffic in this manner helps ensure that end-user and administrative traffic flows do not compete for or interfere with available bandwidth. Remote access to a device will not be compromised when the device needs to be reset or provisioned. This practice also mitigates threats to network security and availability that could be introduced when end-user and administrative traffic share the same interface. Management and user traffic isolation is important in the data center, where desktop virtualization concentrates user traffic in an infrastructure traditionally used for server and management traffic loads.

Each management tool uses a specific set of protocols to communicate with devices. Refer to the vendor documentation for a complete list of protocols and ports used by each tool. Make sure that these ports are open on all intermediary routers, switches, and firewalls.

Table 7 Management Tools

Product	Management Tool	Description	Product Documentation Link
VMware ESX/ESXi and virtual machines	VMware vCenter and vSphere client	Use VMware ESX and ESXi hypervisor manager to create and manage virtual machines	VMware vSphere documentation
VMware View Manager 5.0	VMware View Administrator Console	Create virtual desktop pools, and grant user privileges, and monitor sessions.	VMware View documentation

Product	Management Tool	Description	Product Documentation Link
EMC Unified Storage	EMC Unisphere Management Suite	Provision and monitor the SAN based storage array	EMC Unisphere Management Suite
NetApp FAS 3170	NetApp Virtual Storage Console	Provision and manage NetApp Unified Storage arrays	NetApp Virtual Storage Console
Cisco UCS BSeries Blade Servers	Cisco UCS Manager	Provision and monitor the Cisco UCS B Series Blade Servers	Cisco UCS Manager
Virtual desktops (guest OS)	Standard enterprise desktop and OS management tools(Altiris, Microsoft Systems Management Server [SMS], and System Center Configuration Manager [SCCM])	Provision and monitor the virtual desktops.	Altiris reference documentation
Microsoft Active Directory, Domain Name System (DNS), DHCP	Standard enterprise management tools	Manage end user profiles and perform authentication of user sessions. Provide DHCP services to endpoints.	Microsoft Active Directory and Network services

Modular Data Center Blocks

Cisco has partnered with key storage and virtualization partners to develop and deliver prepackaged, validated infrastructure solutions for data centers. These solutions provide computing, storage, server, management, and virtualization resources in an integrated, modular form. The Cisco Virtual Workspace Smart Solution architecture supports the use of these solutions as effective data center building blocks. These packages simplify planning, facilitate acquisition, speed deployment, and reduce risk. Cisco has partnered on the following solutions:

- **Vblock**: the Virtual Computing Environment Company (VCE) was formed by Cisco and EMC, with investments by VMware and Intel. Vblock platforms integrate Cisco UCS and networking technology with VMware vSphere and EMC storage. VCE offers different size Vblocks for different deployment sizes. See http://www.vce.com/solutions/ for more information on Vblock.
- **Flexpod**: unites Cisco UCS and networking technologies with NetApp storage. These integrated platforms have been validated with hypervisors from VMware, Red Hat, and Microsoft. Flexpod solutions, too, have been pre-configured and validated to speed deployment. See http://www.netapp.com/us/technology/flexpod/ for more information.
- **VSPEX**: Cisco partners with EMC on the VSPEX solution, which integrates Cisco UCS and Nexus equipment with EMC storage and VMware vSphere for server virtualization. VSPEX is a pre-validated and modular platform that provides a complete end-to-end solution. See www.cisco.com/go/vspex for more information on the VSPEX architecture.

Virtualization Aware Network

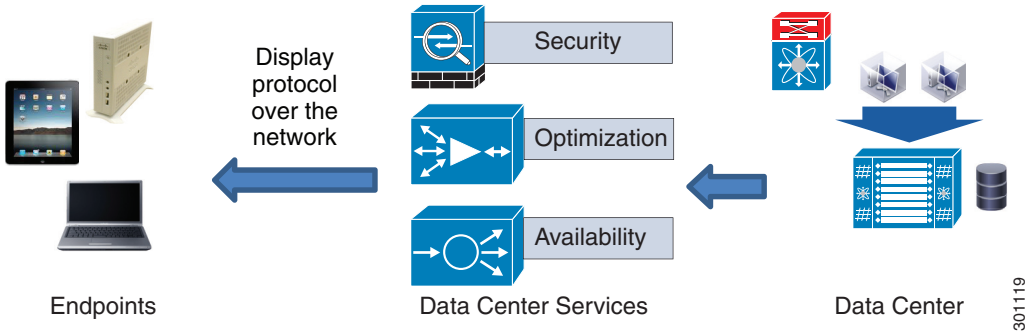
The Cisco® Virtual Workspace network design uses Cisco best practices for deploying campus, branch-office, and data center network infrastructure. However, a Cisco Virtual Workspace network design also must consider virtual desktop display protocols, unified communications requirements, and changing traffic patterns to create a better user experience. Display protocol traffic may be optimized and secured for network transport. The network must provide high levels of availability, reliability, and security especially for virtual desktops and data center applications because a network failure impacts access to all applications.

This chapter, together with the Securing Cisco Virtual Workspace chapter, describes the design of a highly available and secure network for desktop virtualization.

This chapter addresses the following critical elements of a successful design, summarized in Figure 19:

- Campus, branch-office, and teleworker network requirements
- Data center edge considerations
- WAN optimization
- Quality of service (QoS)
- Network management

Figure 18 Elements of a Desktop Virtualization Network



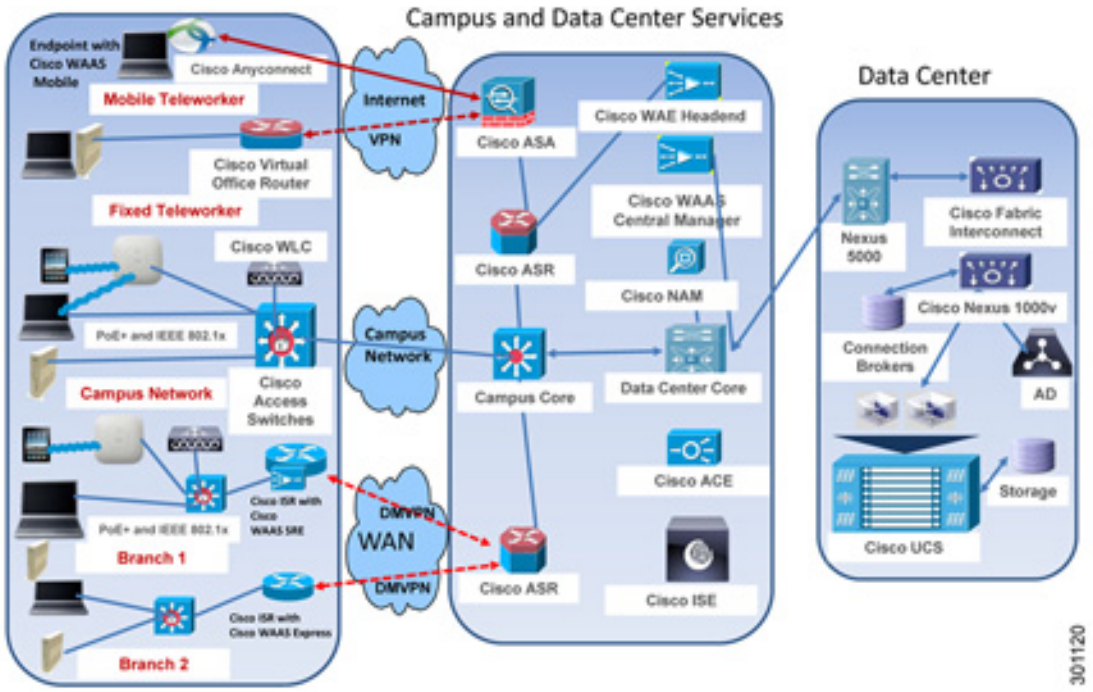
What is New in Release 2.7?

- Cisco Wide Area Application Services (WAAS) 5.1
- Cisco TrustSec 3.0
- Cisco BYOD Smart Solution 1.0.

Cisco Virtual Workspace Network Deployment Models

Figure 19 shows the three most common Cisco Virtual Workspace (VXI) Smart Solution deployment models. Virtual desktop users can be on campus, in a branch office, or fixed or mobile teleworkers. The Cisco Virtual Workspace network is designed to help ensure that users enjoy the same high-quality experience, regardless of physical location. This section discusses Cisco Virtual Workspace (VXI) Smart Solution design requirements for each scenario, and the main technologies used to meet these requirements.

Figure 19 Main Components of Cisco Virtual Workspace Networks



Campus

The campus network connects end users and devices in the corporate network with the data center and WAN. This section describes best practices for connecting endpoints to a first-hop switch, because this switch is where all the endpoint services are delivered in the campus network. For overall campus network design principles (from the first-hop switch to the data center edge), please refer to the Cisco Enterprise Campus 3.0 Architecture recommendations. The Cisco Virtual Workspace campus network provides a comprehensive set of services, such as:

- Power over Ethernet (PoE)
- Secure access control using IEEE 802.1x
- Intelligent and dynamic provisioning of supported endpoints to VLANs with appropriate QoS and security policies
- Endpoint location tracking
- Traffic monitoring and management

To enable these services, Cisco Virtual Workspace endpoints within the campus network should be connected to a wiring closet switch such as a Cisco Catalyst® 4000, 3000, or 2000 Series Switch.

Automatic Port Provisioning

Auto Smartports enable the switch to dynamically provision Cisco Virtual Workspace clients by automatically configuring a port based on the device identification information obtained through the Cisco Discovery Protocol or MAC addresses. Smartport macros are pre-created, customizable configuration scripts based on Cisco best practices that allow administrators to easily set up common switch-port configurations. Auto Smartports help ensure consistent use of security, QoS, and high-availability policies.

With Auto Smartports, the macros are applied by the switch on the basis of end point credentials. The configuration is removed when the link to the end point is dropped or when the user session is terminated. Cisco VXC endpoints do not support Smart-ports based on dynamic macros. Cisco Virtual Workspace testing enabled the Auto Smartports feature to detect Cisco IP Phones and used static MAC address-based macros for Cisco VXC endpoints. The Cisco Catalyst 4500E and 3000-X switch platforms support both types of macros.

Static Smartports macros provide port configurations that can be manually applied on the basis of the device connected to the port. When a static macro is applied, the command-line interface (CLI) commands attached to the macro are added to the existing port configuration. These configurations are not removed during a link-down event on the port.

To use a static macro, configure a MAC address group with a MAC address operationally unique identifier (OUI)-based trigger. The first three octets of the MAC address identify the vendor and can be used to define the MAC address group trigger. Map the trigger to a built-in or user-defined macro, and the macro can be applied on an access switch globally or on an interface. This macro is applied when the Cisco VXC endpoint connecting to the switch meets the trigger condition defined earlier. If the connecting endpoint does not trigger the macro, the port is considered untrusted and can be placed in a guest VLAN by default. The user-defined macro file can be maintained in a remote server location, and all access switches in the network can consistently apply the same macro; this approach is the recommended deployment model for Cisco Virtual Workspace.

The procedure for manual configuration can be found at:

http://www.cisco.com/en/US/docs/switches/lan/auto_smartports/12.2_55_se/configuration/guide/iosaspcg.pdf

Securing Endpoint Access

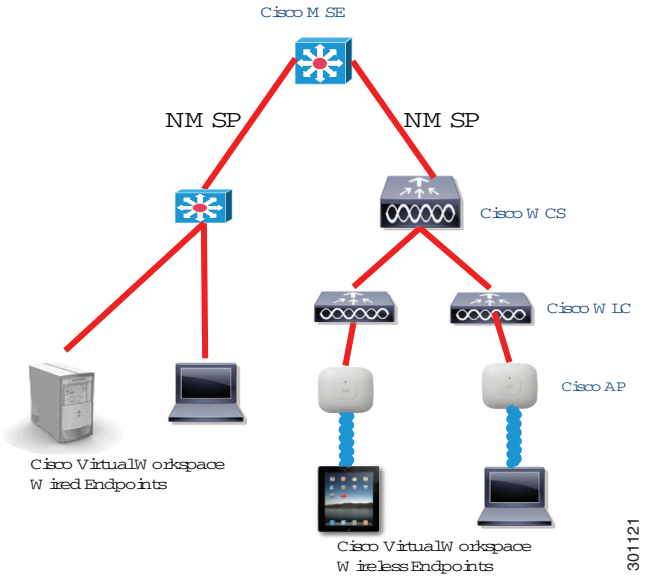
Cisco Catalyst switches provide security and availability functions for Cisco Virtual Workspace endpoints, helping ensure that some traffic types are characterized and secured based on log-in characteristics. For example, contractors can use their own devices, with the contractor's virtual desktop accessed using a thin or thick client. When the contractor connects to the network, the contractor logs in using the username and password and is authenticated using IEEE 802.1x; the Cisco Catalyst switches map the user to a discrete contractor VLAN with access control lists (ACLs) that enable access to the Internet and connectivity to the contractor desktops only. The Cisco Catalyst 4500 E and 3000-X platforms provide IEEE 802.1x and port security features to provide endpoint-level security. The deployment and design details for these security features are discussed in Securing Cisco Virtual Workspace chapter of this guide.

Endpoint Location Tracking

The Network Mobility Services Protocol (NMSP) provides the capability to determine the physical location of a host in the network. It provides the capability to track assets for inventory and compliance purposes. Access switches, such as the Cisco Catalyst 4500E and 3000-X platforms, provide a wired location service feature and transmit location information for connected endpoints to the Cisco Mobility Services Engine (MSE). Cisco MSE is an appliance-based platform for delivering mobility services in a centralized and scalable way across wired and wireless networks.

In the case of a wireless connection, the access point tracks and transmits location information for its connected Cisco Virtual Workspace endpoints to the Cisco MSE through the Cisco Wireless Control System (WCS), shown in Figure 20. Cisco WCS is the management platform for location services and device tracking. It allows you to setup event notifications, and alarms based on location triggers. The actual location information for the switch and access point and the ports on the switch is configured in the switch or access point itself.

Figure 20 Location Tracking Topology



Device information such as the MAC address, IP address, serial number, unique device identifier (UDI), and model number can be obtained by the switch or the access point. Some of Cisco VXC and other supported endpoints do not support Cisco Discovery Protocol (CDP) and Link Layer Discovery (LLDP) Media Endpoint Discovery (MED). In that case, you should use Address Resolution Protocol (ARP) snooping to learn the IP address and MAC address of the endpoint on wired switches and the Wi-Fi endpoint registration sequence on wireless access points. For the Cisco VXC endpoint attached to an IP phone, the location tracking can be performed by using Cisco Discovery Protocol and LLDP-MED from the phone. Switches and access points also transmit state information such as device category (wired or wireless), time of attachment, and operational state (connected or disconnected). Cisco MSE holds the location database and, in this release of Cisco Virtual Workspace, a manually maintained list of MAC addresses tied to each enterprise client. All models of Cisco switches and Cisco access points support location services natively and can be used in Cisco Virtual Workspace. However in the Cisco Virtual Workspace design, only the Cisco Catalyst 4500E and 3000-X were validated.

The guidelines for configuring location tracking can be found at: http://www.cisco.com/en/US/docs/solutions/Enterprise/Data_Center/VXI/configuration/VXI_Config_Guide.pdf.

CLI commands for the above sequence and other useful debug information can be found at:

http://www.cisco.com/en/US/docs/switches/lan/catalyst3750/software/release/12.2_50_se/configuration/guide/swlldp.html

Detailed configuration steps and screen shots for Location based services, also called context-aware-services, can be found at:

http://www.cisco.com/en/US/partner/products/ps9742/products_installation_and_configuration_guides_list.html

http://www.cisco.com/en/US/partner/docs/wireless/mse/3350/7.0/CAS/configuration/guide/CAS_70.html

Wired and wireless endpoints can be looked up by IP address, partial IP address, MAC address, partial MAC address, or VLAN ID. Cisco WCS can be set up to serve notifications when certain conditions about endpoint location or status are met. The links listed in this section discuss these aspects of tracking. For a quick video overview of these functions, please see:

http://www.cisco.com/web/techdoc/wcs/location/client-tracking/tracking_wi-fi_clients.html?referring_site=bodynav

Campus Access High-Availability and In-Service Software Upgrade

Cisco Virtual Workspace requires high availability of the network and reduced downtime during the planned software upgrades. Cisco Catalyst 4500E in a redundant configuration with dual supervisors provides high -availability and allows Cisco IOS® Software to be upgraded or modified while packet forwarding continues using the In-Service Software Upgrade (ISSU) feature. The ISSU feature allows a nondisruptive user experience during a planned upgrade or modification. The ISSU process with the Cisco Catalyst 4500E Supervisor Engine 7-E is reduced to a single-line command, and the traffic interruption is less than 10 milliseconds (ms). The Cisco Virtual Workspace session PoE and data plane stay up during the ISSU process. Access-switch high availability is provided by the Cisco Catalyst 4500 Series Switches modular chassis through ISSU. A detailed deployment guide can be found at

http://www.cisco.com/en/US/prod/collateral/switches/ps5718/ps4324/prod_white_paper0900aecd805e6a95.html

Branch-Office

Branch-office users may access their virtual desktops across bandwidth-constrained or untrusted WAN links and thus present a number of challenges.

Dynamic Multipoint VPNs in Cisco Virtual Workspace

In a Cisco Virtual Workspace system, virtual desktop traffic usually travels from the branch office to data center. Collaboration traffic may be peer to peer between branch offices. Both types of traffic may need to travel across untrusted WAN links. Because of these factors, a solution is needed that can dynamically create and tear down encrypted tunnels based on the traffic destination. Cisco Virtual Workspace recommends Dynamic Multipoint VPN (DMVPN) for this purpose. DMVPN provides both security and optimized route selection in distributed environments.

DMVPN routers use generic routing encapsulation (GRE) -based tunnel interfaces that support IP unicast, IP multicast, and broadcast traffic as well as dynamic routing protocols such as Enhanced Interior Gateway Routing Protocol (EIGRP). After the initial spoke-to-hub or branch-office-to-campus tunnel is active, dynamic spoke-to-spoke or branch-office-to-branch-office tunnels are created, depending on traffic-flow requirements. DMVPN uses Next-Hop Resolution Protocol (NHRP) to find other spokes based on IP destination. The branch-office-to-branch-office tunnels are encrypted using IP Security (IPsec). Idle connections between branch offices are timed-out. Internet Security Association and Key Management Protocol (ISAKMP) provides dead-peer detection (DPD) so that unused or orphaned tunnels do not remain active. IPsec tunneling is also required when DMVPN hubs are behind firewalls implementing Network Address Translation (NAT). Since the tunnels are created behind the physical interface, branch-office routers configured for DMVPN can use the Dynamic Host Configuration Protocol (DHCP) address provided by the service provider. The DMVPN hub, however, must have a static IP address so that all branch-office routers can be created in the initial spoke-to-hub tunnel.

Use Cisco Integrated Services Routers Generation Two (ISR G2) for the spoke (branch office) and use Cisco Aggregation Services Routers (ASRs) and Cisco 7200 Series Routers for the hub (head-end). All the routers must have security plus licensing for Cisco IOS Software to enable DMVPN functions.

WAN optimization is also used to help ensure a high-quality user experience. WAN optimization and DMVPN can be deployed together successfully, but some configuration best practices need to be followed. An excellent resource for DMVPN and Cisco WAN deployments in general is located at

http://www.cisco.com/en/US/solutions/collateral/ns340/ns414/ns742/ns982/c07-610746-02_wanDeploy.pdf.

All other main resources specific to Cisco DMVPN are located at

http://www.cisco.com/en/US/products/ps6658/index.html

Cisco Wide Area Application Services

Cisco WAAS enables WAN optimization to address the challenges of delivering virtual desktops. Refer to the WAN Optimization for Cisco Virtual Workspace (VXI) Smart Solution Vision section of this chapter for details.

Endpoint Access Services in the Branch

Campus access features such as PoE+, location tracking, and Auto Smartport provisioning previously discussed are also supported in the branch-office environment and are recommended in Cisco Virtual Workspace deployments.

Teleworker

The Cisco Virtual Workspace system supports both fixed and mobile teleworkers. Both types will likely access virtual desktops in the same data center. The only difference is the location from which the remote user is connecting.

A fixed teleworker uses a fixed network device, which provides secure connection and comprehensive network services to full- and part-time home-office workers, mobile contractors, and executives. A fixed teleworker in a Virtual Workspace environment can establish a secure connection using either of two methods:

- Cisco Virtual Office with Cisco VXC and all supported endpoints behind a Cisco Virtual Office router
- Cisco VXC 2111 and 2112 endpoints behind a single- and dual-VPN tunnel with a Cisco IP Phone

By providing extensible network services that include data, voice, video, and applications, the Cisco Virtual Office creates a comprehensive office environment for employees. The Cisco Virtual Office solution consists of the following components:

- A Cisco 800 Series Router and a Cisco Unified IP Phone
- A data center presence that includes a VPN router and centralized management software for policy, configuration, and identity control
- WAN optimization for teleworkers using either a thick or thin endpoint that supports Cisco WAAS Mobile. Cisco WAAS Mobile pairs with the Cisco WAAS Mobile server in the data center behind the VPN headend
- Deployment and ongoing services from Cisco and approved partners for successful deployment and integration, and consultative guidance for automating the deployment

A network enabled for Cisco Virtual Workspace provides additional benefits to the fixed teleworker using the Cisco Virtual Office solution. These enhancements provide the capability to deploy Cisco VXC behind the Cisco Virtual Office router. The end user can now access a virtual desktop through a secure VPN tunnel between the Cisco Virtual Office router and the corporate edge router. The virtual desktop

can have a Cisco Unified Communications client installed, enabling the user to control an IP desk phone. More information about the Cisco Virtual Office solution can be found at http://www.cisco.com/en/US/netsol/ns855/index.html.

Cisco Virtual Workspace mobile teleworkers connect to virtual desktops securely from any endpoint that supports a Cisco AnyConnect® Secure Mobility client. Mobile teleworkers are typically in unsecure network locations. One of the advantages of Cisco Virtual Workspace is that the data stays secure, even if the endpoint is stolen, damaged, or lost. Teleworker secure connections are discussed in detail in the Securing Cisco Virtual Workspace chapter.

Cisco WAAS Mobile optimizes virtual desktop traffic over these secure connections to provide a better user experience. The Cisco WAAS Mobile client is installed on the endpoint device, connecting through a VPN or in a home-office router (Cisco Virtual Office router) to the Cisco Virtual Workspace data center. The Cisco WAAS Mobile client pairs with a Cisco WAAS Mobile server positioned behind the VPN. Note that the Cisco WAAS Mobile server is not the same as a Cisco WAAS headend appliance and so cannot pair with the Cisco Wide Area Virtualization Engine (WAVE), Cisco WAAS Express, or Cisco WAAS on Cisco Services-Ready Engine (SRE) products. Cisco WAAS Mobile functions are turned off when a teleworker endpoint is in a campus or branch-office environment, to allow more full-featured WAN optimization. Cisco WAAS Mobile deployments are managed by Cisco WAAS Mobile Manager, which provides the capability to monitor performance and measure return on investment (ROI).

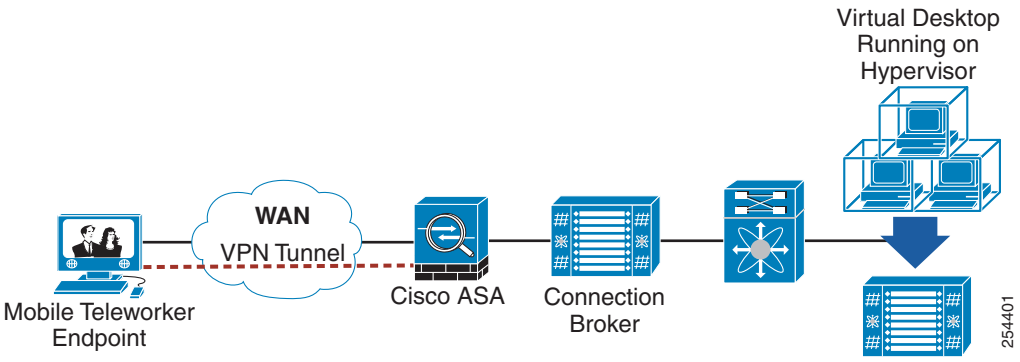
For more information about Cisco WAAS Mobile, see:

http://www.cisco.com/en/US/prod/collateral/contnetw/ps5680/ps6870/data_sheet_cisco_wide_area_application_services_mobile.html

Figure 21 shows the mobile teleworker with a thick client connecting to a hosted virtual desktop from an unsecure location using a secure connection.

- The mobile teleworker establishes an SSL VPN connection with the Cisco Adaptive Security Appliance (Cisco ASA) VPN gateway using the Cisco AnyConnect client.
- The mobile teleworker then uses the desktop virtualization client to communicate across the secure tunnel and authenticate with the connection broker in the corporate data center.
- After successful authentication, the connection broker displays the virtual desktop associated with the mobile teleworker.
- The mobile teleworker then selects and connects to the desired virtual desktop to establish the virtual desktop session. The entire session is now secured by the VPN tunnel.

Figure 21 Teleworker with Thick Client



Cisco Virtual Workspace Data Center Edge

Desktop virtualization relies on always-on network connectivity to the data center. Network availability has become a critical part for every network design. Cisco provides high-availability network design guidance focused on specific areas of the enterprise network, such as the data center, campus, branch office and WAN, and Internet edge. Hierarchical network design is a common strategy for designing a network for high availability. For high-availability network design guidance when building a network enabled for Cisco Virtual Workspace, see http://www.cisco.com/en/US/docs/app_ntwk_services/data_center_app_services/ace_appliances/vA1_7_/configuration/device_manager/guide/UG_ha.html.

Load Balancing, SSL Offloading and Cisco Virtual Workspace Health Monitoring

The connection brokers used to offer virtual desktops in a virtual desktop infrastructure (VDI) environment need to be load balanced. Load balancing has been validated with Cisco Application Control Engine (ACE). For load balancing and SSL offloading, the Cisco ACE is deployed at the aggregation layer in the data center using one-arm mode. In one-arm mode, the Cisco ACE is configured with a single VLAN that handles both client requests and server responses. Routed and bridged Cisco ACE deployments also work in this design. For one-arm mode, client-source NAT or policy-based routing (PBR) needs to be configured. The Cisco ACE then uses NAT to send the requests to the real servers. Server responses return through the Cisco ACE rather than directly to the original clients. This topology is convenient, because the Cisco ACE can be almost anywhere on the network, but its reliance on NAT makes it impractical in some situations. If you are using routed or bridging mode, confirm that there is sufficient bandwidth for all user display connections. In a one-arm topology, the Cisco ACE is connected to the real servers' connection broker through an independent router and acts as neither a switch nor a router for the real servers. Clients send requests to the virtual IP address on the Cisco ACE.

The recommended approach is to use two Cisco ACE appliances. Either the Cisco ACE Module or 4710 Appliance can be used. When deploying the Cisco ACE Module, technologies such as the Cisco Catalyst 6500 Virtual Switching System (VSS) 1440 with multichassis EtherChannel can help improve network resiliency by delivering network convergence in less than a second in the event of failure. Cisco ACE 4710 Appliances connect to the aggregation layer using a PortChannel and are enabled for high availability. The Cisco ACE Module or 4710 Appliance load-balances virtual desktop requests to the connection broker while providing session persistence. The main features implemented on the Cisco ACE for the Cisco Virtual Workspace system are:

- Load balancing of the connection broker
- Health monitoring of the connection broker
- Session persistence based on client IP address

The Cisco ACE periodically checks the health of the connection broker. Using the probe information, Cisco ACE determines if the connection broker can service the user's request with the best performance and availability.

To access a virtual desktop, the user connects to a virtual IP address configured on the Cisco ACE. Then the Cisco ACE forwards the request from the user to the connection broker. Cisco ACE supports several session persistence mechanisms between the client and the connection broker so that a particular client session is always directed to the same server. The Cisco ACE is configured to perform session persistence based on the client IP address. If proxy servers are used, Cisco recommends enabling JSESSIONID cookie persistence.

When deploying a Cisco ACE with connection brokers, Cisco recommends either tunnel or direct mode, or a combination of the two methods.

- In direct mode, an endpoint establishes a connection to the virtual desktop instead of traversing the connection broker for all display protocol activity. Direct mode has many advantages, which include significantly less load (CPU, memory, and network traffic) on the connection broker. The connection broker responds to the initial first-phase connections from the endpoint, which is a very low-resource-intensive operation. Subsequent data passes directly between the endpoint and the agent running on the virtual desktop. Direct mode has some disadvantages as well. Without comprehensive security policies, an endpoint may connect directly to the virtual desktop, bypassing the connection broker.
- In tunnel (proxy) mode, an endpoint establishes a connection to the connection broker for all phases of communication, including the remote display data. Tunnel mode has some advantages, which include tighter access control and policy enforcement and significantly more load (CPU, memory, and network traffic) availability on the connection broker.

Figure 22 shows the deployment of the Cisco ACE 4710 Appliance using direct mode. The Cisco ACE 4710 provides load balancing only for the connection broker. After the connection broker has assigned the user a virtual desktop, the display protocol bypasses the Cisco ACE 4710. The Cisco ACE is configured in one-arm mode. One-arm mode is preferred in a network enabled for Cisco Virtual Workspace. Because the display protocol is between the end user and the virtual desktop, Cisco ACE does not need to be in the middle of the transaction, which saves resources on the Cisco ACE for load balancing incoming end-user requests for an available desktop from the connection broker. Because Cisco ACE is deployed in one-arm mode, you need to configure Source NAT.

Figure 22 Cisco ACE Deployment in the Network

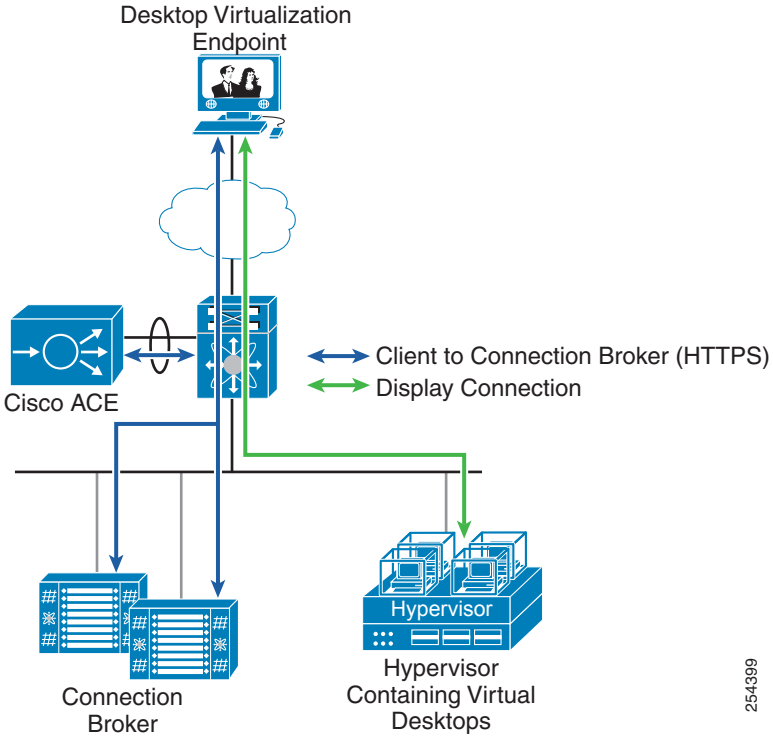


Table 8 provides detailed configurations for enhancing DV server availability. Cisco Virtual Workspace validated the ACE virtualized device environment by creating a virtual context named VMware DV load balancing. The virtualized device environment provided by the Cisco ACE is divided into objects called contexts. Each context behaves like an independent Cisco ACE load balancer with its own policies,

interfaces, domains, server farms, real servers, and administrators. Each context also has its own management. When Cisco ACE is deployed, the Admin context is used for managing and provisioning other virtual contexts.

Table 8 Desktop Virtualization Server Configurations

Function	Traffic Types	Description/Capability	Product Documentation
Admin Context Configuration	Allow only Telnet, SSH, Simple Network Management Protocol (SNMP), HTTP, HTTPS, or ICMP	The Admin context is used to configure the following: • Physical interfaces • Management access • Resource Management • High availability	Configuring Virtual Contexts
One-armed mode configuration	Recommend using One-armed mode for load balancing connection broker	Source NAT is required	Configuring One-Arm Mode
Configuring the Virtual Context for VMware DV	RDP and PCoIP	Load balancing HTTPS connection - Create a layer 4 class-map	Configuring Virtual Server Properties
Configuring Session Persistence		Session persistence based on the source IP address	IP Address Stickiness
Configuring Health Monitoring	HTTP Probe using regex checking	Regex statement inside HTTP Probe expect regex "View Administrator	Configuring Health Monitoring for Real Servers
Configuring the Load-Balancing Algorithm	Use Least-Loaded algorithm	Least number of connection in Serverpredictor leastconns	Load-Balancing Predictors
Configuration of Source NAT	SNAT needed so connection does not bypass Cisco ACE on the response back to the end-user	Use the virtual-address as NAT address	Cisco ACE NAT Configuration

VPN Termination

Many IT departments use VPN solutions to provide secure connections to end users accessing virtual desktops from branch offices or teleworker locations. With solutions such as the Cisco ASA appliances and the Cisco AnyConnect Secure Mobility Client, end users can connect to their virtual desktops through secure tunnels.

The Cisco ASA 5500 Series Adaptive Security Appliance is a purpose-built platform that combines best-in-class security and VPN services. The Cisco ASA is a VPN concentrator and a firewall and is the first point in the data center or campus edge at which all traffic is received. All VPN tunnels from VPN

clients such as Cisco AnyConnect or IPsec-enabled Cisco 800 Series Routers terminate here. Additional network security concerns and detailed information regarding the Cisco ASA appliances with the Cisco AnyConnect Secure Mobility Client are discussed in the Securing Cisco Virtual Workspace chapter.

Many IT departments use VPN solutions to provide secure connections to end users accessing virtual desktops from branch offices or teleworker locations. With solutions such as the Cisco ASA appliances and the Cisco AnyConnect Secure Mobility Client, end users can connect to their virtual desktops through secure tunnels.

The Cisco ASA 5500 Series Adaptive Security Appliance is a purpose-built platform that combines best-in-class security and VPN services. The Cisco ASA is a VPN concentrator and a firewall and is the first point in the data center or campus edge at which all traffic is received. All VPN tunnels from VPN clients such as Cisco AnyConnect or IPsec-enabled Cisco 800 Series Routers terminate here. Additional network security concerns and detailed information regarding the Cisco ASA appliances with the Cisco AnyConnect Secure Mobility Client are discussed in the Securing Cisco Virtual Workspace chapter.

Cisco Trustsec for Cisco Virtual Workspace (VXI) Smart Solution

Cisco Trustsec is a security solution to provide scalable secure architecture to enterprise networks built on existing identity based access layer security. Cisco Trustsec will build a network of mutually trusted network devices and strengthen the security of the enterprise networks. Cisco Identity Service Engine (ISE) is a component of Cisco Trustsec architecture which can provide contextual based access policy enforcement. Cisco Trustsec provides many enhancements to standard identity based access policy enforcement model including Security Group Access (SGA). After a user is authenticated, the resources access policy for authorization can be enforced using the SGA mechanism which is much more scalable than other authorization mechanisms. SGA allow capturing the identity of users and tagging each data packet called Security Group Tagging (SGT). If the network device is not capable of tagging the packet with SGT then the Trustsec solution will use Security Group Tag Exchange Protocol (SXP) to communicate the security tag to the next policy enforcement point. Once the SGT is inserted at the ingress of a data packet, the SGT can be used to enforce the access policy. The details of the Cisco Trustsec solution and its advantages can be found at: http://www.cisco.com/go/trustsec.

Cisco Virtual Workspace recommends using SGT for server segmentation and different access levels enforcement. The recommended access levels are Managed User/Managed Asset (Full Access), Managed User/Non-Managed Asset, Contractor, and Guest (Internet). For each access level a separate virtual desktop pool has to be created. Also the policy has to be enforced at multiple points in a Cisco Virtual Workspace environment and the recommended way to do this is by Security Group Access Lists (SGACL) or Security Group Firewalling (SGFW).

In Cisco Virtual Workspace deployments with Cisco Nexus 1000v switch, the SGT mappings to the servers can be provisioned in Cisco Nexus 1000v and send to the policy enforcement points like Cisco Nexus 7000 and Cisco ASA over SXP. The deployment details of the server segmentation using SGACL and SGFW can be found at the Trustsec Server Segmentation how-to guide at http://www.cisco.com/go/trustsec.

Bring-Your-Own-Device (BYOD)

With the proliferation of smart mobile devices and the trend of Bring-Your-Own-Device (BYOD), enterprises are facing a challenge to provide employees secure access to corporate resources on their personal devices. While BYOD has many advantages like increased work-force mobility and productivity, it poses a lot of challenges in supporting the various platforms and ensuring security of the corporate resources. Cisco BYOD Smart Solution 1.0 is a smart solution that allows users to connect, register and provision their personal devices for corporate use and for the access to the corporate network

in a secure way. Deploying Cisco BYOD Smart Solution to access Cisco Virtual Workspace will allow the mobile device users to access virtual desktops and enable the enterprise to offer a unified workspace to the employees with reduced cost and increased productivity.

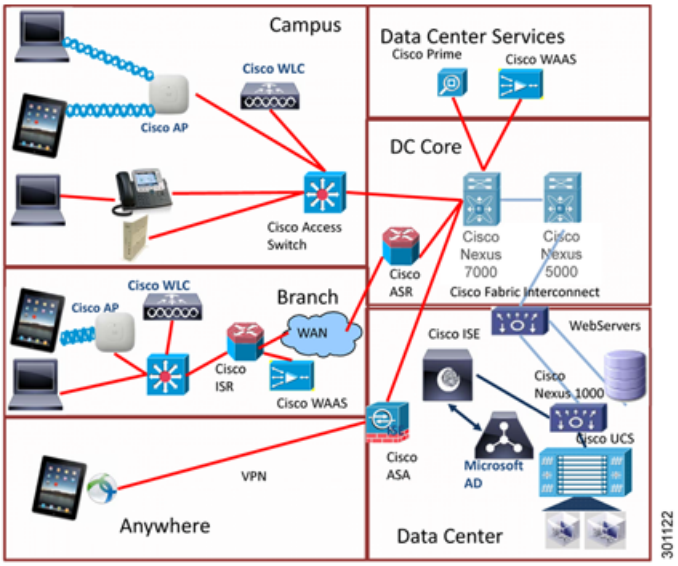
The fundamental building blocks for providing a BYOD solution are:

- Cisco Prime Infrastructure - the workspace delivery through network entities
- Cisco Identity Service Engine (ISE) and Cisco Trustsec Solution - contextual based policy enforcement
- Cisco Adaptive Security Appliances (ASA) and Cisco AnyConnect Mobile Client - Secure remote access to workspace
- Cisco Jabber - Collaboration on BYOD devices

The Cisco BYOD smart solution details can be found at http://www.cisco.com/go/byod.

Cisco Virtual Workspace virtual desktops can be delivered to the BYOD device as part of the workspace delivery when the Cisco BYOD components are added to the Cisco Virtual Workspace infrastructure. A typical Cisco Virtual Workspace deployment with BYOD components is depicted Figure 23 below.

Figure 23 Bring-Your-Own-Device



The constructs and mechanisms defined for workspace delivery and contextual based policy enforcement should be adapted from the Cisco BYOD smart solution for the delivery of Cisco Virtual Workspace to BYOD devices. The Microsoft Active Directory (AD) used in Cisco Virtual Workspace infrastructure for user identity should be added as an identity store in Cisco ISE so that once the user gets the access through the Cisco ISE, a desired access policy can be downloaded to the Wireless LAN Controller (WLC). Cisco Virtual Workspace smart solution recommends provisioning Cisco ISE to provide an authorization policy for the mobile devices just to have the access to Virtual Workspace infrastructure and Internet. Also every WLC deployed in campus and remote-sites should be provisioned to gather authorization policy from Cisco ISE as the BYOD device authenticates with that controller. Also Cisco ISE should be provisioned with authorization policies to deny access for the undesired BYOD devices.

The remote access from the BYOD device can be provisioned as outlined in the remote access section of the Cisco BYOD smart solution guide.

The collaboration services can be provided through Cisco Jabber application and the details of the architecture and design guidelines can be found at the Rich Media, Collaboration and User Experience chapter of this solution guide.

WAN Optimization for Virtual Workspace

The primary challenge in delivering HVD traffic across the WAN is helping ensure that performance is sufficient to meet end-user expectations. The Cisco Virtual Workspace system is designed to mitigate the effects of limited bandwidth, latency, and packet loss often associated with desktop virtualization and WANs.

Cisco WAAS optimizes WAN links through the use of intelligent caching, compression, and protocol optimization. When end users access the virtual desktops through the connection broker, Cisco WAAS compresses the response and then efficiently passes it across the WAN with little bandwidth use and high speed. Commonly used information is cached at the Cisco WAAS solution in the branch office and at the data center, which significantly reduces the burden on the servers and the WAN. Cisco WAAS accomplishes all WAN optimization functions using advanced compression using Data Redundancy Elimination (DRE), Lempel-Ziv (LZ) compression, and Transport Flow Optimization (TFO), which improves throughput and reliability for clients and servers in WAN environments. TFO optimizes TCP-based display protocols even if the display protocol traffic is encrypted and compressed. Some of these algorithms may not apply to optimization of the display protocols, but they are still useful in Cisco Virtual Workspace for the traffic flowing outside display protocols.

Cisco WAAS also enhances the performance and accelerates the operation of a broad range of chatty application protocols, such as Common Internet File System (CIFS), HTTP, SSL, and Messaging Application Programming Interface (MAPI) using application-specific accelerators (also called application optimizers [AO]) over the WAN. This optimization provides additional bandwidth for desktop virtualization and multimedia traffic.

Optimization of desktop virtualization protocols, including Remote Desktop Protocol (RDP) is achieved by applying DRE, LZ, and TFO techniques. Note that Cisco WAAS works only with TCP transport as of now and does not support UDP transport. Thus, protocols that use UDP transport cannot benefit from using Cisco WAAS. However, in combination with optimization that reduces the bandwidth consumed by chatty protocols, even UDP traffic can see noticeable benefits.

Cisco WAAS provides a holistic approach to support for a variety of print strategies, including centralized network printing through a print server, which can dramatically improve printing performance and reduce WAN data by using print-specific optimizations. Cisco WAAS provides print servers locally to branch-office users by running Microsoft Windows print services. When combined with USB redirection, Cisco WAAS offers optimization of the printing traffic redirected to locally attached peripherals (such as USB-connected printers) at the branch-office client device by reducing bandwidth utilization and mitigating WAN latency. Deploying Cisco WAAS on existing branch-office WAN connections also helps save bandwidth and allows both Cisco Virtual Workspaceand other traffic to be optimized.

Table 9 lists Cisco WAAS technologies applicable to various types of applications and traffic types seen in a Cisco Virtual Workspacesystem.

Table 9 Traffic Types

Traffic Types			**Cisco WAAS**		
Application	Protocol	Transport	TFO	DRE	LZ
VMware View 5.x	PCoIP (AES 128-bit encryption)	TCP & UDP port 4172	No	No	No

Traffic Types			Cisco WAAS		
VMware View 5.x	RDP	TCP port 3389	Yes	Yes	Yes
USB Redirection		TCP port 32111	Yes	Yes	Yes
Print	CIFS	TCP port 445	Yes	Yes	Yes
Multimedia Redirection (MMR)	MMR supported by VDI client on the endpoint	TCP port 9427	Yes	Yes	Yes

The Teradici PC over IP (PCoIP) protocol used by VMware View provides network and performance optimizations to address streaming multimedia issues. The PCoIP protocol uses TCP and User Datagram Protocol (UDP) over port 4172. The TCP port is used for session establishment and control, while the UDP port is used for the display protocol. The display protocol is encrypted with 128-bit Advanced Encryption Standard (AES) or 256-bit Salsa20 encryption. Cisco WAAS cannot optimize UDP packets and therefore just passes the packet through. Therefore Cisco WAAS can apply network optimization only for PCoIP session establishment and control session.

Application Aware DRE

In Cisco Virtual Workspace, the optimization over the WAN using application-aware DRE enhances the performance and scalability of Cisco WAAS appliance, and there by the user experience. Advanced compression using context aware DRE reduces disk space and memory. This feature is available in Cisco WAAS Release 4.4 or later. Context aware DRE supports two features: unidirectional DRE and single instance of data. These two features can greatly benefit the optimization of desktop session traffic which is unidirectional in nature and which can be optimized by Cisco WAAS.

Unidirectional DRE eliminates caching of unneeded data on the sending side Cisco WAAS. For specific traffic, cache data is only maintained on the receiving side Cisco WAAS where needed for de-compression. The sending side Cisco WAAS appliance does not keep this data as it is unlikely that the same data will flow in the reverse direction. This feature provides benefits to traffic that are asymmetrical in nature like desktop session and streaming traffic. If the same data flows in the reverse direction, this data will not be optimized on the first transmission but will be optimized on subsequent transmissions. For asymmetrical traffic, this event is unlikely and should not impact performance. Traffic classification indicates to Cisco WAAS which traffic requires unidirectional cache versus bidirectional cache.

Single instance of data further improves caching efficiency by eliminating multiple copies of the same data for multiple connections across multiple branches. This feature reduces memory and disk space usage resulting in greater scalability and performance of the Cisco WAAS appliance. It allows Cisco WAAS to optimize data across user sessions, and will benefit scenarios where the same desktop session data is being transmitted to multiple users (like an audio or video stream for example). This feature demonstrates the distinct advantage Cisco Virtual Workspace has of performing the optimization in the network (by offloading the task from the endpoints and servers). This feature is enabled with both unidirectional and bidirectional DRE.

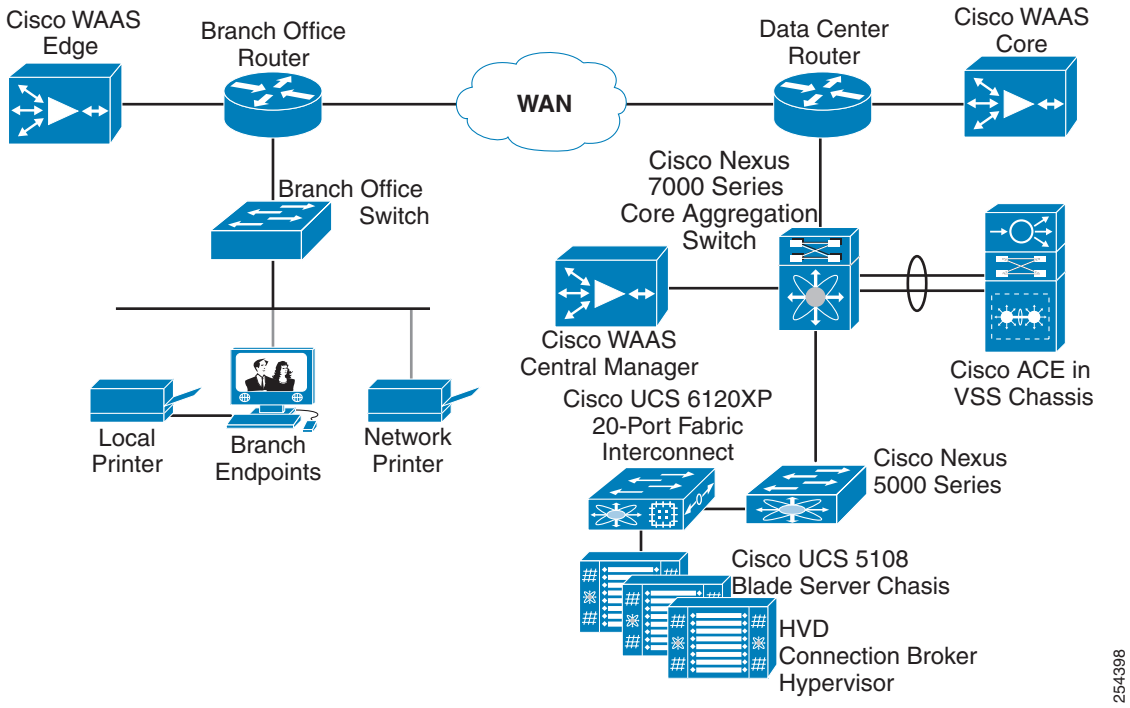
Deploying Cisco WAAS

Because Cisco WAAS is transparent in the network, some form of network interception is required to redirect relevant traffic types to the Cisco Wide Area Application Engine (WAE). Cisco WAAS supports many methods of network interception. The following methods are recommended with Cisco Virtual Workspace.

- **In Path Non-Redundant Interception**: The Cisco WAE is deployed physically between two network devices, most commonly between a router and a switch in a branch office. This allows all traffic traversing the network toward the WAN or returning from the WAN to physically pass through the Cisco WAE, thereby giving it the opportunity to optimize.
- **Off PathWeb Cache Communication Protocol Version 2 (WCCPv2)**: Cisco WAAS devices support WCCPv2, which provides an off-path but virtually inline deployment. With WCCPv2, Cisco WAE devices are deployed as appliances (nodes on the network and not physically inline) on the network. WCCPv2 provides scalability to 32 Cisco WAE devices in a service group, load balancing among Cisco WAE devices, and fail-through operation if all WAE devices are unavailable. It also allows the administrator to dynamically add or remove WAE devices in the cluster with little to no disruption. This is the preferred deployment model for Cisco WAAS in the Cisco Virtual Workspace.
- **Off Path with Cisco Nexus 1000v and vPath**: Cisco WAAS not only supports WCCP but Cisco vPath based on VXLAN. Rather than touching the existing transport infrastructure, Cisco vWAAS may be deployed in the virtualization infrastructure to intercept inbound display protocols and outbound desktop protocols. When deployed with vPath, vWAAS can use the programmatic interface to push pass through traffic into the Cisco Nexus 1000v switch.
- **In Path and Off Path WCCP Redundant AppNav**: The Cisco Application Navigation (AppNav) Controller may be deployed off path using WCCP or in path using integrated NICs designed to cluster WAAS. When deployed off path, the AppNav Controller (ANC) acts as a WCCP client to register with Cisco routers and switches to collect interesting traffic using a simplified WCCP configuration. The AppNav Controller then intelligently clusters the WAAS Service Nodes (SN) more intelligently than is possible with WCCP including the ability to target specific core boxes to branches and applications. When deployed in path, the AppNav Controller NIC offloads much of the processing for uninteresting pass through traffic and integrates well into the network to retain high availability leveraging exsting routing and switching protocols.

In the sample branch-office setup shown in Figure 24, the Cisco WAE appliance is connected to a local router, typically a Cisco Integrated Services Router (ISR), and in the data center Cisco WAAS is connected to a data center WAN edge router, typically a Cisco ASR. The branch-office router intercepts the traffic from Cisco VXC and other supported endpoints and sends it to the attached Cisco WAE, which after optimization sends it out the data center WAN edge router. The data center WAN edge router intercepts the traffic and sends it to the Cisco WAAS core for optimization, from which the traffic reaches the data center.

Figure 24 Cisco WAAS Deployment in the Branch Network



Cisco WAAS Form Factors

Multiple form factors for Cisco WAAS are available for deployment and are listed in Table 20. Network designers can select any form factor independently from an interoperability perspective, unless noted otherwise. All the Cisco WAAS models will work with the data center or campus Cisco WAAS WAE headend appliance, and the choice of branch-office or data center equipment depends on the scale needed and the version of software chosen. Detailed Cisco WAAS sizing guidelines can be found at: http://tools.cisco.com/Cisco WAAS/sizing

Table 10 Cisco WAAS Form Factor

Cisco WAAS Form Factor	Main Characteristics	Design Considerations
Cisco WAAS appliance	• Available in multiple form factors. • Supports all features.	Appropriate for scale of data center and in campus
Cisco WAAS with Cisco ISR G2 or Services-Ready Engine (SRE)	• Provides router-integrated application acceleration. • Provides dedicated onboard processing, memory, and hard drive	• Managed with the Cisco WAAS central manager or Cisco Prime LMS. • Suited for Medium to large branches

Cisco WAAS Form Factor	Main Characteristics	Design Considerations
Cisco WAAS express	Runs natively on IOS with ISR G2 router. Cisco 1941, Cisco 2901, Cisco 2911, Cisco 2921, Cisco 2951, Cisco 3925, and Cisco 3945	• Applicable to small and midsize branch offices • Does not support application optimization • Requires Cisco WAAS 4.2.1 in the data center • Appropriate for WAN links T1, E1, and 3G or serial links
Cisco WAAS mobile	• Ha small foot print • Runs as a Microsoft Windows application	Requires Cisco WAAS Mobile Manager
Cisco Virtual Cisco WAAS (Cisco vWAAS)	• Is a virtual appliance • Runs in the data center • Support WCCP redirection	Recommended where Virtual Security Gateway is not used.

Cisco WAAS Central Manager

Cisco WAAS requires a central manager to manage the Cisco WAAS solution from a central point. The Cisco WAAS Central Manager resides on a Cisco WAE appliance. When the administrator applies configuration or policy changes to a Cisco WAE device or a group of Cisco WAE devices, the Cisco WAAS Central Manager automatically propagates the changes to each of the managed Cisco WAE devices. Each Cisco WAE device can be configured either as an application accelerator or a central manager. The best practice is to deploy both a primary and a standby central manager.

Configuring Cisco WAAS in a Cisco Virtual Workspace System

Table 11 provides the configuration information needed to configure Cisco WAAS on a Cisco Virtual Workspace network.

Table 11 Cisco WAAS Configuration Information

Function	Traffic Types	Description and Capability	Product Documentation Links
Configure the Cisco WAAS central manager	HTTPS		Introduction to the Cisco WAAS Central Manager GUI
Configure WCCP on the branch and data center router	Interception of interesting traffic	WCCP Supported IOS versions	Configuring Traffic Interception
Configure the application accelerators	VMware RDP	Disable RDP encryption	Configuring Application Acceleration
Configure the context aware DRE	VMware RDP, USB, and MMR	Enable unidirectional caching of desktop session traffic.	Configuring Context aware DRE

Function	Traffic Types	Description and Capability	Product Documentation Links
Configure the branch and Data Center Cisco WAE			Configuring Network Settings
Optimizing printing using Cisco WAAS	CIFS		Configuring and Managing Cisco WAAS Print Services

Configuring Traffic Interception for Cisco WAAS Using WCCP

WCCP, as discussed earlier in this chapter, is used to configure the router to intercept all "interesting" traffic and forward it to Cisco WAAS. WCCP services 61 and 62 direct the router to reroute traffic from the interface to the WCCP group. Service 61 redirects ingress traffic, and service 62 redirects egress traffic. Services 61 and 62 are both needed to redirect bidirectional traffic flow.

You should exclude the Cisco WAE subnet from interception because this configuration uses a single interface to intercept incoming and outgoing packets. The interception exclusion is required because the router does not differentiate between traffic from the Cisco WAE for the client or server. Traffic from the Cisco WAE should not be redirected again by the router because doing so will create a loop.

Configure WCCP interception service 61 on the ingress interface and service 62 on the egress interface. All ingress and egress packets from the interface are forwarded to the Cisco WAE for optimization.

For more information, please read the Cisco WAAS section in the WAN deployment guide at http://www.cisco.com/en/US/docs/solutions/Enterprise/Borderless_Networks/Smart_Business_Architecture/February2012/SBA_Mid_DC_DataCenterDeploymentGuide-February2012.pdf.

Cisco WAAS appliances and Cisco vWAAS are expected by design to use the entire available RAM on the device or virtual machine. This condition is normal, and you should configure the low-memory alert with a higher threshold value to avoid unnecessary messages.

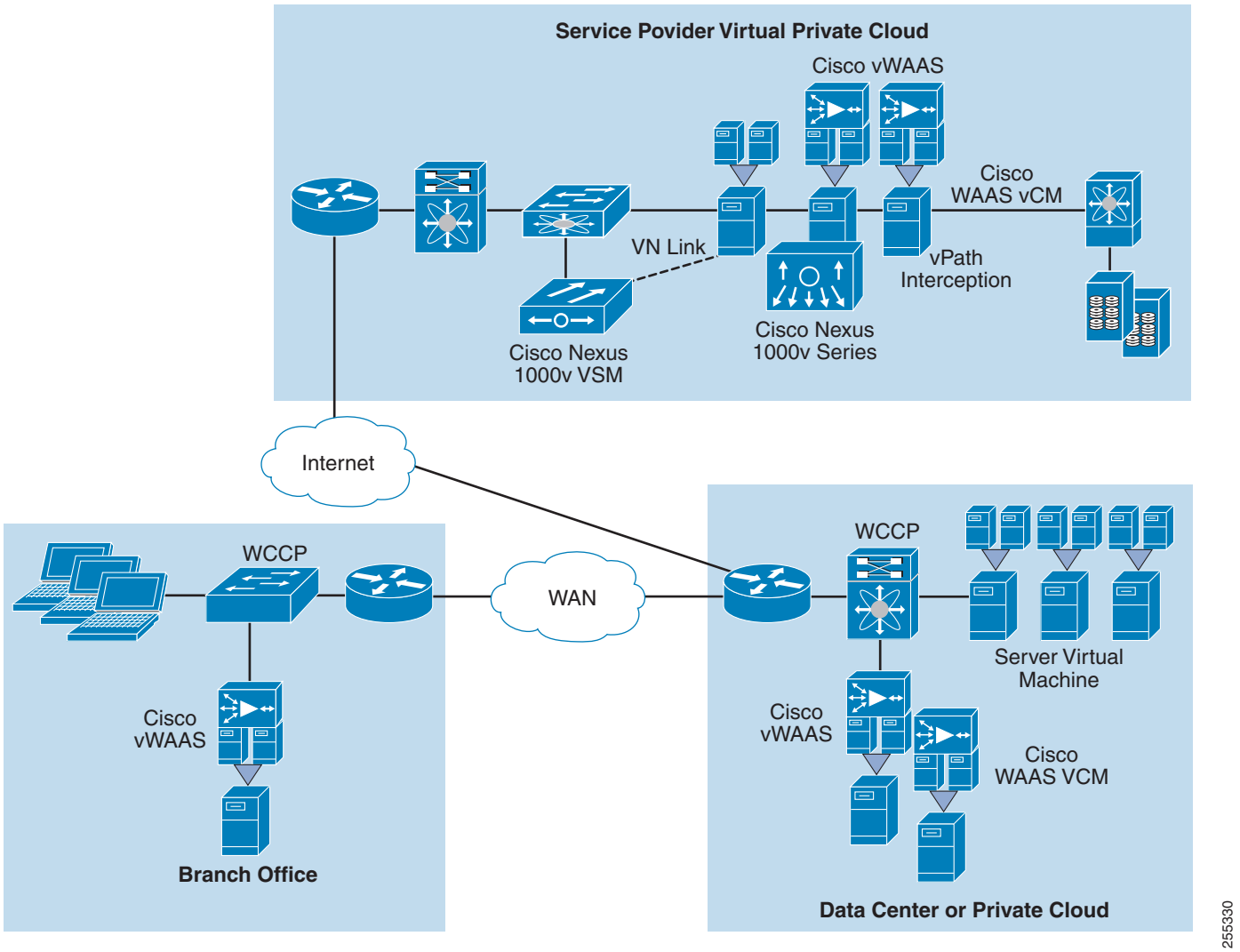
Cisco VSG and vWAAS are both virtual appliances that require a Cisco Nexus® 1000V Series Switch to function. Both these virtual appliances need to intercept traffic on the Cisco Nexus 1000V Series Switch and are not simultaneously supported as of this writing. As a workaround, if Cisco VSG is installed in the data center, then the Cisco WAAS appliance form factor can be used. Cisco WAAS appliances in the data center typically intercept traffic at the data center edge and do not require the Cisco Nexus 1000V Series for operation.

When Cisco vWAAS is used in the data center, the traffic needs to pass through the Cisco ASA at the data center edge. To allow Cisco ASA Software to recognize, permit, and secure traffic optimized by Cisco WAAS, a new CLI is provided. The new CLI should be enabled as part of the class modular policy framework (MPF) to facilitate Cisco ASA Software interoperability with Cisco WAAS.

Virtual Machine–Based Network Optimization: Cisco vWAAS

Cisco vWAAS is a virtual appliance that accelerates business applications delivered from private and virtual private clouds. Cisco vWAAS runs on Cisco Unified Computing System™ (Cisco UCS®) servers and the VMware vSphere hypervisor, using policy-based configuration in the Cisco Nexus 1000V switch. Cisco vWAAS can be associated with application server virtual machines as these are instantiated or moved. With Cisco vWAAS, cloud providers can rapidly provision WAN optimization services with little to no configuration or disruption (Figure 28).

Figure 25 Cisco vWAAS



Cisco vWAAS supports two deployment options:

- WAN edge deployment with out-of-path interception
- Deployment deep in the data center with virtual path (vPath) interception

The WAN edge is the traditional location for WAN traffic interception to be optimized. The out-of-path model uses WCCP to intercept traffic. In this model, the hypervisor hosts with Cisco vWAAS virtual machines connect to WCCP-enabled switches or routers. Multiple Cisco vWAAS virtual machines can operate in a cluster, optimizing all traffic intercepted by the router. The Cisco vWAAS virtual machines can be spread across single or multiple servers, and both physical and virtual Cisco vWAAS appliances can be mixed in a cluster.

In the deep data center deployment model, it is recommended to place Cisco vWAAS VMs next to server VMs in the same Hypervisor host. Using the policy-based virtual services of the Cisco Nexus 1000v Series switch, Cisco vWAAS can be deployed on a per-application or per-server-group basis. vPath interception in the Cisco Nexus 1000v Series intercepts all traffic to and from these servers and forwards it to the Cisco vWAAS virtual machine for optimization.

Cisco vWAAS requires storage access to store the data redundancy elimination byte cache and the Common Internet File System (CIFS) cache. In the data center, it supports either direct-attached storage or remote shared storage in a SAN. Cisco recommends the use of the SAN option to enable advanced hypervisor features such asVMWare vMotion, VMware Storage vMotion ®, and VMware High Availability.

WAN Stability and Path Optimization

In a Cisco Virtual Workspace deployment, the availability and efficiency of HVDs can be affected by suboptimal routing and path selection. Cisco recommends using Performance Routing (PfR) to reduce the effects of network instability and link quality changes. PfR improves application performance by selecting the best path across the WAN. PfR uses network metrics such as application reachability, delay, loss, and jitter to help select the best path based on application needs. It measures the network performance and dynamically reroutes the traffic when the metrics do not meet requirements. PfR thus provides path optimization and advanced load balancing for Cisco Virtual Workspace traffic over the WAN. PfR and Cisco WAAS provide different forms of WAN optimization, and when used together they can complement each other to provide a superior user experience.

PfR has two logical components: a master controller and a border router. The master controller acts as a central processor and data collection point for PfR and runs on the router or in standalone mode. The master controller gathers network metrics from all the border routers and determines whether traffic classes adhere to configured policies. On the basis of this determination, the master controller can instruct the border router to stay on the current WAN link or change paths. Path selection is performed on the border router by using route injection or dynamic policy-based routing injection. In small deployments, the master controller can be deployed on the border router itself.

PfR operations have three logical steps:

1. Identify the traffic flow and understand its requirements.

   You do this by using a combination of traffic class, port, protocol, src/dst prefix information. For Cisco Virtual Workspace traffic TCP/UDP ports can be used to identify the display protocol traffic.

2. Measure and monitor network performance.

   You do this either passively by using Cisco Netflow based monitoring on the BR or actively by sending network probes to measure health of the WAN link. Both methods can be used together for better accuracy.

3. Determine routing and path selection.

   You do this by either using traffic class optimization or link optimization. For Cisco Virtual Workspace traffic, traffic class optimization is recommended since the aim is to improve the performance of a specific traffic type. Link based optimization optimizes based on link cost and will result in optimum load balancing of multiple traffic flows. Link cost optimization is the primary focus for this approach.

Cisco Virtual Workspace design recommendations for PfR are as below:

- Enable traffic-class performance optimization of PfR using display protocol TCP and UDP ports. These ports are listed earlier in this chapter.

- Because Cisco Virtual Workspace traffic is highly interactive, an average one-way delay target across the WAN of 100 ms should be used when PfR profiles and policies are defined.

The PfR function is available on Cisco ISR and ASR platforms. In Cisco Virtual Workspace branch-office environments, the border router function should be deployed using the Cisco ISR in the branch office and the Cisco ASR in the data center. The master controller should be enabled on the data center Cisco ASR.

Note PfR does not work with Cisco WAAS clustering since Cisco WAAS appliance that serviced the packets before a PfR triggered route change will not service the subsequent packets.

Cisco PfR can be deployed with Cisco WAAS and DMVPN. Refer to the links below for interoperability details.

Enhancing the WAN Experience with PfR and Cisco WAAS:

http://www.cisco.com/en/US/prod/collateral/iosswrel/ps6537/ps6554/ps6599/ps8787/prod_white_paper0900aecd806c5077.html

Using PfR in Redundant VPN networks:

http://www.cisco.com/en/US/products/ps8787/products_ios_protocol_option_home.html

PfR deployment involves multiple logical steps. Each phase and its relevant documentation is shown in Table 12.

Table 12 PfR Deployment Steps

PfR Deployment Step	Configuration Reference
Traffic Profiling: How to identify/define traffic flows.	http://www.cisco.com/en/US/docs/ios/oer/configuration/guide/pfr-profile.html
Measure: Measuring network metrics for traffic flows above.	http://www.cisco.com/en/US/docs/ios/oer/configuration/guide/oer-trf_lnk_util.html
Policies: Mapping network metrics to thresholds and triggers. Deciding for action.	http://www.cisco.com/en/US/docs/ios/oer/configuration/guide/oer-cfg_ap_polcy.html
Control: Take route control actions based on policy decisions.	http://www.cisco.com/en/US/docs/ios/oer/configuration/guide/oer-trf_rte_ctl.html
PfR deployment step	Configuration reference

Quality of Service (QoS)

Providing QoS in desktop virtualization deployments can be a challenge. Protocols are often proprietary and encrypted. Applications are encapsulated in the desktop virtualization display protocol and may be difficult to differentiate. Some traffic (for example, video) may be transported in a separate channel, but implementations may be proprietary and not suitable for traditional QoS approaches. Cisco Virtual Workspace relies on differentiated services code point (DSCP) marking to provide QoS for such traffic.

Table of QoS Markings

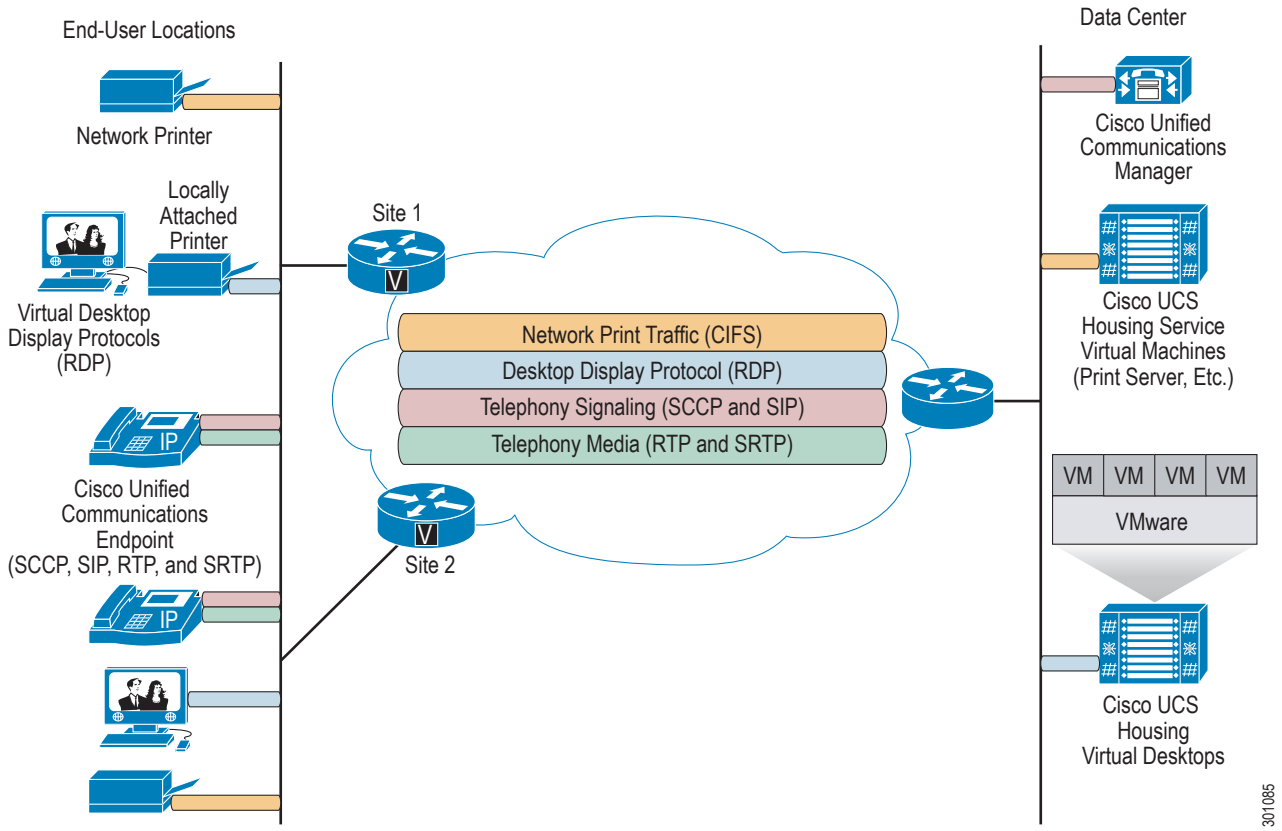
The following sections describe how to recognize, mark, and optimize the protocols used in desktop virtualization environments. Table 13 and Figure 26 shows the protocols in a typical Cisco Virtual Workspacesystem.

Table 13 QoS Markings for Protocols Used in Virtual Workspace

Protocol	TCP/UDP Port	DSCP/CoS Value
Desktop Virtualization Protocols		
Remote Desktop Protocol Version 7 (RDPv7)	TCP 3389	DSCP af21 and CoS2
PC Over IP (PCoIP)*	TCP & UDP 50002 TCP 4172	DSCP af21 and CoS 2
USB redirection (PCoIP*)	TCP 32111	DSCP af11 and CoS 1
Multimedia redirection (MMR)	TCP 9427	DSCP af31 and CoS 4
Other Protocols used within Cisco Virtual Workspace		
Network-based Printing (CIFS)	TCP 445	DSCP af11 and CoS 1
• Unified communications signaling (Skinny Client Control Protocol[SCCP]) • Unified communications signaling Signaling (Session Initiation Protocol[SIP]) • Unified communications signaling Signaling (CTI)	• TCP 2000 • TCP 5060 • TCP 2748	DSCP cs3 and CoS 3
Unified communications signaling Media (RTP(Real-time Transport Protocol), and Secure RTP [SRTP])	UDP 16384 to 32767	DSCP ef and CoS 5

Note *PCoIP is moving away from using port 50002 and will move to using 4172 for the future

Figure 26 Overview of Protocols within Cisco Virtual Workspace Network



Data Center QoS

Classification and marking should be performed in the data center as close to the application servers and virtual desktops as possible. If proper QoS policies and queuing priorities are in place, these markings should be maintained and traffic handled appropriately throughout the network.

Markings

Many applications do not mark traffic with DSCP values. For even those that do, the marking may not be appropriate for every enterprise's priority scheme. Therefore, hardware-based classification (using a Cisco Catalyst® or Cisco Nexus® Family switch) should be used. In testing, marking was implemented on a Cisco Nexus 1000V Switch whenever possible. DSCP values and associated ports are based on Table 25. For more information and examples, see the Cisco Virtual Workspace (VXI) Smart Solution As-Built Reference Guide.

Queuing for Optimum Cisco UCS Performance

Cisco recommends marking intra-data center traffic so that desktop virtualization traffic does not starve the Cisco UCS for storage resources. The Cisco UCS is uniquely equipped with the capability to mark and queue outbound packets to improve its performance with other data center entities (such as storage). Cisco UCS queuing strategies are discussed in the Cisco UCS GUI configuration guide, at

http://www.cisco.com/en/US/docs/unified_computing/Cisco UCS/sw/gui/config/guide/1.3.1/UCSM_GUI_Configuration_Guide_1_3_1.pdf

and Cisco Unified Communications Manager systems guide, at

http://www.cisco.com/en/US/docs/voice_ip_comm/cucm/admin/8_0_2/ccmsys/accm.pdf

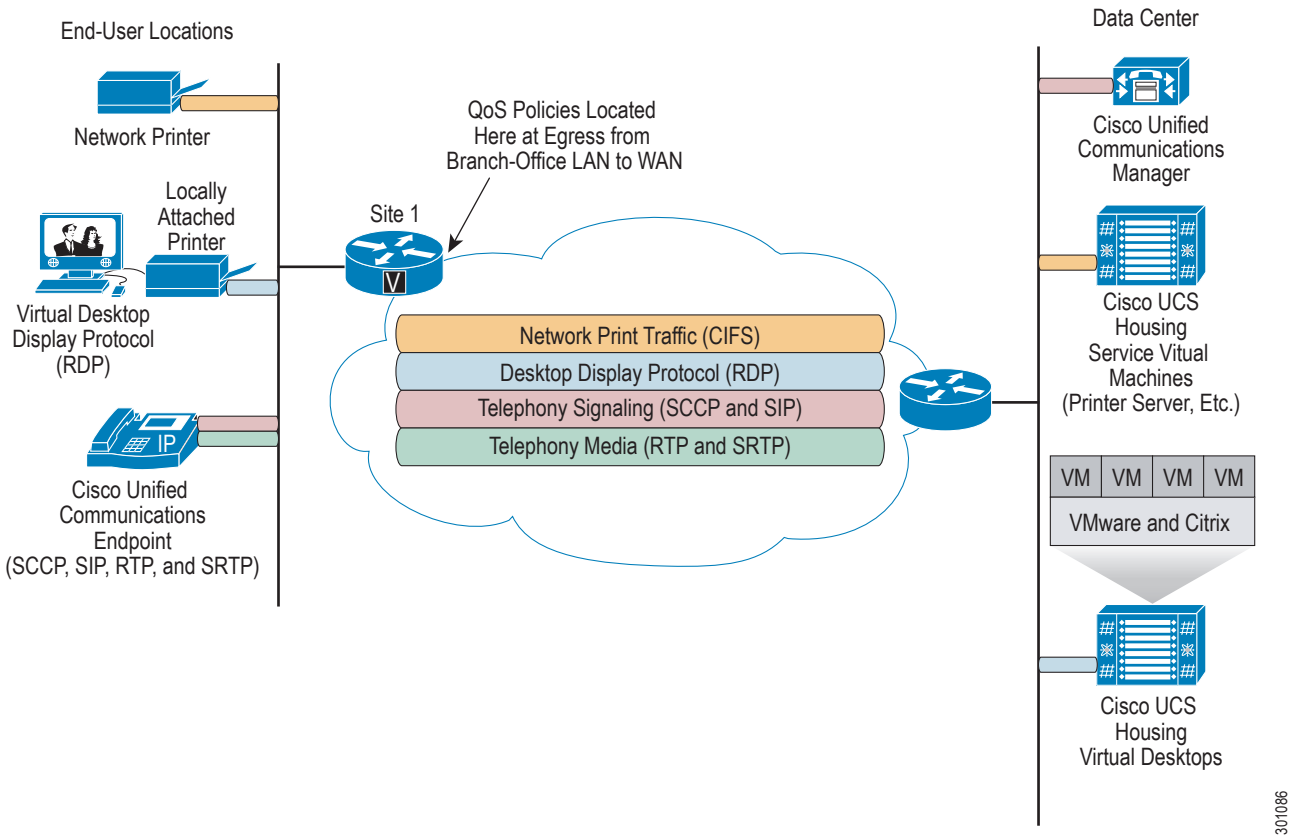
Network QoS

Network devices are responsible for routing and queuing the traffic based on the priority of the traffic decided at the branch office or data center. Traffic is prioritized and forwarded based on the DSCP markings. Although remarking can take place within the network, if the endpoint locations and data center deploy the suggested configurations, remarking likely will be unnecessary.

In a campus network, bandwidth-related resource contention should be minimal. Bandwidth contention is most likely at the egress point from the high-speed connections of the branch-office LAN to the slower-speed links of the WAN. Service policies that constrain the amount of bandwidth dedicated to a given protocol are defined and applied at this point. These same queuing and bandwidth configurations can be placed anywhere that Cisco VXC and other supported endpoints are concentrated, to enforce the appropriate response in the event of traffic congestion (Figure 27).

To view an example of the bandwidth service policies, see the Cisco Virtual Workspace (VXI) Smart Solution As-Built Reference Guide.

Figure 27 Location of Service Policies for QoS



Endpoint QoS

The Cisco VXC 6215 does not mark desktop session traffic. Therefore, the same marking that was performed on the Cisco Nexus 1000V Series Switch in the data center for outbound desktop virtualization traffic must be performed at the branch office on behalf of the endpoints for the traffic returning to the data center virtual machine. The configuration on the branch-office switch will look similar to the configuration presented in the Cisco Virtual Workspace (VXI) Smart Solution As-Built Reference Guide.

Admission control is another means of contention reduction that is required by some applications to enforce QoS. Using the Cisco Jabber™ platform for hard-phone control of an IP phone, voice-over-IP (VoIP) traffic is supported outside the display protocols. This traffic is admitted into the priority queue. Admission control is required to prevent this priority queue from flooding.

Cisco Unified Communications Manager supports two kinds of admission control: location based and Resource Reservation Protocol (RSVP) based. Although location-based communications admission control is extensively deployed in many Cisco Unified Communications Manager installations, it cannot use redundant links or a mesh topology. RSVP-based admission control is a network-aware admission control system. With the help of proxies, called RSVP agents, that run in Cisco 2800, 2900, 3800, and 3900 Series ISRs, RSVP is a much more robust solution for controlling hard phones in Cisco Virtual Workspace.

For information about implementing RSVP-based admission control in Cisco Unified Communications Manager, please see Part 2, Chapter 9, in the Cisco Unified Communications Manager Release 8.5(1) System Guide.

Cisco VXC Client QoS Considerations

Cisco Virtual Workspace validation includes the Cisco VXC 4000, a unified communications software appliance, and the Cisco VXC 6215, a unified communications–enabled thin client that uses a single network interface to transmit both unified communications (voice and video media streams) and desktop virtualization traffic using the assigned single IP address and VLAN. Because the endpoint separates these traffic types into separate streams, you should perform QoS marking on the access switch. For specific QoS guidance about how to integrate these endpoints into a Cisco Virtual Workspace deployment, please refer to the Cisco Virtual Workspace Clients chapter of this guide.

Managing the Cisco Virtual Workspace Network

Cisco Virtual Workspace customers use a broad portfolio of management tools to administer these systems (see the Management and Operations chapter). This section provides design guidance for deploying and using the Cisco Network Analysis Module (NAM) and Cisco NetFlow, which are essential for managing a Cisco Virtual Workspace network.

Cisco Network Analysis Module

The Cisco NAM monitors traffic, collects packet traces, and generates historical reports. It provides comprehensive traffic analysis and insightful troubleshooting information in real time, which can be used to increase the efficiency of the network and applications. Cisco NAM is used to validate QoS planning assumptions and help ensure that service levels are met, and to assess the user experience with transaction-based statistics.

Cisco NAM includes an embedded, web-based traffic analyzer GUI for accessing configuration menus. Easy-to-read performance reports about web, voice, and video traffic use are accessed through the browser. Cisco recommends dedicating a separate management interface on the Cisco NAM appliance for accessing the GUI.

The Cisco NAM probe placement varies according to the task being performed:

- Any location that is the ingress or egress point of a logical network boundary (aggregation layer, core, or campus edge) can offer valuable insights into the network activity within that partition and is usually a good choice for Cisco NAM deployment.
- Cisco NAM is often located in the data center core to monitor sessions connecting to critical server farms, such as the connection manager. This location allows monitoring of sessions initiated by end users located in all parts of the enterprise network.
- You can place Cisco NAM in a branch office to troubleshoot a particular user group connecting across a WAN. The WAN edge is the optimal location for monitoring sessions from multiple branch offices.

Cisco NAM comes in several form factors to facilitate deployment in different network locations In the branch office, Cisco NAM can be deployed as a network module on a Cisco ISR branch-office router, or in the core network as a blade on a Cisco Catalyst 6000 Series core switch or as a separate appliance (Cisco NAM 2204 or 2220 Appliance.) Cisco NAM can also be deployed as a software module on the Cisco Nexus 1000V or the Cisco WAAS device. In the data center core, you should deploy a separate Cisco NAM appliance, which uses dedicated hardware to provide the required level of performance. Refer to the Cisco NAM deployment guide: [http://www.cisco.com/en/US/partner/prod/collateral/modules/ps2706/white_paper_c07-505273.html] for a comprehensive discussion of considerations for Cisco NAM deployment.

Use Cisco NetFlow data export from a remote router to the Cisco NAM for network traffic use reports. The Cisco NetFlow data can be exported from a Cisco WAAS appliance (to monitor the traffic across the WAN), router, switch, or Cisco ASA appliance. For example, NetQoS flow data from Cisco WAAS can provide information about application latency. The traffic use reports can help identify traffic types that will benefit from QoS policies and Cisco WAAS optimization. These measurements can also be used for growth forecasting and planning.

Using Cisco NAM to Troubleshoot Virtual Desktop User Sessions

The main use of Cisco NAM is to troubleshoot and analyze a user session that fails to set up correctly or a user session with poor quality. To display a specific conversation, a packet capture can be filtered based on the endpoint, desktop controller, or virtual desktop network address. It can also be filtered based on VLAN, protocol (port), or DSCP marking in the stream.

Note that Cisco NAM does not currently decode desktop virtualization protocol packets, because the protocol format is proprietary and may be encrypted. Cisco NAM can identify and label the stream based on the port numbers used.

Table 14 lists the protocols and ports used by desktop virtualization sessions. This information can be used to identify the user streams in a packet capture.

Table 14 Protocols and Ports Used by Desktop Virtualization Sessions

DV Protocol	Protocol	Port
RDP	TCP	3389, 32111(USB) and 9427(MMR)
PCoIP	UDP/TCP	4172, 32111(USB) and 9427(MMR)

Switched Port Analyzer (SPAN) sessions on a core switch capable of monitoring multiple source ports, such as the Cisco Catalyst 6000 Series or Cisco Nexus 7000 Series Switches, should be used for packet captures. Use a SPAN port (preferably 10 Gigabit Ethernet) on the data center core switch (Cisco Nexus 7000 Series) to monitor all traffic to and from the data center servers. Note that there is a limit of 2 SPAN sessions on the switch.

Protocol/Port Information

The VMware View session setup (desktop virtualization endpoint to VMware View Manager stream) uses TCP (port 80) and Transport Layer Security Version 1 (TLSv1; port 443). The communication between VMware View Manager and View Agent uses TCP (port 4001). The VMware vCenter server's VMware vSphere client and software development kit (SDK) interface uses HTTP (80) and HTTPS (443). VMware vCenter communication to the VMware ESXi host uses a TCP (port 902) stream. The VXC Manager defaults to HTTP (80) and HTTPS (443) for communication with desktop virtualization endpoint agents. The administrator should make sure that these ports are open on all firewall devices.

Packet capture of the traffic from a virtual desktop running Cisco Unified Personal Communicator with an active voice call does not include any Real-Time Transfer Protocol (RTP) traffic, since the Cisco Unified Personal Communicator client is running in desk phone control mode. The UC signaling traffic uses the following ports: SCCP TCP 2000 and SIP TCP/UDP 5060.

The desktop packet capture does include Cisco Jabber client communications with Cisco Unified Presence and Cisco Unified Communications Manager through the Computer Telephony Integration (CTI) protocol. The packet capture of a Cisco IP Phone that is co-located with the desktop virtualization endpoint does include the RTP traffic. Cisco NAM can also collect RTP metrics used for audio-quality measurements from the Cisco Unified Management Suite.

Cisco NetFlow

Cisco Netflow can be used to monitor, troubleshoot, and perform capacity planning for a large-scale Cisco Virtual Workspace deployment. Cisco Netflow data can be collected for multiple Cisco Virtual Workspace traffic flows. This data can be aggregated and a report on use can be generated. Reports can be used to determine peak rate, average rate, and peak volume per day or week, average volume per day or week for different traffic types. The data can be used for network capacity planning when adding additional network interfaces (such as WAN links) or network elements (routers and switches) to the network. Other useful tasks are to identify top Cisco Virtual Workspace users, setup alerts when certain traffic thresholds are reached, and obtain traffic distribution reports for services accessed via Cisco Virtual Workspace sessions (for example, monitoring virtual desktop traffic).

- Cisco Netflow can be enabled on the routers (such as Cisco ISRs) located in the branch office to gain visibility into desktop protocol and session traffic; on aggregation routers (Cisco ASRs, 7200 Series Routers, and Cisco WAAS) located on the WAN edge to monitor desktop protocol and session traffic; on the enterprise core switches (Cisco Nexus 7000 or Cisco Catalyst 6500 Series) located in the data center to monitor desktop protocol and session traffic and virtual desktop traffic; or on the virtual switch (Cisco Nexus 1000v Series) to monitor desktop protocol and virtual desktop traffic.
- Enabling Netflow on all network elements and aggregating the data will result in an information overload unlikely to provide useful data. It is better to enable Netflow selectively on network elements located at specific points in the network (based on the traffic pattern/path of the session to be monitored) and collect the data using distributed or centralized set of collectors.
- A simple strategy for identifying Cisco Virtual Workspace session traffic is to use the IP address (of the thin client, virtual desktop, or desktop controller), protocol type (TCP or UDP), port (PCoIP, RDP, HTTP, or HTTPS), and type of service information in the Cisco Netflow records.

Specific traffic flows of interest include Cisco Virtual Workspace session traffic between endpoint and desktop controller, desktop protocol traffic between endpoint and the virtual desktop, virtual desktop to desktop controller traffic, the virtual desktop traffic to the servers and the Internet, as well as Cisco Unified Personal Communicator traffic patterns including CTI and RTP traffic.

Cisco Netflow Version 9 export format is a flexible and extensible means for carrying Cisco Netflow records from a network node to a collector. Cisco Netflow Version 9 has definable record types and is self-describing, which allows for easier Netflow Collection Engine configuration.

Note Cisco Netflow records are exported using UDP on a user configured port, on a best effort basis and use available network bandwidth. Cisco Netflow records can be reliably exported using Stream Control Transmission Protocol (SCTP) on certain routers and switches. Consult Cisco router documentation for more information.

Note Cisco Netflow has little negative effect on the CPU of the network element on which Cisco Netflow is enabled. There are techniques to reduce CPU, memory, and storage impact on the router, collector, and network including adjusting aging timers, using sampled NetFlow, and the use of Flow masks. Additional methods include using aggregation schemes, filters, data compression, and larger collection bucket sizes. One method to reduce network bandwidth usage is to locate the collector and router on the same LAN segment. Refer to Cisco Netflow documentation for more details on the use of these techniques.

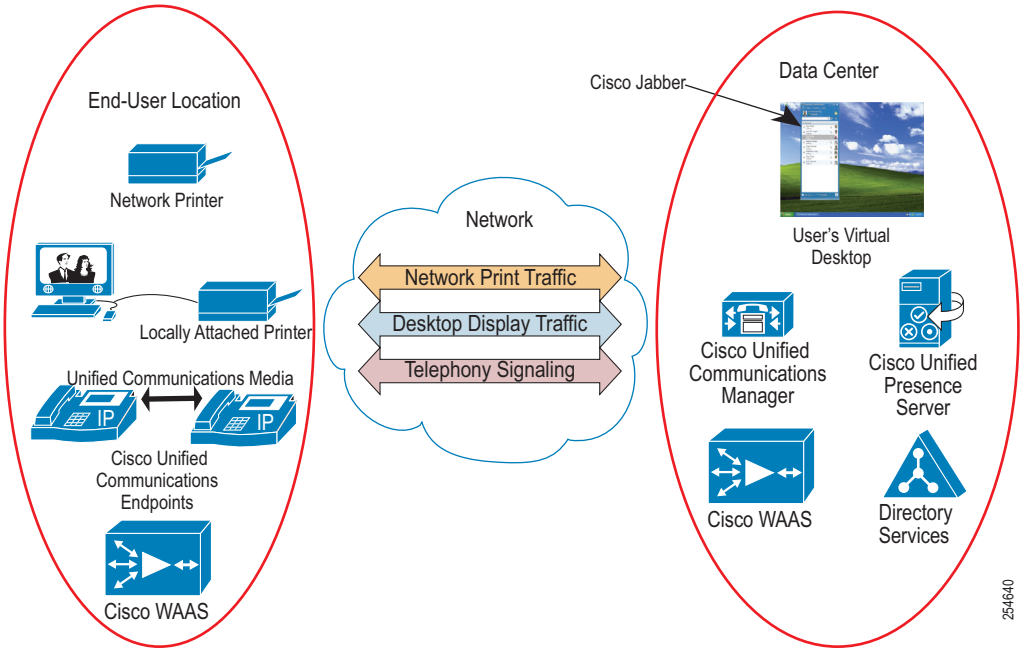
Please refer to the Cisco Virtual Workspace (VXI) Smart Solution As-Built Reference Guide at: http://www.cisco.com/en/US/docs/solutions/Enterprise/Data_Center/VXI/configuration/VXI_Config_Guide.pdf for a sample Cisco NetFlow configuration on Cisco IOS Software routers and Cisco Catalyst Family switches.

Cisco Virtual Workspace Clients

The Cisco® Virtual Workspace (VXI) Smart Solution system includes the Cisco VXC 6215, which offers desktop virtualization and collaboration services. Cisco Virtual Workspace (VXI) Smart Solution also integrates Cisco Unified Communications endpoints to deliver voice and video communications. Cisco Jabber™ provides the functional integration between the virtual desktop and the unified communications endpoint by allowing user control of a desk phone, or a unified communications-enabled VXC 6215 through its deployment on the user's hosted virtual desktop (HVD).

Figure 28 provides an overview of endpoint data flow in a virtual desktop environment.

Figure 28 General Endpoint Data Flow Diagram



Note Please consult the Scaling and High Availability chapter and the Cisco Virtual Workspace (VXI) Smart Solution 2.7 Release Notes for more information about Cisco Jabber validation. Please consult the product documentation for these unified communications applications to verify support in a virtual desktop infrastructure (VDI) environment. Cisco recommends that organizations fully validate these applications before deployment.

The Cisco Virtualization Experience Client (Cisco VXC) 6215 is a thin client shown in Figure 29.

Figure 29 Cisco VXC 6215 Thin Client

What is New in Release 2.7

The firmware release 9.0 and later for Cisco VXC 6215 supports the Virtualization Experience Media Engine (VXME) add-on. VXME allows Cisco Jabber on the virtual desktop to make voice and video calls through the Cisco VXC 6215.

Cisco VXC 6215 Thin Client

The Cisco VXC 6215 connects with VMware View version 4.6, 5.0, or 5.1 environment using PCoIP or RDP (versions 7.0 or 7.1) as display protocols. Along with the base firmware, the software on the Cisco VXC 6215 supports an add-on software known as the Cisco Virtualization Experience Media Engine (VXME) that delivers optimized unified communications voice and video capabilities. This add-on extends Cisco Jabber on the virtual desktop to enable users to collaborate through instant messaging, email, presence, and voice and video conferencing using a single desktop device. With the help of VXME, the thin client intelligently separates unified communications media traffic from display protocol traffic so that each media type can be managed and prioritized appropriately by the network. The combination of Cisco Jabber running in the user's HVD and VXME running in the thin client is designed to replace the user's desk phone and PC while maintaining a good user experience. This solution is well suited for campus workers and remote workers, or in branch locations where complete unified communications survivability is not required.

Note The Cisco VXC 6215 VXME add-on is not available for RDP based HVD sessions currently.

The Cisco VXC 6215 can operate in either the basic VDI mode or the VDI with Cisco Unified Communications integration mode. In the basic VDI mode, the Cisco VXC 6215 supports RDP 7, VMware View Client and works with VMware View versions4.6, 5.0, and 5.1. For the VDI with Cisco Unified Communications integration mode, the Cisco VXC 6215 requires VMware View Client and VMware View version 5.1, Cisco Unified Communications Manager Release 7.1.0 or later, and optionally, Cisco Unified Presence Server release 8.0 or later. When the Cisco VXC 6215 is in the VDI with Cisco Unified Communications integration mode, the user can access collaboration tools through Cisco Jabber running on the HVD, while using the audio and video capabilities of the local Cisco VXME. Note that the Cisco Unified Personal Communicator and Unified Communications Integration for Microsoft Lync running in the virtual desktop do not control the audio and video capabilities of VXME and may only be used for Presence and Instant Messaging applications, or to control a Cisco IP deskphone.

Note Cisco Unified Survivable Remote Site Telephony (SRST) is not supported in the current release of Cisco VXC 6215 with VXME. However in case the VXC 6215 loses contact with Cisco Unified Communications Manager or the HVD, it will maintain existing calls.

Note When deploying the thin client, verify that you are running the correct version (release 9.1 or later) of Cisco Jabber in the HVD so that it can recognize and control the Cisco VXME. There are other configuration requirements for the HVD and Cisco Unified Communications Manager.

The Cisco VXC 6215 requires an external power supply. It connects directly to the access switch, and queries the switch for both data and voice VLANs using Cisco Discovery Protocol (CDP). It places IP telephony traffic in the Voice VLAN, if available, and VDI traffic in the data VLAN. For more details, please refer to the Virtualization Aware Network chapter in this document.

The Cisco VXC 6215 end user cannot select which USB devices to redirect to the HVD and which USB devices to connect locally on the thin client. For use with Cisco VXME for unified communications applications, devices such as a USB camera and headset should not be mapped to the HVD. The mapping itself is controlled in-group policy definitions through VMware View administration. If certain devices are needed for applications on the HVD, you can have them passed explicitly to the HVD by specifying them in the Cisco VXC 6215 ini file. End users are also not allowed to customize the Cisco VXC 6215 OS. Only administrators using Cisco VXC Manager can customize the OS.

In general, any USB device or analog audio device used in conjunction with the local audio and video capabilities of the thin-client unified communications software appliance should not be redirected to the HVD. For more information, see the product documentation at http://www.cisco.com/go/vxc and the Cisco VXC 6215 administration guide at http://www.cisco.com/en/US/docs/voice_ip_comm/Cisco VXC/english/vxc_6215_1-0/admin/AdminGuide_VXCM.html#wp1074754.

The Cisco VXC 6215 uses a VMware PCoIP virtual channel to provide a transparent user experience between the HVD running in the data center and VXME running locally in the thin client. Also, through the display protocol virtual channel, the Cisco VXC 6215 receives user and profile login information from Cisco Jabber running in the HVD to register unified communications services and user profiles with Cisco Unified Communications Manager. This approach allows a single Cisco VXC 6215 device to be reused among multiple users, enabling an efficient shared desk environment. The thin client also uploads logs and diagnostics to Cisco Jabber running in the HVD so that the existing Cisco Jabber Problem Reporting Tool (PRT) can be used for troubleshooting.

VXME supports basic supplementary services (call transfer, call forwarding, hold and resume, and conferencing) plus the advanced services supported by Cisco Jabber installed on the HVD. All provisioned services are invoked and managed within Cisco Jabber. Video conferencing requires Cisco Unified Video Conferencing multipoint control units. Cisco TelePresence® Server provides telepresence interoperability between the Cisco VXC 6215 and a telepresence system. If necessary, a Cisco Media Experience Engine (MXE) may be used for telepresence interoperability.

The Cisco VXC 6215 supports a number of headset devices and webcams, which are listed in the product release notes. Please consult the Cisco Virtual Workspace (VXI) Smart Solution Release Notes for information about the limitations of scaling the thin-client environment to a large number of endpoints.

For configuration details for Cisco VXC, see: http://www.cisco.com/go/vxc.

Cisco VXC 6215 Media Termination Capability

The Cisco VXC 6215 includes the Virtualization Experience Media Engine (VXME) that allows media to terminate directly on the thin client. VXME enables the Cisco VXC 6215 to render both audio and high-definition video. The direct media termination approach allows the media stream to take an optimum path between endpoints and allows native network-based QoS to be applied to the media stream. Having VXME handle unified communications media locally at the client also avoids user-experience degradation when running unified communications applications in soft-phone mode inside the HVD. This degradation is the result of a hairpin effect, in which the unified communications media stream travels through the display protocol all the way back to the HVD in the data center and then out again to the remote endpoint, which results in excessive bandwidth and server resource use and an increase in latency.

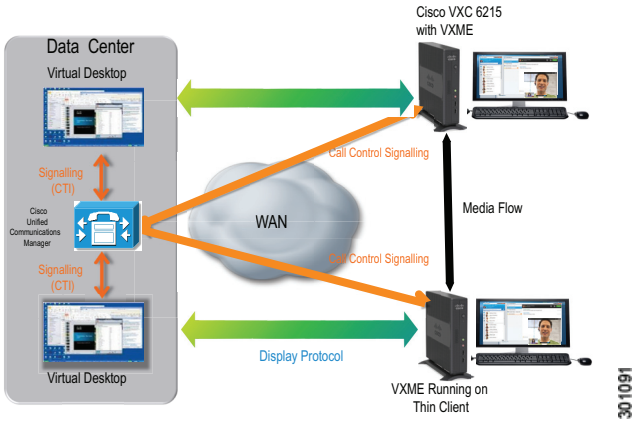
For any HVD client, the network providing connectivity back to the data center is the most important part of a virtual desktop solution. VXME is designed to leverage the network by intelligently separating unified communications media traffic and placing it outside the display protocol traffic (PCoIP) allowing for network-awareness of each traffic type (Figure 30). VXME for Cisco VXC 6215 has the capability to tag both Unified Communications and display protocol traffic appropriately with the correct VLAN values.

VXME optimizes the media traffic flows further by allowing unified communications media traffic to travel directly between the UC enabled end-points, while the desktop virtualization traffic flows between the Cisco VXC 6215 and the datacenter hosting the HVD. Cisco VXC 6215 with VXME's ability to allow for different media flows to be separated along with network awareness of each media flow delivers a best in class user experience. Cisco Jabber running in the HVD, accessed by the Cisco VXC 6215 with VXME add-on, enables VXME to register with the Cisco Unified Communications Manager in place of Cisco Jabber on HVD itself.

This separation on the Cisco VXC 6215 results in better call-admission-control (CAC), emergency (E911) and codec selection for media traversing different types of links. This is in contrast to softphone-based solution where softphone runs in HVD and Cisco Unified Communications Manager has no awareness of the location of the endpoint.

Note VXME uses ports between 16384 and 32766 for unified communications media. The audio port is randomly selected from the bottom half of this range whereas the video port is selected randomly from the top half of this range.

Figure 30 Separation of Media using Cisco Jabber with VXME



For more information on Cisco VXME, please refer to the Rich Media, Collaboration and User Experience chapter in this document.

Endpoint Management

Managing endpoints involves the following processes:

- Software image distribution and upgrades
- Initialization and configuration
- Monitoring and troubleshooting
- Configuration to set capabilities, enable local applications, and allow peripherals

In addition, endpoint management presents challenges related to the variety of endpoints available, the scale of the deployment, and the geographical distribution of the managed endpoints. An effective management strategy must address all these processes and challenges.

Cisco VXC Manager

Cisco VXC Manager provides centralized, enterprise-scale manageability for Cisco VXC deployments. It is used to efficiently provision, monitor, and troubleshoot the Cisco VXC endpoints; perform asset management and inventory; remotely control and access the devices; update firmware and configurations using an enterprise policy-based approach; and generate reports. Cisco VXC Manager supports advanced device control, scripting, and imaging capabilities. It supports secure (HTTPS based) imaging, compression, and bandwidth throttling for efficiency; device health-status reporting, shadowing, remote control and scripting, and default configuration; device discovery, with detailed device hardware and software asset information; and a client cloning feature that allows the administrator to capture the image and configuration on a reference client and deploy it across the enterprise. It also supports a distributed scalable architecture that includes the use of master and remote software repositories for deploying images.

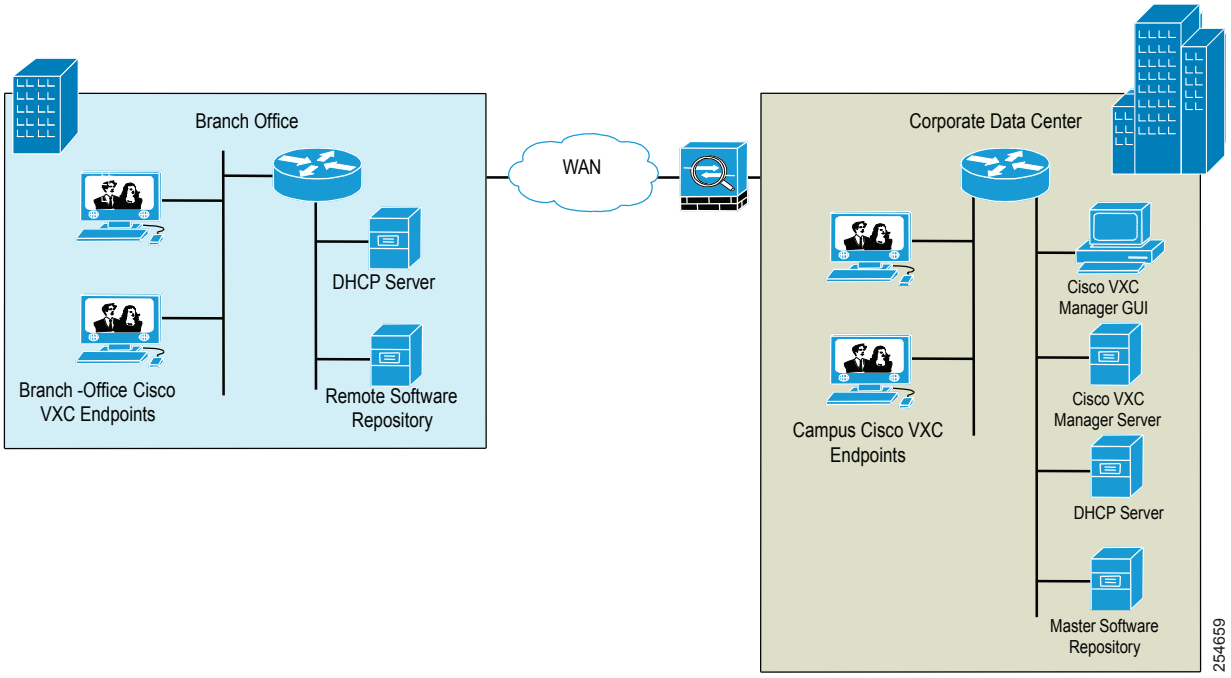
Cisco VXC Manager includes the following components: management console GUI (Microsoft Management Console-based snap-in) for administration; device management services for communicating with endpoints (includes HTTP, Dynamic Host Configuration Protocol [DHCP] proxy, and Trivial File Transfer Protocol [TFTP] support); a database for storing configuration and device information; and a master software repository for storing images, configuration files, and packages. Cisco VXC runs a web agent that enables communication with the Cisco VXC Manager server.

The Cisco VXC Manager console includes device, package, update, report, and configuration managers. The device manager displays the health status of all devices that check in with the Cisco VXC Manager server. The device status is green, yellow, or red, depending on the check-in status of the device. The device checks in to the server periodically to report the device status and check for updates (such as a configuration or firmware update). The default device check-in interval is 60 minutes for Cisco VXC clients. This value can be configured in the Cisco VXC Manager preferences. A device that is powered off will appear in red in the device listing.

Note that the device status does not include information about whether the device is idle or is actively being used in a remote-desktop session. The device manager also provides detailed information about the device hardware and software profiles (including the OS version), software installed, network settings, and packages installed. Devices can be logically grouped according to OS type, geographical location, building, floor, department, or subnet. This grouping makes it easier to view the status, apply policies, and deploy packages to a particular group of devices.

The device discovery feature allows desktop virtualization endpoints to be manually added to the device manager or discovered dynamically by the device manager by providing an IP address range or subnet within which to perform the discovery. Cisco VXC Manager can remotely control a desktop virtualization endpoint, including reboot, shutdown, and wakeup (using Wake on LAN [WoL]) operations. Shadowing allows the administrator to remotely access the desktop virtualization endpoint console to verify settings and set up a remote-desktop session. Shadowing can be useful for troubleshooting endpoint problems in real time. Management features available for endpoints depend on the endpoint OS and device type.

Figure 31 VXC Manager Deployment Model



The Cisco VXC Manager includes the following components: management console GUI (MMC based snap-in) for administration, device management services for communicating with endpoints (includes HTTP, DHCP proxy and FTP support), a database for storing configuration and device information, and a master software repository for storing images, configuration files, and packages. The Cisco VXC clients run a web agent that enables communication with the Cisco VXC Manager server.

The Cisco VXC Manager management console includes a device, package, update, report, and configuration manager. The device manager displays the health status of all devices that check-in with the Cisco VXC Manager server. The device status indicates green, yellow, or red depending on the check-in status of the device. The device checks-in to the server periodically to indicate device status and check for updates including configuration or firmware. The default device check-in interval is 5 minutes for PCoIP based Cisco VXC clients. Other devices check-in with the Cisco VXC Manager every 60 minutes. The check-in interval can be configured in Cisco VXC Manager. A device that is powered off will appear red in the device listing. Note that the device status does not include information about whether the device is idle or is actively being used in a remote desktop session. The Device manager also includes detailed information on the device hardware and software profile (including OS version), software installed, network settings, and packages installed. Devices can be logically grouped according to OS type, geographical location, building, floor, department or subnet. This makes it easier to view the status, apply policies, and deploy packages to a particular group of devices.

The device discovery feature allows desktop virtualization endpoints to be manually added to the device manager or discovered dynamically by the device manager by providing an IP address range or subnet within which to perform the discovery. The Cisco VXC Manager can remotely control a desktop virtualization endpoint, including reboot, shutdown, and wakeup (using Wake on LAN) operations. Shadowing allows the administrator to remotely access the desktop virtualization endpoint console to verify settings and set up a remote desktop session. Shadowing can be useful for troubleshooting endpoint issues in real time. Management features available for endpoints depend on the endpoint OS or device type.

The Package Manager is used to deploy packages to Cisco VXC endpoints. A package is a collection of files (scripts, configuration files, or firmware images) that instruct the device to perform a set of actions (modify device settings, upload a configuration, download a firmware image, shutdown, reset,). The Cisco VXC Manager script builder utility can be used for creating scripts to be included in packages.

Note When using this utility, the administrator needs to select the OS for the target device (for instance ThreadX for PCoIP VXC clients).

It is recommended to validate packages on a test device before deploying it across the enterprise. Note that the web agent update in package manager does not apply to Cisco VXC clients. The web agent running on a Cisco VXC client can be updated by changing the firmware of the device.

The update manager feature of the Cisco VXC Manager can be used to schedule the deployment of packages to devices at a prescribed time (for instance during off-peak hours). The scheduled time is automatically adjusted to the time zone of the client device. For example, a package can be deployed to a group of devices to update the firmware at a scheduled time or to reboot a group of devices in order to clear user sessions at 6pm every day. The report manager can be used to generate a log report to troubleshoot administrative changes to the Cisco VXC Manager; a device listing report which include a complete list of devices being managed; a package distribution report which includes package deployment logs and status (this report indicates successful package download to devices).

Use the configuration manager feature of the Cisco VXC Manager to customize settings on the Cisco VXC Manager server, create device groups, setup subnets and IP ranges, and provision remote software repositories. The Cisco VXC Manager settings include device check-in interval, enabling device security, changing logging levels, timeout preferences, and enabling default device configuration (DDC). Timeout preferences can be set globally for all endpoints or a subset of endpoints based on subnet. For example, the timeout can be adjusted for endpoints located across high latency WAN links. It is recommended to limit the number of simultaneous image updates across low bandwidth links (so that they are not slowed down to the point of failure due to timeout). This setting can be provisioned in configuration manager. In addition, the Cisco VXC Manager supports role based access and provisioning of administrators with associated permissions. For example, read-only permissions can be assigned for a specific group of administrators. Integration with Active Directory (AD) allows addition of multiple administrative users who can be delegated to manage separate pools of endpoints organized by logical views, endpoint properties, OS and device type and subnet.

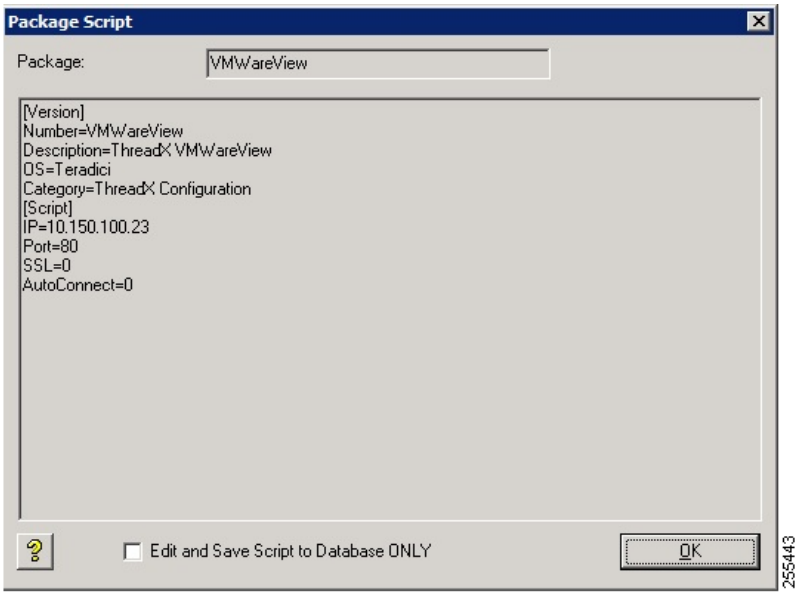
The repository creation and administration feature allows the administrator to easily build and administer a repository of software, images, and configuration updates for distribution. The Cisco VXC Manager server includes a master software repository, which acts as the image and package server. The master repository can be synchronized with a remote software repository that is located in a branch office. For remote locations such as a branch office, it is recommended to use remote file repositories that are local to the endpoints to conserve WAN bandwidth. Since the file transfers are completed using TCP based protocols (FTP and HTTP), any file transferred across the WAN would benefit from the acceleration, caching, and compression features of Cisco WAAS.

The Cisco VXC Manager Default Device Configuration (DDC) is a policy management tool that allows administrators to create rules for automatic deployment of OS images, scripts, packages, and other settings to thin-client end-points. DDC allows an administrator to configure default software and device settings for a group of devices to simplify and fully automate the management of their endpoints. For example, an administrator for a global organization can use DDC to map a subnet to a specific country. When an endpoint is connected to the network for the first time, DDC uses the subnet information to provision the DV endpoint with the correct language settings and configuration. The DDC feature ensures device configuration and image version conformance for all devices in a particular group (like building or subnet) running a particular OS. It is useful for automatically implementing the device configuration change based on a change in the subnet, physical location or device group of a device. For

example, a Cisco VXC client can be easily migrated from one VMware View environment to another and be re-configured automatically simply be changing the VLAN assignment of the device or physical port on a switch. Note that a DDC is applied to a device during the device bootup phase when it checks-in to the Cisco VXC Manager server.

A package for PCoIP Cisco VXC client includes a script to change configuration settings and firmware and is applied to the device during the device check-in to Cisco VXC Manager or by an administrator using the package manager.

Figure 32 VMware View Package Script



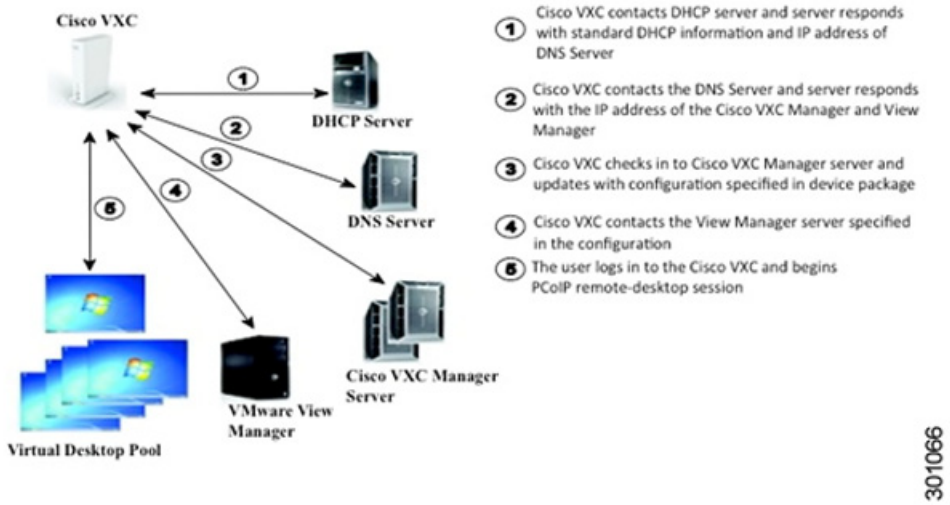
Use the Cisco VXC Manager device manager dashboard to modify the following configuration on a PCoIP Cisco VXC client: VMware View client, label, video, and time zone settings.

Guidelines for using Cisco VXC Manager

When deploying a large number of desktop virtualization endpoints, a centralized configuration approach is recommended. This approach allows an endpoint to automatically detect the Cisco VXC manager and desktop controller locations and update itself with the correct image and settings. Endpoints should be provisioned with network settings, locations of Cisco VXC manager server, and the location of a desktop controller during boot-up using DHCP, DNS, and FTP services. Refer to the VXC client and VXC Manager documentation for more information about the specific DHCP options to use here.

In addition, file server locations can provide configuration files instructing the endpoint to download updated images and configurations. The desktop virtualization endpoints can be reset or rebooted by end users. Alternatively, the administrator can centrally manage the power on Cisco VXC clients (using PoE switches with an EnergyWise compatible management application) or the administrator can use the Cisco VXC Manager to centrally initiate this task and push configurations and image updates to endpoints. Remember that these settings (filer server, the Cisco VXC Manager server, and desktop controller locations) can also be provisioned manually by accessing the endpoint through the console or web interface. Consult the Cisco VXC documentation for specific instructions for setting up a master file server repository, including the directory structure for storing configurations and images and the format of the configuration files for each desktop virtualization endpoint type being deployed.

Figure 33 VXC Client Boot Process



The PCoIP Cisco VXC endpoint can be managed using a web interface to remotely access configuration settings, using on screen display available on the device console, using the device manager dashboard, or by using package manager feature of the Cisco VXC Manager. The Cisco VXC endpoints can learn the Cisco VXC Manager server location using DHCP options or using standard service (SRV) and DNS records that identify the location of the server. Use DHCP options (12 and 15) to specify the Cisco VXC Manager server location (IP address, port, and secure port). Alternatively, Cisco VXC endpoints will query the DNS server for the Cisco VXC Manager server location (The hostname '_pcoip-tool' is used by Cisco VXC PCoIP clients). Note that the Cisco VXC PCoIP clients will also query the DNS server for the VMware View Manager location (The hostname is `_pcoip-broker'). The use of DNS can be useful in a deployment scenario where the DHCP service cannot be modified (for example, on a branch router located at a customer site). The use of manual or auto discovery of the endpoint in the Cisco VXC Manager device manager is another method of informing the device of the Cisco VXC Manager location. Use the subnet manager to define subnets and IP ranges on which to auto discover devices. This can be useful when the devices cannot be remotely rebooted in order to discover the Cisco VXC Manager server.

FTP should be enabled on the master software repository during the Cisco VXC Manager installation to support image deployment to Cisco VXC PCoIP endpoints.

The recommended method for creating and modifying package configuration files for Cisco VXC PCoIP devices is to use a text editor.

Ensure that all communication ports used by the Cisco VXC Manager are open on the network to allow it to communicate with endpoints and management console: 80,280, and 443 (HTTP); 21 (FTP); 1433 (database), and 5900 (VNC) Please consult Cisco VXC Manager documentation for complete list of ports used by the server.

The Cisco VXC Manager can be installed on a single server or multiple servers for large deployments (by distributed services across multiple servers using the customized installation procedure). Note that multiple management consoles can be installed to facilitate remote access from multiple administrators.

Use the power management features to power on and power off endpoints at scheduled times to conserve energy. Note that only devices that support WoL can be powered on by the Cisco VXC Manager.

Ensure that remote software repositories are synchronized with the master software repository. Use Microsoft synchronization services to synchronize the software repositories. When using remote software repositories across a WAN, it is recommended to deploy a remote repository in branches with slower links or with relatively large number of users.

The recommended troubleshooting tools include Wireshark to monitor device and server communications, Microsoft DebugView to log events and errors on the server including database communications, and Cisco VXC Manager server log reports.

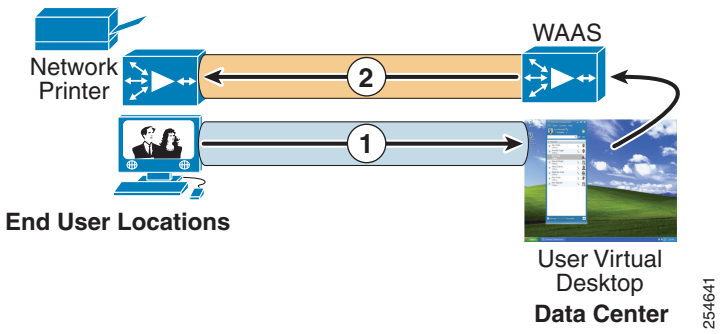
Custom reports can be generated using any tools that are Open Database Connectivity (ODBC) compliant (for example, Crystal reports, Microsoft Access, and Microsoft SQL client tool) and that can access the Cisco VXC Manager database.

Desktop Virtualization Endpoint Printing using Network Printer

When a user sends a document to a network printer co-located with the user's endpoint, the print flow actually originates from within the data center. The print flow to the network printer actually travels outside the desktop display protocol, and can be optimized with Cisco WAAS. Cisco WAAS can recognize the printing protocol, compress it to reduce WAN bandwidth consumption, and send the resulting print file to the remote branch. If the remote printer is busy, Cisco WAAS offers queuing functionality. Details of this print optimization and queuing can be found in configuration guide for Cisco Wide-Area Application Services:

http://www.cisco.com/en/US/docs/app_ntwk_services/waas/waas/v421/configuration/guide/cnfgbook.pdf

Figure 34 Network Printing Data Flow



1. The request for document to print at local network printer travels within the desktop display protocol.
2. From DV Desktop in data center, the print job is sent to the print server, also within the data center. It is then streamed to the printer.

Note This protocol stream can be recognized and optimized by Cisco WAAS.

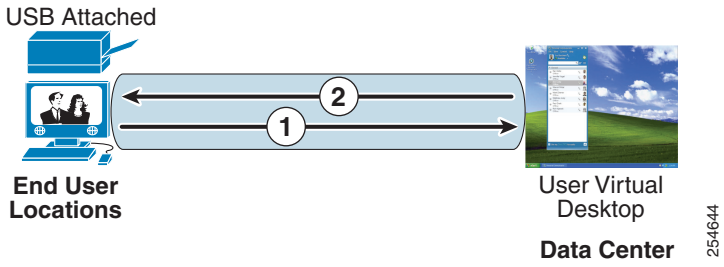
Desktop Virtualization Endpoint Access to USB-attached Peripherals (Storage or Printing)

DV endpoints often feature USB ports allowing connections to storage devices (such as, hard disks, "thumb" drives) and printers. Although physical access to the USB port may allow a user to connect such peripherals, the connection broker controls the logical enabling of the port.

USB port access is a policy that follows the user, not the device. For example, one user may be allowed to use storage peripherals on a DV endpoint, log out, and then when another user logs into the same DV endpoint, storage access is disabled. When USB peripheral access is enabled for a given user, the flow of data to the USB peripheral originates in the user's Hosted Virtual Desktop, in the Data Center; all file storage or print operations to a USB-attached peripheral thus result in a data flow across the network to the DV endpoint, encapsulated within the display protocol.

Details on enabling/disabling the USB access can be found in the Securing Cisco Virtual Workspace chapter of this document.

Figure 35 USB Printing Data Flow



Rich Media, Collaboration and User Experience

A distinguishing feature of the Cisco Virtual Workspace (VXI) Smart Solution offering is its built-in support for voice and video telephony and network supported multimedia. This support is derived from the following architectural emphasis:

- Media streams should not be mixed with display protocol data
- Quality of service (QoS) for interactive voice and video must be provided and enforced in the network
- The best user experience results from point-to-point media flow and not hairpinning through the data center
- Network bandwidth resources should be used optimally

What is New in Release 2.7

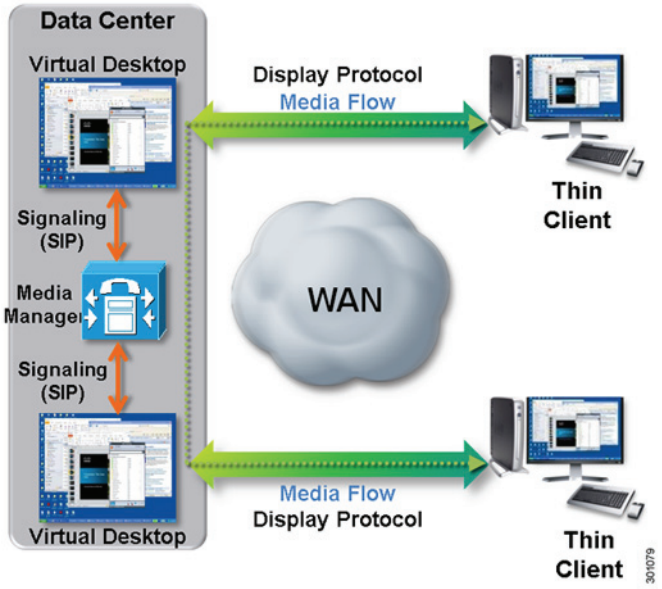
This chapter describes the collaboration features integrated into the Cisco Virtual Workspace (VXI) Smart Solution.

The Cisco Virtualization Experience Media Engine, as an extension of Cisco Jabber for virtualized environments, is now available as an add-on to the VXC 6215 base firmware version 9.0.

Rich Media and Collaboration in Traditional Virtual Desktop Infrastructure Environments

A traditional hosted virtual desktop (HVD) connection relies on an IP-based display protocol connection to carry both graphical user input (mouse, keyboard, etc.) and multimedia information (such as voice and video telephony) between the endpoint and the HVD. This approach places all information types (voice, video, display updates, USB-based information, etc.) in the same IP connection between the endpoint and the HVD. Because the display protocol is opaque to network services, QoS, call admission control (CAC), and codec negotiation cannot be applied to the individual media streams within the display protocol, possibly resulting in a suboptimal user experience.

Figure 36 IP Telephony in traditional Virtual Desktops

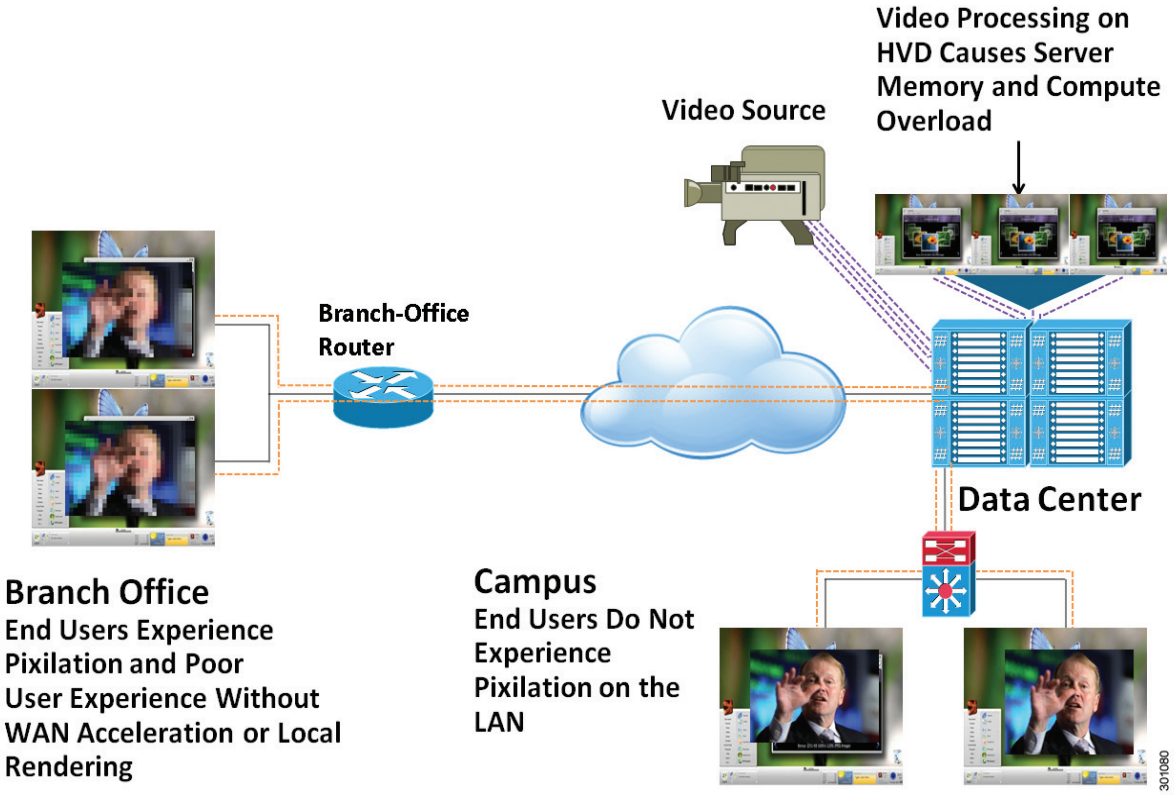


In Figure 36:

- An audio/video call is made using a softphone running in the virtual desktop in the datacenter to another softphone running on a different virtual desktop
- Signaling and media for the call run between the media manager in the datacenter to the virtual desktops also in the datacenter
- Voice and video are rendered using the display protocol on the user's thin clients located across the WAN from the datacenter

Now consider the case where a video is being streamed to virtual desktops. This situation is illustrated in the Figure 37.

Figure 37 Streamed Media in traditional Virtual Desktops



In Figure 37, video from the source is sent from the data center to virtual desktops, also in the data center. Video displayed on these desktops is rendered on the users' endpoints using the display protocol. User endpoints that are located in the campus and delivered from the data center across the LAN are able to view the video with no pixilation. However, video delivered to users across the WAN suffers from pixilation, resulting in a poor-quality user experience.

The following problems may occur with the above methods to deliver IP-telephony and other multimedia to virtual desktop users especially those located across a WAN.

Sub-optimal Routing

In a video call between traditional desktops, media is routed optimally from one endpoint to the other. When using a hosted virtual desktop, all application and OS processing is performed in the datacenter. The end clients collect audio data from the audio or a USB port, and video data from a USB port and redirect it to the application running in the datacenter. This effectively creates a hub-and-spoke network. This form of communication is appropriate for the desktop itself as it is actually located in the data center and needs to be displayed on the endpoint. It is not optimal for media.

Much work went into the design of the SIP protocol to ensure that even though call signaling may use a central point, media flows optimally from one endpoint to another. The issue of data passing through a central point instead of routing optimally through the network is referred to as hairpinning.

There are some techniques for optimizing the flow of media within the VDI session itself, but these do not fix this fundamental issue.

Encapsulation in VDI Protocols

With a traditional desktop, applications run on the computer and establish independent network connections, often using well-defined ports and protocols. This enables the network to identify and act upon those network flows, providing QoS, caching, WAN optimization, and other services.

In contrast, hosted desktops and applications run in the data center. The user then accesses that desktop remotely via a single virtual display protocol. This protocol encapsulates:

- The display of the desktop itself
- Keyboard and Mouse information
- Signaling, session authentication, and control
- Remote USB data from local peripherals

The intermixing of different types of traffic in a single stream makes it much harder to manage and optimize the media traffic appropriately in the network.

Media Re-processing

In a video call between two traditional desktops, media is encoded once by the transmitting endpoint, and decoded once by the receiver.

In a video call between hosted virtual desktops, the local camera video is encoded as USB data by the VDI endpoint and sent to the hosted desktop. On the desktop, the video application is running and sees the camera as just another USB device. It reads the USB data and encodes it as RTP. The data is then sent to the other video application, which may be within the same data center. There the process is reversed as the data is decoded, re-encoded as USB, and sent to the other VDI endpoint.

This decoding and encoding significantly increases the CPU load of the hosted desktop. This impacts both the user experience and the scalability of the data center.

Increased Latency

Hairpinning media through the data center, rather than directly from one endpoint to another, increases the latency of the connection – the time it takes for audio or video from one endpoint to be played or displayed on the other.

Consider a scenario where two end clients are situated within the same building in San Jose, CA and the datacenter is situated in a different geographically location, Boston, MA. The hub-and-spoke architecture will require all real-time audio/video data to travel from End Client1 (San Jose) to datacenter (Boston) and then from the datacenter (Boston) to End Client2 (San Jose).

In addition, re-processing the media in the hosted desktop creates not only CPU load, but also additional latency. These two combine to have a significant impact on user experience.

For an acceptable call, Cisco and others recommend less than 150 milliseconds latency, with an absolute maximum of 400. Beyond this, a conversation becomes difficult as people talk over each other and miss the other's visual cues.

Optimizing the flow of audio and video within the virtual session offers minimal improvement as the fundamental issues of routing and media processing remain.

Increased Bandwidth Consumption

The network bandwidth consumed by VDI is already a challenge, especially over more constrained links like WAN connections. Adding media to VDI seriously increases the impact.

The first increase is due to the hub-and-spoke routing issue. This increases not only latency, but also bandwidth, as the media flows to the data center and back instead of flowing directly between two endpoints.

As an example, consider a phone call between two employees in a branch office. Instead of flowing entirely within the local network, the media will be sent over the WAN link to the data center, and then back. This doubles the amount of data, and moves it to a constrained link.

The encapsulation of media data within the VDI protocol makes the problem even worse by sending it to the data center as USB data rather than compressed RTP.

The combined effect of this can increase the data transfer rate in a virtual desktop environment to almost 100 times higher than that in traditional desktop environment (see Figure 39).

This is a problem with individual users in normal use, but the effect is multiplied when many people make calls at the same time. This could be due to a company event or conference, a crisis of some kind, or some other unique event that generates a lot of calls. The problem is magnified just when optimization is needed the most.

Loss of Network Optimization

When rich media applications are deployed, the network is designed to provide Quality of Service (QoS), traffic engineering, and routing of the application flows. The encapsulation of media within a VDI protocol makes it difficult or impossible for the network to identify the flows and act upon the media. The traffic from all applications is combined into one or more VDI streams. For example: the data traffic from a print application is encapsulated along with the rich media traffic from a video application. When this is done, the network cannot sufficiently recognize and differentiate to provide the optimum Quality of Service needed for rich media flows.

Such lack of differentiation also reduces the ability of the network administrator to monitor and troubleshoot the network. For example: while an administrator was able to estimate, measure, and thereafter adapt their network for rich media applications, they can no longer do so because they have no visibility in what constitutes the rich media traffic in a VDI session.

Unable to provide Call Admission Control

While provisioning the right amount of bandwidth in the network for rich media sessions is critical, it is also necessary to apply call admission control so that too many sessions (audio or video) do not overrun the available bandwidth. Call admission control can be applied by either on-path mechanisms such as Resource Reservation Protocol (RSVP) or off-path mechanisms such as Locations-CAC in Unified Communication Manager.

When the rich-media traffic is encapsulated within the VDI protocol, call admission control cannot be effectively applied because all traffic appears as part of the VDI stream.

Unable to use Voice VLAN

Voice VLANs are recommended for a variety of reasons such as preventing broadcast storms for voice devices, a trusted means of providing QoS for rich media traffic, allowing multiple devices to be daisy-chained and connected to the network on a single port and so on. This benefit is lost if the media is encapsulated within the display data stream.

Loss of 911 emergency capabilities

Enhanced 911 Services enable 911 operators to:

- Locate the user based on the calling number
- Callback the 911 caller if a disconnect occurs

If the entire softphone is running in the datacenter then 911 services will be unavailable or will incorrectly identify the location of the caller as being in the datacenter.

Increased CPU load in the datacenter

Processing the media for the call in the data center increases the CPU load on the data center servers in several ways:

- Decoding USB audio and video from the endpoint
- Encoding that media as RTP for the call
- Decoding incoming media from RTP
- Encoding incoming media as USB to be sent to the endpoint

Media encoding and decoding is one of the most CPU-intensive tasks. For a call between traditional endpoints there is only one encode and one decode. The hosted desktop, however, has to do more significantly increasing its CPU utilization. Furthermore, while there is generally excess CPU capacity in a traditional desktop, hosted desktops are not generally designed with latent capacity. A large CPU processing requirement from one user can not only impact the performance for that user, but also for others who share the same pool of resources.

The problem is most acute in times of high call load such as a crisis situation or important event, just when optimization is most needed.

Loss of network management tools

Many tools exist to manage networks and VoIP deployments. In many cases, these are already deployed to manage an existing VoIP deployment at a customer site.

Network management and diagnostic tools rely on being able to identify and mark packets associated with a stream according to the network treatment appropriate for the stream. The flow of multimedia streams in a VDI session not only 'hairpin' through the data center but are also mixed either directly into the VDI desktop image generated in the data center or are tunneled through the VDI protocol connection as a 'Virtual Stream'. In both cases, it is not possible for network management tools to differentiate between VDI desktop image packets, multimedia packets or even USB redirection packets. Typically, in VDI all of these streams are intertwined and delivered to the endpoint via a single IP connection.

Media within VDI sessions renders most network management and monitoring tools ineffective, and decreases the value of the customer's investment in tools, training, and experience, making the deployment harder to manage and troubleshoot.

No Survivable Remote Site Telephony (SRST)

Cisco Unified SRST provides telephony backup services to help ensure that the branch office has continuous telephony service over the network infrastructure deployed in the branch location. Call-processing redundancy in the branch office is particularly critical in an emergency (which may be the actual cause of the WAN outage).

Cisco Unified SRST functions in the branch-office router to automatically detect a failure in the network and initiate a process to auto-configure the branch router, providing call-processing backup redundancy for the IP phones in that office and helping ensure that the telephony capabilities stay operational. Upon restoration of WAN connectivity, the system automatically shifts call processing back to the primary Cisco Unified Communications Manager cluster.

If the video/phone application is running in the data center then SRST in the branch no longer helps. If the WAN link goes out then employees in the branch lose not only their desktops, but also their rich media capabilities.

General user-experience and management issues

The issues outlined above add up to a seriously flawed user experience, especially at times of peak load when optimization is most needed.

These issues also impair the ability of the administrator to manage their network and communications system. Existing tools, techniques, and experience are no longer applicable, and new tools are not available.

There is a need for an architecture that provides an equivalent user experience to a traditional desktop, while preserving the IT investment in their existing network.

Some of the issues listed above can be observed by measuring the CPU and Network usage in the virtual desktop when a call is made in the traditional manner and the media is routed through the datacenter and carried by the display protocol. These measurements, taken in a test setup, are shown below. On an idle virtual desktop, CPU and Network utilization averaged 6% and 20 kbps respectively. These values increase multiple folds when on a video call. As shown in the Figure 38 and Figure 39 below, CPU utilization increased to 51% while Network utilization increased to 44 Mbps.

Figure 38 Resource Monitor of an idle virtual desktop: CPU 6%, Network 20 kbps

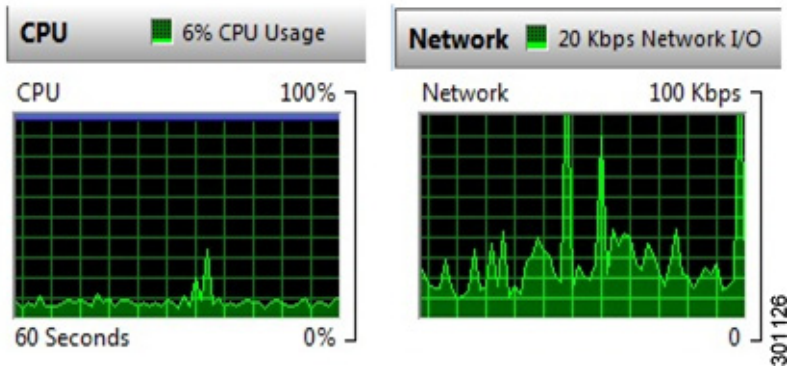
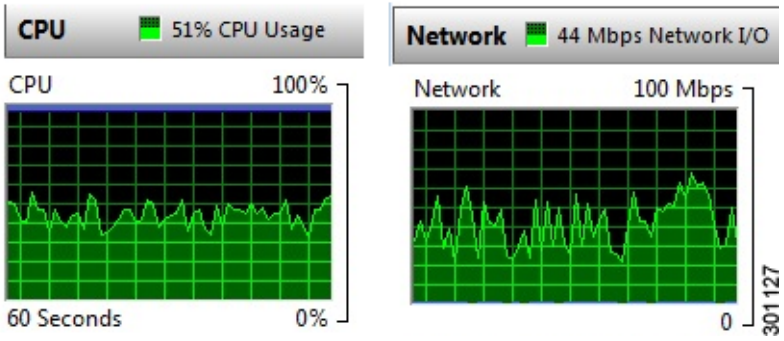


Figure 39 Resource Monitor of a virtual desktop during a point-to-point call: CPU 51%, Network 44 Mbps



Rich Media and Collaboration in Cisco Virtual Workspace (VXI) Smart Solution

User Experience is one of the most important architectural guiding principles for Cisco Virtual Workspace (VXI) Smart Solution. The Cisco Virtual Workspace architecture aims to:

- Avoid hairpinning, that is, avoid media traversing the data center as much as possible
- Separate the media streams from the display protocol so that proper QoS policies may be applied

This section provides a discussion of how these techniques are used to realize a better user experience.

IP Telephony and Interactive Voice and Video Traffic Separation

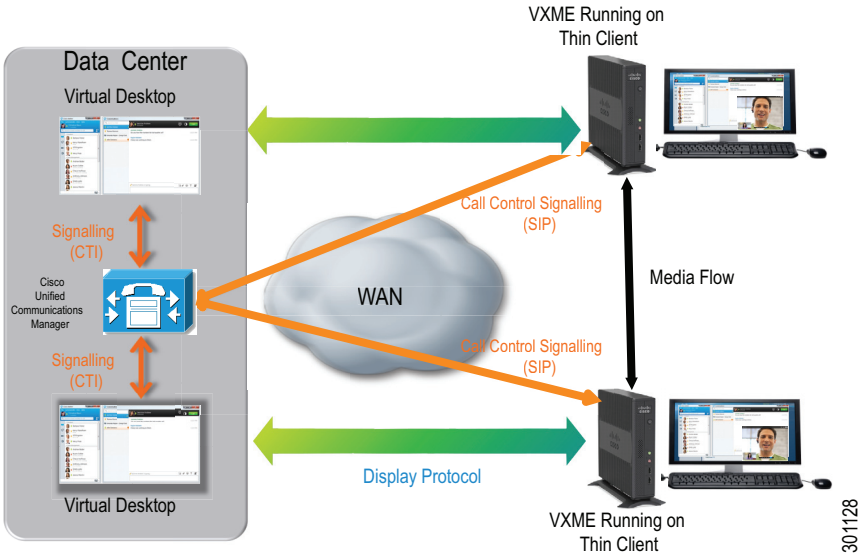
Cisco Virtual Workspace (VXI) Smart Solution adds intelligence in the end clients as well as the data center. Cisco Virtual Workspace (VXI) Smart Solution enables the end clients to separate rich-media data from the display protocol, to process rich-media data locally and to communicate directly with other end clients.

While the media processing moves to the endpoint, the user interface remains on the hosted desktop to preserve a seamless experience for the user. Interaction with other applications works as if everything was running in the virtual desktop.

This approach addresses the problems enumerated, in a way that maintains the value of existing investment in network and communications, while merging seamlessly with virtual desktop deployments.

Figure 40 demonstrates the flow of the separated traffic.

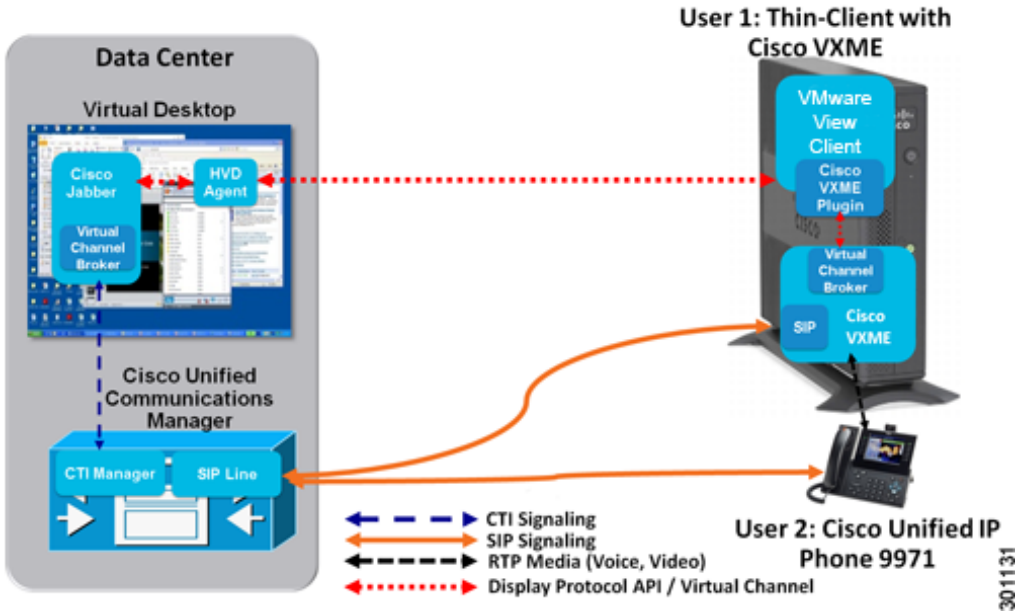
Figure 40 Flow of Separated Traffic



Cisco Virtual Workspace (VXI) Smart Solution implements this architecture by leveraging Cisco VXME, which enables Cisco Jabber for virtualized environments.

Cisco VXME, which acts as the local extension of Cisco Jabber running on the virtual desktop, handles voice and video traffic. The interaction between Cisco Jabber and Cisco VXME is shown in Figure 41.

Figure 41 Cisco Jabber in a Virtualized Environment with VXME



As shown in Figure 41, the control process is as follows:

1. With prior configuration, Cisco Jabber on the User 1 virtual desktop detects that it is running in a virtualized environment and is being accessed from a thin client running VXME
2. Cisco Jabber, through the VMware View agent running on the desktop, establishes a virtual channel through the VMware View client in the thin client to Cisco VXME residing within the thin client
3. Cisco VXME receives login information from Cisco Jabber through the virtual channel and registers with Cisco Unified Communications Manager as a Client Services Framework (CSF) device
4. When the user places a call through Cisco Jabber, Cisco Jabber uses CTI to instruct Cisco Unified Communications Manager, which directs Cisco VXME to place the call.
5. After the call is answered, both audio and video media flow directly between Cisco VXME and the remote phone (which may be another thin client running VXME, a gateway, or any other Cisco Unified Communications Manager endpoint).

Note that this architecture provides native extension mobility. It enables user independence from a specific thin client. A user, such as User 1, can log in to their desktop from thin client-1, and the desktop will automatically configure VXME within the thin client-1 to register with Cisco Unified Communications Manager on the user's behalf. If User 1 logs out from thin client-1 and User 2 logs in, then VXME in thin client-1 will automatically register as User 2. Similarly, the User 1 desktop will automatically configure VXME in thin client-2 if User 1 logs in to the desktop from thin client-2.

Note Cisco Unified Personal Communicator and Cisco Unified Communication Integration for Microsoft Lync cannot control VXME and hence must be run in deskphone control mode when running on a virtual desktop.

With the capabilities of Cisco Jabber and Cisco VXME, the rich-media challenges as outlined in this chapter can be solved.

Latency

Placing the rich-media processing directly on the endpoint removes the hairpinning and enables media to flow directly from endpoint to endpoint. It restores the ability of the network to optimally route traffic, minimizing latency.

Encoding and decoding using codecs that are optimized for real-time voice and video rather than generic USB data further reduce latency.

Finally, the total encoding/decoding processes required for a video call is reduced by a factor of 3, from 24 to 8. With rich-media in VDI, each HVD must decode and re-encode the data. With VXME system, each video and audio stream is encoded only once by the sender, and decoded once by the receiver.

Bandwidth

Bandwidth consumption is reduced by routing media directly from one endpoint to the other rather than hairpinning through the data center. For example, if one VXME endpoint is used to call another in the same branch office, the media will flow between the endpoints in the branch with no impact to the WAN link.

In addition, the media is encoded and encrypted optimally for voice rather than as part of VDI USB data.

Network Optimization

The VXME architecture makes the thin client traffic appear to the network just like traffic from a regular Cisco desk phone with an attached PC. This means that all network optimization and diagnostics work as designed over the same network that is used today. The VXME architecture can be deployed as a seamless addition to the existing network. The same capabilities that the network provides for phones work in a UC enabled VDI deployment.

Call Admission Control

Moving the session management and rich-media functions to the client enables existing bandwidth allocation tools and call admission control to work as they do with regular phones. Call admission control can be applied by either on-path mechanisms such as Resource Reservation protocol (RSVP) or off-path mechanisms such as Locations-CAC in Unified Communication Manager.

QoS, Voice VLAN, and Medianet Support

VXME supports voice VLAN segregation. It uses CDP to determine the voice and access VLANs, if configured, on the access switch port and places media in the voice VLAN and all other traffic in the access VLAN. VXME obtains configured DSCP values from the Cisco Unified Communications Manager after registration.

VXME is also Cisco Medianet compatible: it can send metadata detailing its media usage and port numbers that may be used by the access switch to recognize and apply proper QoS markings to the traffic.

Decreased CPU Load

In Cisco Virtual Workspace (VXI) Smart Solution, only the user-interface elements run in the data center and no media processing is done there, dramatically reducing CPU usage. All media processing is done on the client, eliminating that load from the virtual desktop used with that endpoint.

It is important to note that the CPU load is not moved from the data center to the endpoint, but rather completely eliminated. Encoding and decoding is now required only on the endpoint, enabling better performance along with decreased load.

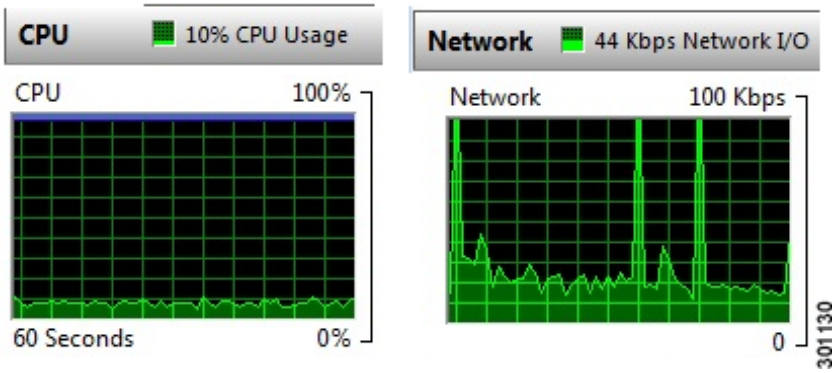
Network Management

Many tools exist to manage networks and Unified Communications deployments. In many cases, these are already deployed to manage an existing VoIP deployment at a customer site. The Cisco solution enables the existing network and UC management tools to be used with a VDI deployment.

Test Results

Cisco Virtual Workspace (VXI) Smart Solution dramatically reduces the required CPU and network resources. Figure 42below shows CPU and network utilization in the virtual desktop in midst of a video call.

Figure 42 Resource Monitor during a point-to-point call with Cisco Virtual Workspace (VXI) Smart Solution



Cisco VXME QoS Considerations

Cisco VXME, when paired with Cisco Jabber, supports a number of QoS enhancing features. These features are:

1. Separate Virtual LANs (Data and Voice VLAN) for VDI and IP telephony traffic
2. Configurable DSCP markings for VDI and IP telephony signaling, voice, and video
3. Compatibility with Cisco Medianet architecture allowing for automatic network recognition of media streams and application of QoS policies

The thin-client hosting VXME can be thought to contain two distinct identities – a PC identity consisting of the virtual desktop (VDI) traffic, and the telephony identity of VXME. Each of these identities has their own MAC address and obtain their own IP addresses.

The ports in the access switch, which the thin client connects to, may be configured with both the access (data) VLAN, and the voice (auxiliary) VLAN. On boot-up, the thin client running VXME obtains both VLAN IDs from the switch port by using a Discovery Protocol (CDP or LLDP-MED). The thin client then marks outgoing packets with the proper VLAN ID – access for VDI traffic and voice for telephony traffic.

Note The Cisco VXC 6215, if used as the thin client does not use the Link Layer Discovery Protocol – Media Endpoint Discovery (LLDP-MED) to obtain VLAN values from the access switch.

Having a separate VLAN for telephony traffic creates a logically separate sub-network for it on the same, shared infrastructure. This helps the network correctly identify the traffic and apply the right QoS policies to it. It also makes the thin client identical to the traditional model of the combination of an IP

phone with a PC connected to its back. By the virtue of this similarity, the same network design principles that are advised for Unified Communications are now applicable for the thin client deployment also.

VXME registers with the Cisco Unified Communications Manager using the CSF device type created for the user. As part of the registration process, VXME downloads the configuration file created by the Unified Communications Manager for the CSF device. This configuration file contains DSCP values for signaling, voice, and video packets. VXME marks signaling and media packets appropriately with these values, which helps prioritize this traffic through the network. The DSCP values used by VXME can be set as desired in the Communications Manager configuration.

Cisco Medianet is an end-to-end IP architecture designed to enhance user experience through a smarter transportation of media across the network. The architecture recognizes the value of multimedia communications in the business world. Medianet specifications place requirements on endpoints and the network consisting of switches and routers. Medianet compatible endpoints implement a set of protocols called the Medianet Service Interface (MSI). When such an endpoint is placed into a network designed for Medianet, it communicates with the access switch and makes it aware of its capabilities. The switch then automatically applies the pre-configured QoS policies. The Cisco VXME implements the Medianet Service Interface making it Medianet aware and capable.

For more information on Cisco Medianet architecture, please refer to http://www.cisco.com/go/medianet.

Cisco Medianet services can be used in situations where media traffic cannot be differentiated using VLANs. For example, a worker from home using a Cisco Virtual Office (CVO) router to connect to the corporate network, who is not able to use VLANs, may be better served with QoS via Medianet.

Also see the Quality of Service (QoS) section in the Virtualization Aware Network chapter for specific desktop virtualization traffic marking recommendations). QoS marking should also be consistently applied between the access point and the HVD to provide a high-quality user experience.

All traffic (desktop virtualization and voice and video) is sent through the same interface. The relative size of the various traffic flows can potentially create situations in which the network interface bandwidth is oversubscribed. Telephony calls generate a constant Real-Time Transport Protocol (RTP) media flow commensurate with the codec being used; for example, a G.722 audio call will consume approximately 80 kbps for the duration of the call, in each direction (this is a duplex call). Table 15 shows the nominal bandwidth consumption of the audio and video codecs supported by VXME.

Table 15 Codec / Bandwidth Consumption

Codec	Bandwidth Consumption
G.729 (a, b, and ab)	24kbps, symmetrical
G.711 and G.722	80kbps, symmetrical
iLBC	Up to 80 kbps
H.264	~384 kbps

Desktop virtualization bandwidth is more asymmetrical. The bandwidth consumption is also related to the amount of activity on the user's hosted virtual desktop. For example, an idle desktop generates only a small, quiescent amount of traffic, whereas a desktop running visually robust applications (for example, Adobe Flash video) consume more bandwidth. (Note that interactive video and audio traffic related to unified communications video conferencing is not included in the desktop virtualization bandwidth calculation because it travels outside the display protocol)

You can limit the User Datagram Protocol (UDP) port range for voice traffic by configuring media properties for the Cisco VXME on its device profile defined in the Cisco Unified Communications Manager. The default port range is 24576 to 32768. Please refer to the Campus section in the Virtualization Aware Network chapter of this guide for implementation details and special considerations.

Contact Center Applications in Cisco Virtual Workspace (VXI) Smart Solution

In many respects, contact center and desktop virtualization are complementary. Consider the following characteristics of a contact center:

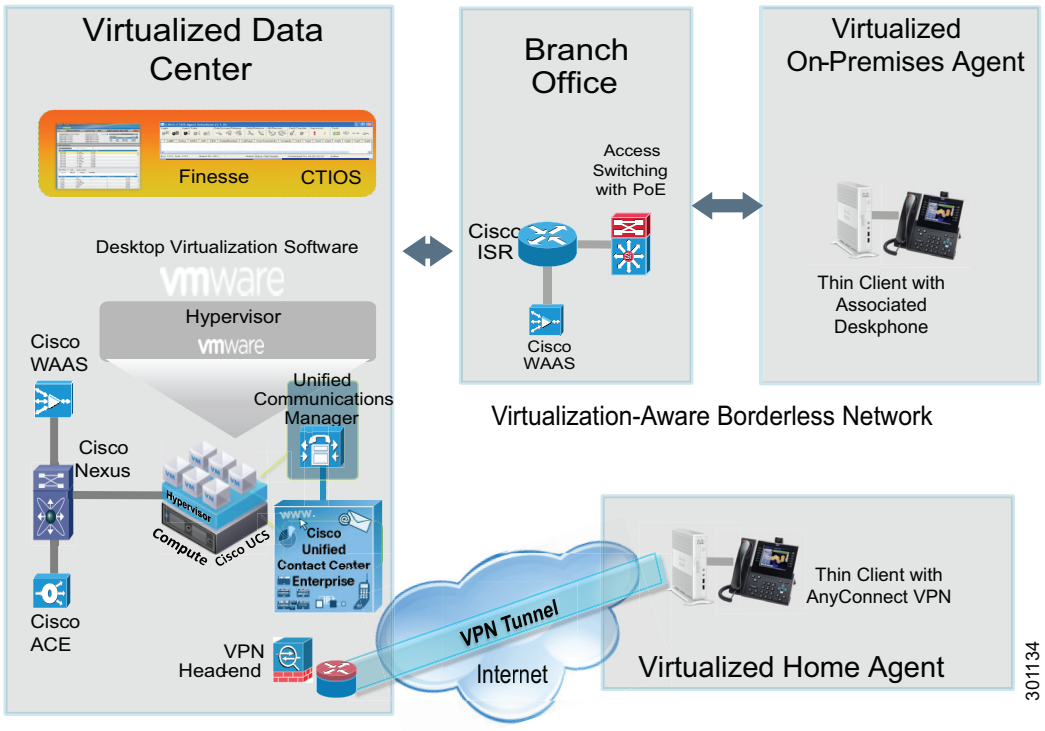
- Contact center agents usually are "task workers" who typically run a single or a small set of applications on their desktops
- Contact center operations are increasingly becoming distributed with agents working from home or outsourced locations around the world
- The contact center industry typically sees a lot of turnover, with agents moving relatively quickly between opportunities
- The contact center industry faces seasonal fluctuations in demand or may need to respond to increase in demand due to new initiatives
- Data security, confidentiality, and business resiliency are extremely important in the contact center
- Applications used by contact center agents change or need to be ungraded frequently

Cisco Virtual Workspace (VXI) Smart Solution provides a comprehensive solution for hosting agent desktops, securing data, and helping ensure business continuity with integrated unified communications capabilities.

Design Considerations

Figure 43 shows the integration of Cisco Unified Contact Center Enterprise (UCCE) with Cisco Virtual Workspace (VXI) Smart Solution.

Figure 43 Integration of Cisco Virtual Workspace (VXI) Smart Solution and UCCE



The following Cisco Unified CCE components are supported for a Cisco Virtual Workspace (VXI) Smart Solution deployment:

Agent Desktops: Desktops based on Computer Telephony Integration Object Server (CTIOS) Toolkit and Finesse are supported on VMware View virtualization technology.

Server Virtualization

All the servers that make up the complete Cisco Unified CCE solution can be hosted on Cisco UCS blade or rack servers. These include:

- Cisco Unified Communications Manager
- Cisco Unified CCE
- Cisco Unified Customer Voice Portal (CVP)
- Cisco Unified IP Interactive Voice Response (IVR)

Cisco Unified Intelligence Center

For more information about this support and design considerations of deploying a contact center, please refer to the Cisco Unified CCE Solution Reference Network Design (SRND) at http://www.cisco.com/en/US/docs/voice_ip_comm/cust_contact/contact_center/ipcc_enterprise/ipccenterprise8_5_1/design/guide/uccesrnd85.pdf.

Agent Desktops & Deskphones

The virtualized agent desktop may either use CTIOS-based or the Cisco Finesse™ desktop applications. Cisco Agent Desktop (CAD) is not supported in the virtualized environment.

The virtualization of the desktop does not affect integration with unified communications. The desktops use CTI to control the agent's associated desk phone. However, these desktops cannot currently use the Cisco VXME and a separate deskphone must be used.

Note If the Cisco Finesse application is being used as the agent desktop, Cisco IP Phones 8961, 9951, and 9971 cannot be used as associated deskphones.

On Premise Agents

Agents based on company premises either in the main campus (across a LAN from the data center) or in a branch office, can use a thin client to access virtual desktops in the data center.

Remote Agents

Remote agents working from their home offices can use a thin client to access their virtual desktops in the enterprise data center. The connection to the data center can be established using a Cisco Virtual Office router.

Cisco Unified Communications Applications

In a Cisco Virtual Workspace (VXI) Smart Solution system, Cisco Unified Communications is supported through the deployment of desktop applications that utilize Cisco's Client Services Framework and are thereby capable of controlling a Cisco Unified IP Phone. Cisco Jabber is such an application. Cisco Cisco Jabber can run within the hosted virtual desktop (HVD), and can be used to control the user's deskphone. With all of these applications, the Cisco Virtual Workspace (VXI) Smart Solution end-user has the capability to manage contact information (directory lookup), obtain real-time availability status of colleagues and co-workers (presence), use online chat (instant messaging) and voice and video IP telephony and conferencing.

Cisco Jabber can also be configured in either the deskphone control mode or softphone mode. Versions of Cisco Jabber prior to 9.1.4 should only be operated in desk-phone control mode on HVDs for reasons explained earlier in this chapter. Cisco Jabber, version 9.1.4 or later, with VXME firmware release 9.0 automatically detects that it is running on a virtual desktop, establishes a communication channel with VXME, and configures it to act as an media extension of itself. When used in this manner, Cisco Jabber may be used in softphone mode, while still maintaining the separation of media from the display protocol traffic.

Cisco Jabber is supported on many different platforms, providing a uniform user experience and capabilities. It works with both cloud-based and on-premises services to provide voice, video, presence, instant messaging, desktop sharing, and conferencing services. For a full description of Cisco Jabber features and supported platforms, see http://www.cisco.com/web/products/voice/jabber.html.

For more information, see:

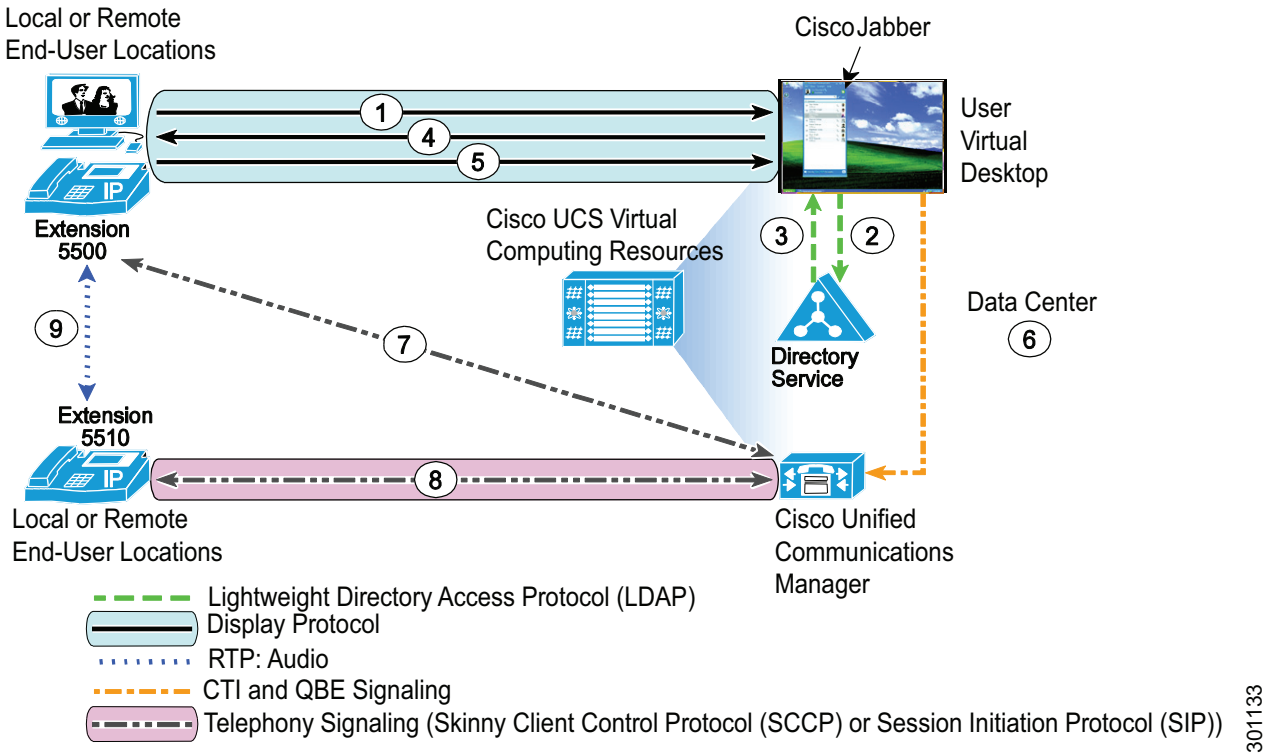
http://www.cisco.com/en/US/partner/docs/voice_ip_comm/cucm/admin/8_5_1/ccmsys/a08video.html

http://www.cisco.com/en/US/partner/docs/voice_ip_comm/cucm/docguide/8_5_1/dg851.html

For information about the software version of the supported unified communications applications, see the Cisco Virtual Workspace (VXI) Smart Solution As-Built Reference Guide at: http://www.cisco.com/en/US/docs/solutions/Enterprise/Data_Center/VXI/configuration/VXI_Config_Guide.pdf.

A typical Cisco Jabber in deskphone control mode call flow is shown in Figure 44.

Figure 44 Data Flow for Placing a Cisco Jabber Call from the HVD



In Figure 44, the desktop virtualization endpoint provides the user with a visual representation of their individual virtual desktop environment, and supports user input via keyboard, mouse and/or touch screen input device. The user's desktop is configured with an instance of Cisco Jabber, offering instant messaging, presence, and directory lookup and telephony control over the desk phone. The desk phone or unified communications enabled Cisco VXC client offers the voice and video media handling.

Note that all the user's actions are relayed within the display protocol to the virtual desktop in the data center. Likewise, all visual updates to the Cisco Jabber screen, like any other screen update, are delivered to the user within the display protocol from the virtual desktop to the HVD endpoint's screen.

User input is processed by Cisco Jabber, which runs on the virtual desktop in the data center. In turn, Cisco Jabber communicates with the appropriate collaboration servers, which are also within the data center and possibly within the same Cisco UCS domain. Notice that because the desktop is now running on a platform within the data center, all communications controlling the call remain within the data center. The complete data flow for this call process is listed here:

1. The desktop virtualization user types the name John Doe in the Cisco Jabber window. Through the display protocol, the information is relayed to the user's virtual desktop instance. Cisco Jabber receives the input as though the user had a keyboard locally connected to the guest OS.
2. Cisco Jabber uses Lightweight Directory Access Protocol (LDAP) to query the LDAP-compatible directory service for John Doe's number.
3. The contact information query results are returned to Cisco Jabber.
4. After processing by Cisco Jabber, this information is relayed back to the user's display through the remote display protocol.
5. If the desktop virtualization user initiates a call to a contact through the Cisco Jabber GUI on the user's desktop, a similar display protocol update is transmitted to the user's desktop virtualization instance in the data center.

6. CTI control data is sent from Cisco Jabber to Cisco Unified Communications Manager, requesting a call be placed from phone extension 5500 to 5510 (number 6 in Figure 44).
7. Cisco Unified Communications Manager performs dial-plan processing for the call request, resolves the called number to a destination phone, and instructs the desk phone to place a call to the call recipient. This operation is performed through a call-control exchange with both phones (numbers 7 and 8 in Figure 44).
8. The telephony call's media is established between the two IP telephones directly (number 9 in Figure 44), without going through the data center's virtual desktop. This aspect is critical to maintaining the quality of the user experience: the network's QoS policies protect the telephony media flow against dropped traffic, jitter, and latency impairments that may otherwise affect the flow if it were encapsulated within the display protocol. Note also that the IP telephony endpoints can connect their media streams directly across the most direct available network path.

Note The call is subject to CAC verification by Cisco Unified Communications Manager before the call is placed. If the network does not have enough bandwidth to allow the call to go through, the call may be processed through Cisco Unified Communications Manager's automated alternate routing (AAR) feature, be allowed to proceed with no QoS guarantees, or be blocked, depending on the Cisco Unified Communications Manager policy configuration.

Note The type of call placed (audio or video) and the codec used to place the call depend on Cisco Unified Communications Manager policy configuration and endpoint capabilities. If both the calling and called endpoint support video, the call may be placed as a video call, depending on Cisco Unified Communications Manager policy and CAC verification. If the endpoint capabilities, Cisco Unified Communications Manager policy configuration, or network condition allows only the placing of an audio call, the codec selection will be based on the Cisco Unified Communications Manager configuration.

Note Please consult the Cisco Virtual Workspace (VXI) Smart Solution Release Notes and Scaling and High Availability chapter of this guide for more information about the validation of the Cisco Jabber Cisco Virtual Workspace (VXI) Smart Solution. You should consult the product documentation for the unified communications applications to verify support in a VDI environment. You also should fully validate these applications within your environment before deployment.

Unified Communications Endpoint Single Sign-on

When a user logs into an HVD, a single sign-on (SSO) process can be used in which the user needs to enter his or her credentials only one time. The user enters the required credentials on the thin client GUI sign-on screen; this entry logs on the user to the appropriate HVD connection broker and automatically logs on and launches the connection to the HVD. After the user is connected to the HVD, Cisco Jabber can launch automatically. The unified communications user credentials are saved after the initial sign-on is performed manually, enabling automatic sign-on to the unified communications desk-phone control application. SSO for the HVD is supported in two ways: it can be configured manually on each thin client or by using a thin client management tool. When using Cisco VXC Manager and supporting services, .ini files are downloaded by Cisco VXC 6215 during the boot-up and login process containing global, device-specific, and user-specific configuration settings, including sign-in credentials and the HVD connection information For more information about Cisco VXC Manager, see Cisco Virtual Workspace Clients chapter.

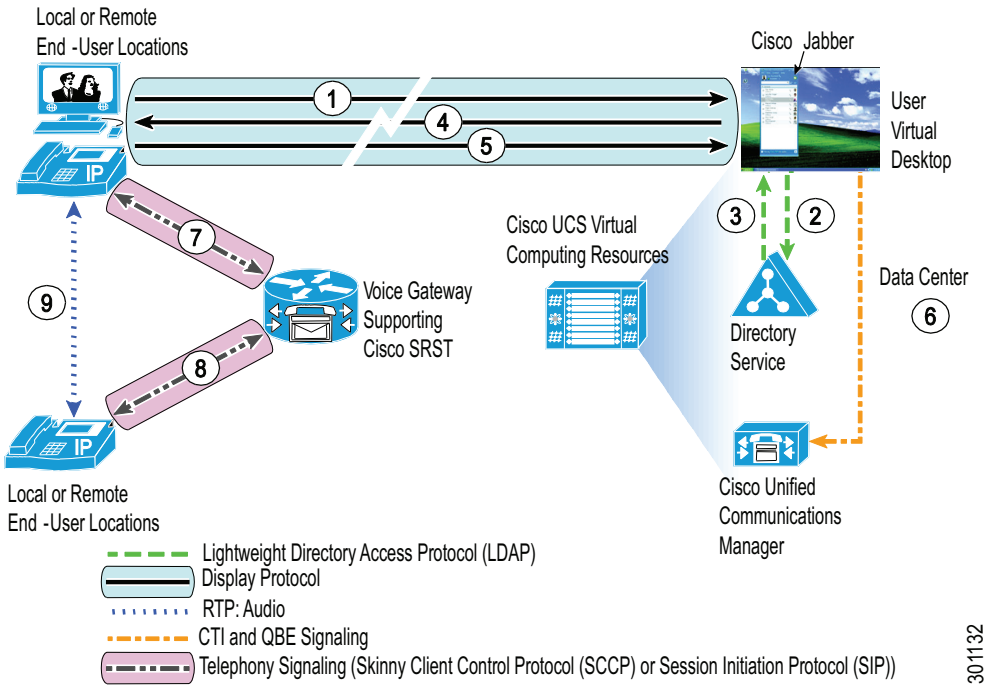
Note SSO for the HVD and unified communications applications is supported only in deployments in which each end user has a dedicated desk phone or unified communications endpoint. SSO is not available in cases in which extension mobility is implemented. If extension mobility is used, the end user will have a two-step log-on process in which the user will need to log on to the hard desk phone and then log on to the HVD environment.

Cisco UC Applications, HVD and SRST

Cisco Unified SRST is supported in Cisco Virtual Workspace (VXI) Smart Solution if a deskphone is being used for telephony. Figure 45 shows the call flow during a failover situation in which the WAN link between the desktop virtualization and unified communications endpoints and the data center is lost. The HVD with Cisco Jabber will continue to have contact with the Cisco Unified Communications Manager co-located in the data center. However, Cisco Unified Communications Manager connectivity to the hard phone will be lost. If the end-user location is equipped with a voice gateway capable of supplying Cisco Unified SRST services, the phones can reregister with the gateway. In this way, the phones will still be capable of placing and receiving calls.

Note In the current release, the Cisco VXME is not supported in the Cisco Unified SRST configuration.

Figure 45 Cisco Unified SRST Signaling and Media Flow



It is difficult to predict how the hard-phone control configuration shown in Figure 45 will respond when network service to the data center is restored. The Cisco unified communication application has never lost contact with Cisco Unified Communications Manager. With connectivity restored, the phone can now resume its connection to Cisco Unified Communications Manager. The unpredictable behavior arises from the timing of the end user's reestablishment of connectivity with the desktop. If this occurs before the phone reconnects with Cisco Unified Communications Manager, the end user may attempt to control the hard desk phone before Cisco Unified Communications Manager is aware that the phone is

present. The best way to avoid such a problem is to verify that the Cisco unified communication application correctly reflects the status of the phone by taking the phone off the hook and confirming a status change in the application. In some rare cases, the status may not resynchronize, and the Cisco unified communication application may need to be restarted within the HVD to reestablish the proper status of the phone.

For more detailed information about Cisco Unified SRST, please see the Cisco Unified SRST configuration guide at http://www.cisco.com/en/US/docs/voice_ip_comm/cusrst/admin/srst/configuration/guide/SRST_SysAdmin.pdf.

For basic configuration information to deliver a unified communications video and voice call solution using desk-phone control in a hosted virtual environment, see the Cisco Virtual Workspace (VXI) Smart Solution configuration guide at http://www.cisco.com/en/US/docs/solutions/Enterprise/Data_Center/VXI/configuration/VXI_Config_Guide.pdf.

Cisco Virtual Workspace (VXI) and BYOD Smart Solutions

Advances in technology over the past few years, especially in computing and mobility, present new possibilities for workers and allow them to be more productive away from the office. Yesterday's IT department provided and managed uniform computing devices (Microsoft Windows based laptops and desktops) are giving way to a variety of alternatives such as tablets and smart-phones that are more likely to be user owned and managed. IT departments rely on virtualization technologies such as VDI to enable such workers to access their corporate work spaces and applications from these user-owned devices using VMware View client.

Cisco BYOD Smart Solution provides the architecture and design guidance necessary for secure inclusion of workers' non-enterprise provided computing devices into the enterprise. Cisco BYOD and Cisco Virtual Workspace (VXI) Smart Solution, therefore, can provide both the flexibility and the worker productivity beneficial to the enterprise.

Devices that are brought on-board rely on VMware View client for virtual desktop and applications access and on Cisco Jabber for collaboration. This access is generally done over wireless networks. Quality of Service over wireless networks is analogous to its wired counterpart but applications need to specifically be enabled for it. The Cisco BYOD Smart Solution specifies the wireless QoS profiles that should be used for good user experience. However, individual applications need to mark their packets appropriately to take advantage of QoS provided by Cisco Wireless LAN Controllers. At this time, the VMware View client does not use IEEE 802.11e and is not able to use Wireless Multimedia (WMM) prioritization features. Cisco Jabber, however, is IEEE 802.11e enabled and its traffic can be prioritized over others in the wireless medium. For this reason, as well as to avoid hairpinning as discussed in this chapter, users should run Cisco Jabber not on their virtual desktops but natively on their access devices.

Cisco BYOD Smart Solution also defines Full, Limited, and Basic network access for wired and wireless devices. These recommendations can be used by your organization's policy on access. For example, if you wanted to allow employee-owned on-boarded devices to only be able to access virtual desktops created for the employee, you can define appropriate Access Control List (ACL) commands to restrict access for that device. Please also see BYOD section in the Virtualization Aware Network chapter in this document.

The Cisco BYOD Smart Solution Design Guide is available here,

http://www.cisco.com/en/US/docs/solutions/Enterprise/Borderless_Networks/Unified_Access/byoddg.html

For more information on the Cisco BYOD Smart Solution see,

http://www.cisco.com/go/byod

Unified Communications Enabled Accessories

With VXME several optional accessories such as a keyboard, a mouse, a handset, a hands-free speakerphone, and a high-definition camera are offered. The user may choose to use the handset, the hands-free speakerphone, or a headset for calls. The keyboard has an LCD display that shows the status of VXME and has special keys that lets the user choose between their handset, hands-free, and headset options. The user can also answer incoming calls, mute their microphone or video using keys on the keyboard.

For more details, please go to:
http://www.cisco.com/en/US/prod/voicesw/uc_endpoints_accessories.html

Securing Cisco Virtual Workspace

Overview

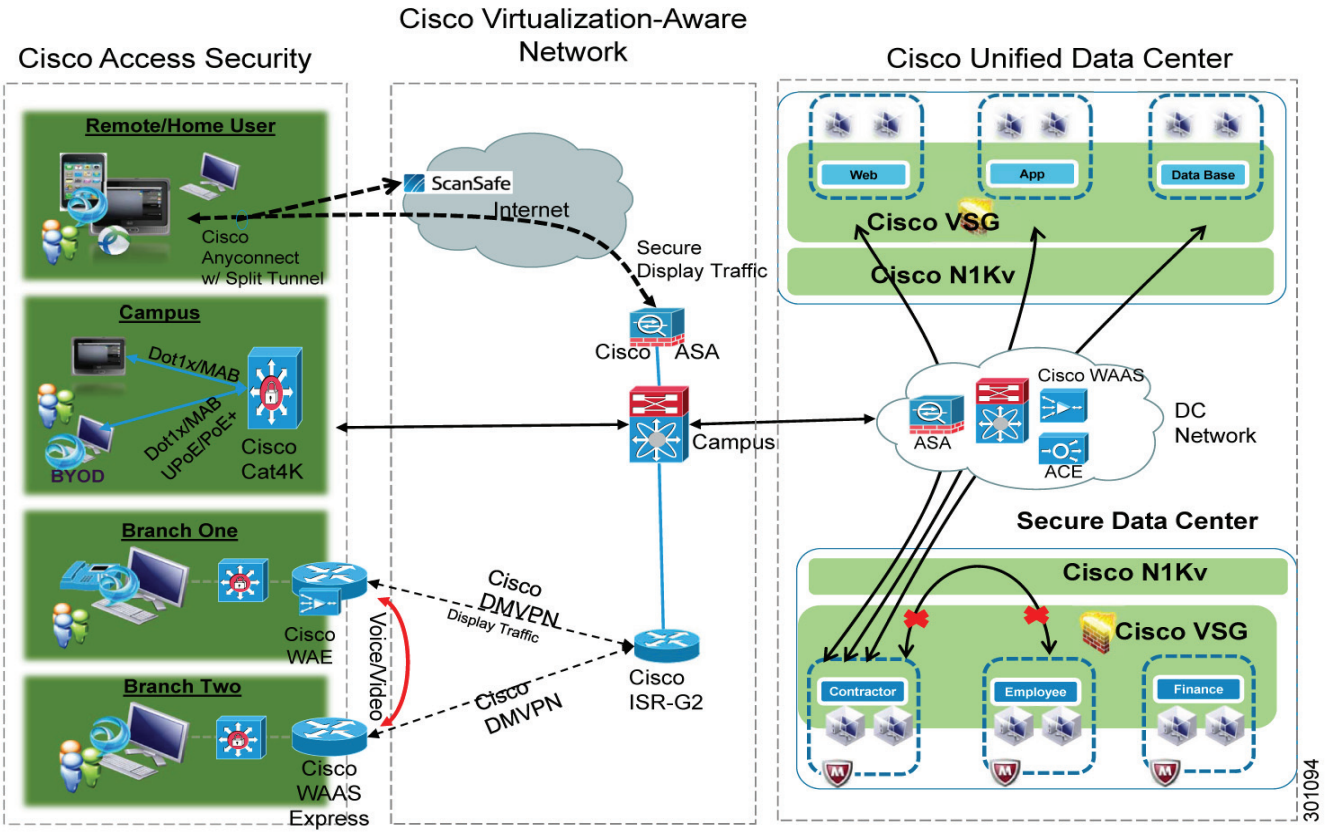
Virtualization presents a shift in the location of user data and so requires a complementary shift in thinking about security. IT data centers were traditionally built with the idea that an end user, in most scenarios, does not need to access data center components directly, and only user applications over controlled access mechanisms interact with servers in the data center. This idea is no longer true, and the separation between the user desktops (which are untrusted computing environments) and data center components needs to be ensured. In addition, because all user virtual desktops are now clustered together in the data center server, the effects of a security breach can be much more widespread than in traditional desktop environments. The Cisco® Virtual Workspace (VXI) Smart Solution addresses these new security challenges by leveraging a robust set of security features and guidelines. This chapter discusses these features along with the best practices for securing a Cisco Virtual Workspace (VXI) Smart Solution. Note that these security mechanisms are consistent with traditional, existing enterprise security technologies.

Cisco Virtual Workspace End-to-End Security Architecture

Figure 46 shows three logical sections of the Cisco Virtual Workspace that need to be secured:

- Access security focuses on securing everything originating from the endpoint to the first-hop network element, and spans all types of deployment scenarios.
- Network security covers data travelling over both untrusted (such as the Internet) and trusted (such as campus) networks.
- Data center security includes securing virtual machine access, segmentation and zoning among virtual servers, and user and data security on hosted virtual desktops (HVDs) and hosted shared desktops (HSDs) in addition to data center network security

Figure 46 End-to-End Cisco Virtual Workspace Security



The rest of this chapter discusses specific technologies and design strategies of the Cisco Virtual Workspace (VXI) Smart Solution that address these security needs.

Secure Unified Workspace: Design and Best Practices

Characteristics of a secure unified workspace include:

- Stateless endpoints
- Endpoint lockdown
- Secure access by the endpoint to the enterprise

Stateless Endpoints

Although thin and zero clients can be secured more easily than other devices, protection of the embedded OS of thin clients is of concern. If this OS were compromised, the client might not be able to reach the data center and associated data and desktop applications. The easiest way to address these concerns is to eliminate write access to the OS of the thin client. Whether or not this feature is enabled by default depends on the client manufacturer, but when it is enabled, this feature prevents write operations to the local boot media. This feature does not prevent viruses for being installed into memory, but it does mean that a simple reboot of the unit will clear the virus and return the unit to its clean state.

In an ideal situation, the endpoint would be completely stateless: no memory, no storage, no OS, and no local configuration required. In the endpoints that were tested with the Cisco Virtual Workspace (VXI) Smart Solution, this was not the case, but management software and centralized configuration can again reduce any negative effects, as described in Table 16.

Table 16 Making an Endpoint Stateless

Endpoint Vulnerability	Possible Work-around
Local memory	Typically not available for use directly; if local memory is infected, it often can be corrected with a reboot of the client
Local storage	Not present in most zero clients and thin clients (most boot from flash memory or from the network)
Local OS	Less likelihood of direct infection of the OS without a browser or email
Local applications	Access to the local desktop for installation can be controlled (steps required vary by vendor)
Connection Configuration & Endpoint Settings	Should be centralized through an endpoint management package.

Endpoint Lockdown

Controlling the local resources that an endpoint is allowed to use is a primary security concern when the desktop doesn't "own" the data. Different types of endpoints limit these resources in different ways. For example, clients can limit access to local printing resources, or they can disable USB ports so that they cannot be used to capture data on USB drives or external hard drives.

Table 17 Endpoint Device

Endpoint Vulnerability	Possible Work-around
Unused USB Ports	Group Policy Settings are used to control USB Redirection: http://pubs.VMware.com/view-50/index.jsp?topic=/com.VMware.view.administration.doc/GUID-0AD7962F-22DC-4FC1-B31B-D48946BF1D47.html
Local OS / hard drive for downloading "personal" applications	Approaches vary by thin client vendor.
Firmware Updates Required	Firmware updates should be handled by centralized endpoint management software such as Cisco VXC Manager.

Secure Access into the Enterprise

Access security in Cisco Virtual Workspace focuses on protecting data and identity while it travels between the endpoint and the data center. Since the enterprise supported applications and OS are present in the data center and not locally, the impact of a breach of endpoint security depends on how end-to-end security is implemented all the way into the data center.

Table 18 Endpoint Location

Endpoint Location	Security Implementation and Monitoring
Campus	• Endpoint authentication: – Cisco ISE with IEEE 802.1x – Cisco AnyConnect with Network Access Manager (NAM)* – MAC address bypass (MAB) – WebAuth if device supports a browser • User authentication: – Microsoft Active Directory (AD) – SmartCards
Branch to Campus	• DMVPN (IPSec) Tunnels to encrypt traffic to Campus • Endpoint Authentication Locally (see above)
Fixed Teleworker	• CVO Router (VPN) • VXC VPN (Dual and Single Tunnel Modes)
Mobile Worker	• Device Authentication via ISE • Cisco AnyConnect 3.1 (VPN Access)

*Cisco AnyConnect 3.1 with NAM can be used if a native IEEE 802.1x supplicant is not available on the endpoint.

If the endpoints deployed do not support any form of IEEE 802.1x supplicant, then IEEE 802.1x-based MAB feature on the access switches should be employed. In MAB, a database of trusted endpoint MAC addresses maintained in the RADIUS server running on Cisco ISE 1.1.1 is used to authenticate the MAC address of the endpoint. For more information about configuring and using IEEE 802.1x with MAB, see http://www.cisco.com/en/US/docs/switches/lan/catalyst4500/12.2/31sg/configuration/guide/dot1x.html

In addition to authentication of the device as a network entity, devices should be monitored for continued authenticity as close as possible to their access point as possible. This is accomplished through the configuration of port security features such as Dynamic Host Configuration Protocol (DHCP) snooping, dynamic Address Resolution Protocol (ARP) inspection, and IP source guard discussed later in this chapter.

Table 19 lists a variety of endpoints supported in Cisco Virtual Workspace (VXI) Smart Solution and corresponding VPN and device authentication methods.

Table 19 Cisco Virtual Workspace (VXI) Smart Solution Supported Endpoints

Endpoints	VPN and Device Authentication	Special Considerations
Thick clients (Microsoft Windows 7 and XP)	Cisco AnyConnect 3.1 for SSL VPN (tunnel mode) and IEEE 802.1x for authentication.	Cisco Catalyst® 3000 or 4000 Series switches at access (see note at the bottom of this table for Cisco AnyConnect 3.1 installation recommendations).
Apple iPad	Cisco AnyConnect 2.5	Available from Apple App Store

Endpoints	VPN and Device Authentication	Special Considerations
Linux and MacOS thick clients	Cisco AnyConnect 3.1 for SSLVPN remote access only	Cisco ASA profile required to support this mode (see note at the bottom of this table for Cisco AnyConnect 3.1 installation recommendations).
Microsoft Windows Mobile platforms	Cisco AnyConnect 3.1 for SSLVPN remote access only	Cisco ASA profile required to support this mode (see note at the bottom of this table for Cisco AnyConnect 3.1 installation recommendations).
Thin clients - that cannot load Cisco AnyConnect 3.1, but have native IEEE 802.1x support	SSLVPN clientless remote access to the Cisco ASA.	Endpoints in this category: Wyse ThinOS clients
Thin and zero clients with no IEEE 802.1x, No Cisco AnyConnect	MAB for port authentication; Thin clients with a browser can use WebAuth to do user authentication at the port level.	Cisco Catalyst 3000 or 4000 Series switches for access.
Administrators can download the Cisco AnyConnect 3.1 from Cisco.com. Administrators can upload it to their Cisco ASA device, and it will then be downloaded to the end-points. This option is the preferred approach. Alternatively, AC 3.1 can be deployed on the endpoint by using a MicroSoft installer package distributed using existing software distribution mechanism.		

Policy Based Network Access Control Using Cisco Identity Services Engine

Policy-based network access control in Cisco Virtual Workspace (VXI) Smart Solution allows endpoints accessing virtual desktops to be controlled based on device type, location of the network attachment point, etc. For example:

- Contractors connecting their own laptops to the network may not be given access to Cisco Virtual Workspace (VXI) Smart Solution resources
- Employees with enterprise-provided endpoints that connect to the same network port will get automatic access to their virtual desktops.

Cisco Virtual Workspace (VXI) Smart Solution also enables the enterprise to deploy bring-your-own-device (BYOD) models in which the end users are responsible for the physical endpoint and the local data, and the enterprise helps ensure that the user's work environment is securely delivered to the endpoint on a per request basis. Cisco BYOD Smart Solution and Cisco Virtual Workspace (VXI) Smart Solution allow the enterprise to reduce endpoint hardware and software management costs, and provide robust data security options.

The main components involved in creating a policy-based network access solution are Cisco ISE 1.1.1, Cisco access devices compatible with Cisco ISE (most Cisco Catalyst access switches and wireless access points are capable), and a network designed to isolate traffic. The following example describes the process for giving a personal device and an enterprise asset access to different parts of the network based on policy decisions made by Cisco ISE.

1. Cisco ISE device profiling allows the network administrator to match the device against single or multiple pre-defined attributes in Cisco ISE, and determines whether the device is personal or an enterprise asset.

2. The access switch is configured with IEEE 802.1x or MAB to authenticate the endpoints trying to access the network. When the device tries to access the network for the first time, the switch sends the MAC address and, optionally, other attributes such as DHCP requests to Cisco ISE.
3. Cisco ISE profiles and identifies the device by looking up the device data-base.
4. The employee-supplied device is allowed restricted access that transports all endpoint traffic to the data center and directly to their virtual desktop. The enterprise device is restricted only by the users' credentials.
5. To further enhance data security, the connection broker serving the restricted subnet can be pre-configured to not allow any USB-attached devices to the end device. This configuration helps ensure that no data is transferred from the HVD.

For more information on Cisco ISE and details around device isolation techniques please refer to:

http://www.cisco.com/en/US/docs/security/ise/1.1/compatibility/ise_sdt.html

http://www.cisco.com/en/US/docs/solutions/Enterprise/Security/TrustSec_2.0/trustsec_2.0_dig.pdf

When using IEEE 802.1x for device authentication, you can use multiple authentication techniques:

- Extensible Authentication Protocol (EAP)
- Protected EAP (PEAP)
- Wi-Fi Protected Access (WPA)

Devices that support IEEE 802.1x can be integrated, for example when deploying BYOD, but you should validate such devices for the environment. A combination of four attributes is used to narrowly identify the endpoint origin and authenticity. These attributes are used for device authentication and not for user authentication. The attributes are MAC address organic user interface (OUI), MAC address, DHCP user class identifier and DHCP class identifier. All the attribute information is sent to the Cisco ISE device profiling service. If more attributes are used, the access switches need to be configured appropriately as described in the Cisco ISE administration guide: http://www.cisco.com/en/US/docs/security/ise/1.1/compatibility/ise_sdt.html.

Secure Borderless Networks: Design and Best Practices

Branch-Office Access Security

Branch-office access is part of the corporate security environment, but it is separated from it by a WAN link. Branch offices typically are deployed with Cisco Integrated Services Routers Generation 2 (ISR G2) with encrypted tunnels to the campus or data center. Because the branch office typically services its own local DHCP requests, the following inspection configurations should be deployed on the access-layer switches:

- DHCP snooping
- dynamic ARP inspection
- IP source guard

MAB and IEEE 802.1x for device authentication should also be deployed locally, through branch-office switches or switch modules in the Cisco ISR.

Dynamic Multipoint VPN

Cisco Virtual Workspace supports multiple VPN options for transmission from the branch office to the campus. DMVPN is the only IPsec site-to-site solution that officially supports integration with quality of service (QoS). Cisco Wide Area Application Services (WAAS) and QoS need to be configured for each branch office to optimize the voice and video traffic for all the branch-office users.

DMVPN requires a Cisco Aggregation Services Router (ASR) at the headend and supported versions of Cisco ISR G2 routers in the branch offices.

Refer to the Virtualization Aware Network chapter for information about the deployment of DMVPN. Cisco Security Manager must be used to configure DMVPN. Also refer to the following documentation:

http://www.cisco.com/en/US/products/ps6658/index.html

http://www.cisco.com/application/pdf/en/us/guest/netsol/ns171/c649/ccmigration_09186a008075ea98.pdf

The network security requirements for the campus are exactly the same as in any existing campus environment. The details of this deployment can be found in Chapter 8 of the Cisco SAFE Reference Guide, at: http://www.cisco.com/en/US/docs/solutions/Enterprise/Security/SAFE_RG/SAFE_rg.pdf.

Secure Access for Fixed Teleworker and Home-User Environments

Cisco Virtual Office

The Cisco Virtual Office solution uses a pre-existing Cisco Virtual Office deployment design that allows VPN access from a home office by means of an access router. This deployment creates an encrypted VPN tunnel over which the endpoint can access the virtual desktops in the data center. The local network behind the access router can be considered trusted because deployment of a home-based router is an extension of the secure corporate network into the employee's home. All devices attempting access to corporate network services require device authentication using IEEE 802.1x or MAB. Note that some configurations of Cisco Virtual Office allow untrusted endpoints to directly access the public Internet outside the VPN tunnel. Many routing platforms are supported for Cisco Virtual Office in teleworker and branch-office environments. The appropriate platform is selected based on throughput required. For a typical teleworker environment the Cisco 881 ISR platform is recommended. For general information about Cisco Virtual Office solution, please visit: http://www.cisco.com/go/cvo.

For zero clients that do not support a local browser and are unable to authenticate using a web-based proxy, the Cisco Virtual Office router can be configured with specific ports opened, along with MAB. The administrator has three options listed below to ensure secure deployment:

- Use IEEE 802.1x to protect all the ports and dynamically assign the VLAN based on the device (IEEE 802.1x authentication or MAB). This is the recommended deployment option. This option performs the following operations:
  - The router tries to determine whether the connected endpoint is IEEE 802.1x capable. Table 19 earlier in this chapter shows the support matrix.
  - If the supplicant is not available, or if invalid credentials are supplied, the device's MAC address is verified with Cisco ISE.
  - If MAB fails, the port is provided with only guest access, or it can be configured to be disabled.
- If the endpoint does not support IEEE 802.1x and MAB is not acceptable, the authentication proxy feature can be used after the authentication proxy inbound access list is modified to open the ports listed in Table 19. This method requires a local browser on the endpoint for user authentication.

- Scenarios in which the preceding two options cannot be used, dedicated switchports placed in specific VLANs to allow traffic destined for specific ports can be used. This option does not authenticate the user to the network, but it can be deployed in scenarios in which it is acceptable.

Table 23 lists the ports that need to be opened for authentication proxy or dedicated switchports.

Table 20 Authentication Proxy

Traffic Description	Port Number
VMware View web	80
VMware View https	443
PCoIP USB traffic	32111
PCoIP old port, being phased out	50002
PCoIP TCP traffic	4172
PCoIP UDP traffic	4172

Secure Remote Access for Mobile Environments

Secure remote access for mobile environments can be provided with:

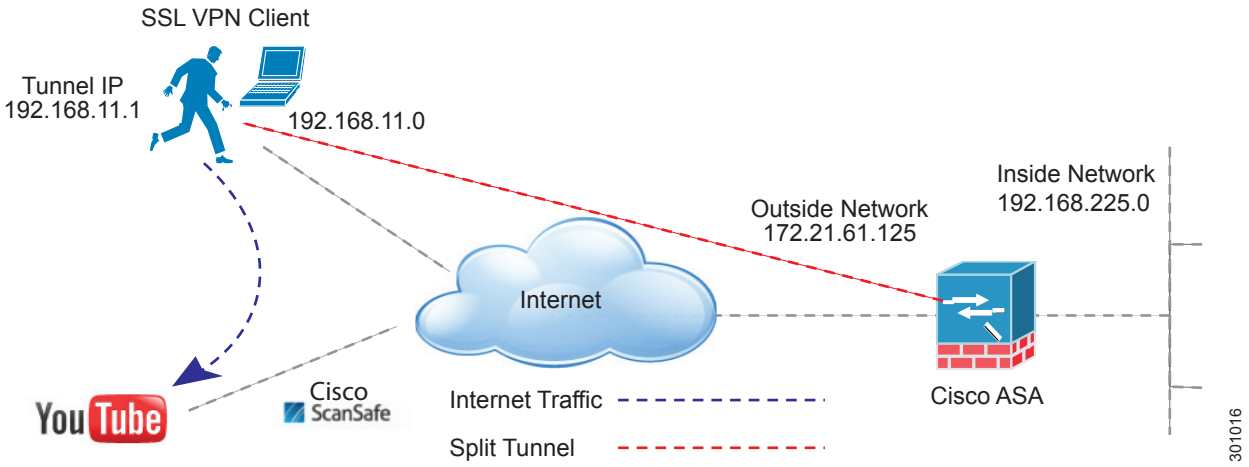
Cisco AnyConnect 3.1 and Cisco ScanSafe

Cisco AnyConnect 3.1 and Cisco ScanSafe

In a mobile teleworker environment, virtual desktops typically are accessed over an untrusted public network. A mobile worker is expected to use a thick endpoint such as the Cisco VXC 4000 Series, a laptop, or a third-party tablet to access the virtual desktop. Display traffic in this scenario needs to be protected starting at the endpoint. Cisco recommends creation of a VPN tunnel into the corporate network using Cisco AnyConnect 3.1 (or 2.5 in the case of Apple iPads) and authenticating against a directory or network access control database. Further, to increase cost effectiveness, all Internet traffic not required to go through the corporate network should be routed directly to the Internet outside the VPN session and protected using Cisco ScanSafe. The Cisco AnyConnect client should be configured to automatically launch the desktop client after the tunnel is established. For a sample configuration, see http://www.cisco.com/en/US/products/ps6120/products_configuration_example09186a00808efbd2.shtml.

The Cisco AnyConnect connection typically terminates at a Cisco ASA device at the Internet access edge of the data center. A user challenge (passcode request) is presented, authentication takes place, and an encrypted tunnel is built to allow the user access. Because display traffic is latency sensitive, you should use DTLS. The Cisco AnyConnect client and Cisco ASA can be configured to set up DTLS as a primary connection option. Using secure split tunneling and the Cisco AnyConnect 3.1 web security module for Cisco ScanSafe, all nonenterprise-bound Internet traffic (ports 80, 443, and 8080) is routed outside the VPN tunnel through a Cisco ScanSafe cloud (Figure 47). Cisco ScanSafe allows the enterprise to extend policies for Internet-bound traffic to the cloud without handling the traffic directly. This capability has a direct effect on the cost of the bandwidth in the enterprise and provides better control. Secure split tunneling installs routes on the local VPN adapter to allow only “interesting” traffic to the Cisco ASA, with the rest of the traffic remaining local.

Figure 47 Using Cisco AnyConnect with Cisco ScanSafe



Data Center Security – Design and Best Practices

User and Data Security in the Virtual Desktop

SData center security includes:

- Securing virtual desktop access to the virtual access network
- Controlling unwanted access among virtual desktops
- Controlling unwanted access between virtual desktops and application servers

Securing Virtual Desktop Access

Typically, authentication of a computing device (such as a desktop PC) is considered to occur at the access level in the network. The Cisco Virtual Workspace (VXI) Smart Solution also has an access level in the data center. With the deployment of virtual machines on large server farms, you must be sure to monitor not only the physical connections, but also the virtual connections. Monitoring should start within the hypervisor-based virtual switch, just as if it were a physical switch in any other network. The virtual desktop addresses can be dynamic and require the same level of surveillance as for desktop devices outside the data center infrastructure. To reduce the vulnerability to spoofing, Cisco recommends implementing an intelligent Layer 2 virtual switch within the enclave of virtual machines. You can deploy a Cisco Nexus® 1000V Switch, with its capability to employ the safeguards of port security, DHCP snooping, dynamic ARP inspection, and IP source guard. This strategy can provide an effective first line of verification to help ensure that the machine originally associated with a port is not replaced by a rogue machine.

Cisco Nexus 1000V Switch

Cisco Nexus 1000V is central to providing advanced security features and is a required component for Cisco VSG deployments. After the traffic from the virtual machines is switching in the virtual environment as expected, Cisco Nexus 1000V security features should be turned on. Port security, DHCP snooping, dynamic ARP inspection, and IP source guard are the features that should be enabled to secure the virtual environment. Information about configuring these features is presented in the Cisco Nexus

1000V security configuration guide at http://www.cisco.com/en/US/docs/switches/datacenter/nexus1000/sw/4_0_4_s_v_1_3/security/configuration/guide/n1000v_security.html.

For the general Cisco Nexus 1000V deployment guide, see http://www.cisco.com/en/US/prod/collateral/switches/ps9441/ps9902/guide_c07-556626.html.

- **Port security**: This feature is used to restrict inbound Layer 2 access from a restricted set of MAC addresses, and it also helps ensure that a MAC address is allowed from only one port in the same VLAN. Enable the sticky MAC address learning method for port security so that when a fresh virtual desktop is created, its MAC address is learned for the life of the virtual machine. Also, where possible, you should keep all similar virtual machines in a cluster in the same VLAN. Port security should be turned on by default on all virtual Ethernet (vEth) ports on the Cisco Nexus 1000V.
- **DHCP snooping**: This security feature protects the DHCP servers and IP resources from being compromised by learning and filtering DHCP messages for each port. All the vEth ports are untrusted by default, and in a virtual environment this default setting should never be changed. If the Cisco Nexus 1000V Virtual Supervisor Module (VSM) is installed on the same hypervisor as the Cisco Nexus 1000V Virtual Ethernet Module (VEM), the VSM port should be configured as trusted. This is the only exception to the preceding recommendation.
- **Dynamic ARP inspection**: Dynamic ARP inspection is a highly recommended security feature for the environment, especially given the number of virtual desktops that can be affected by a simple ARP spoofing attack from a compromised virtual machine. DHCP snooping is required for ARP inspection to work. Before enabling ARP inspection, the trusted uplink ports and PortChannels should be identified carefully. After ARP inspection is enabled, all the uplink ports must be placed in trusted mode. As with DHCP snooping, all host vEth ports must always be set to the untrusted state. You should create a virtual service domain on the Cisco Nexus 1000V and place all the uplink and PortChannels in this domain to help ensure that the port state (trusted) is consistently applied across the deployment.
- **IP source guard**: This feature allows filtering of all IP traffic whose IP address and MAC address does not match the DHCP snooping binding. After this feature is enabled, all untrusted ports on the Cisco Nexus 1000V can start filtering the IP traffic. DHCP snooping is a prerequisite for this feature. Note that when a port comes up for the first time, it may take a few seconds before IP traffic is allowed; this time is needed because DHCP bindings are created as a first step.

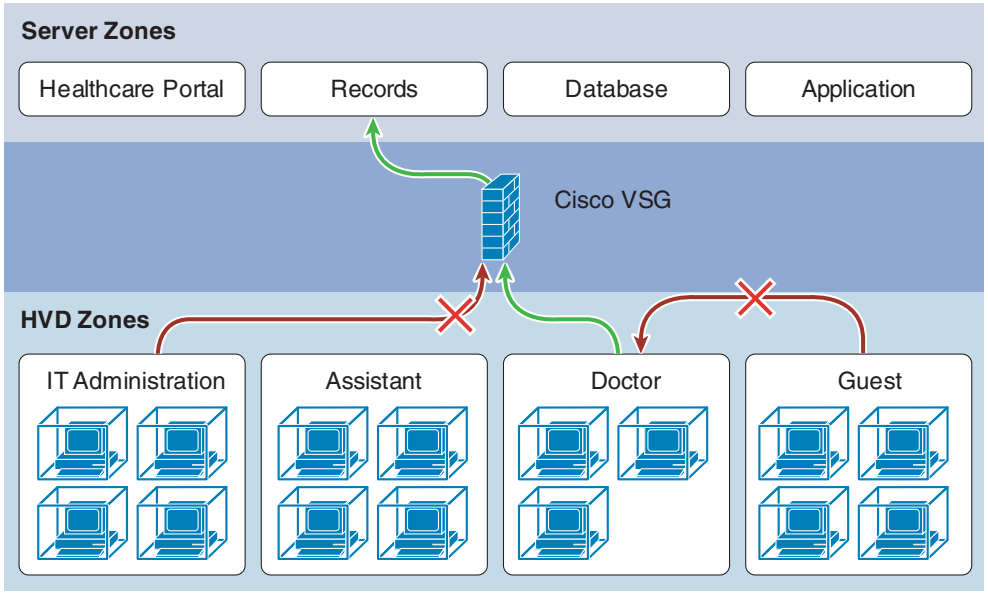
Controlling Unwanted Access Between Virtual Machines

In the Cisco Virtual Workspace (VXI) Smart Solution, a large number of virtual desktops are present in the same data center and very likely connect to the same VLAN spanning multiple servers using a distributed virtual switch. Traffic to and from a given virtual desktop needs to be isolated and controlled from other desktops and from existing critical data center resources. Performing these actions close to the source of traffic, preferably at the virtual network interface card (vNIC) level, provides the best traffic control and eliminates the need to send the traffic to a physical firewall for inspection. Virtual machines containing the virtual desktops may move within a data center for various reasons, and these movements need to be transparent from both the user experience and security perspectives. Automatic transition of security policies applied at the vNIC level, so that policies move with the virtual machines, is required in this environment. To meet all these critical security requirements, a virtual zone-based firewall that can handle movements of virtual machines automatically is required. A firewall for the Cisco Virtual Workspace (VXI) Smart Solution is best implemented by deploying Cisco VSG. This virtual firewall is tightly integrated with the Cisco Nexus 1000V distributed virtual switch and provides the critical security features.

Virtual Security Gateway (VSG)

Figure 48 shows an example of Cisco VSG zoning in a healthcare setting. For detailed configuration information and deployment considerations, including up-to-date version compatibility information, see http://savbu.cisco.com/index.php/nexus-1000v-homepage/n1kv-sales/n1kv-sales-files/1001-nexus-1000v-virtual-security-gateway.

Figure 48 Virtual Security Gateway (VSG)



Policy recommendations specific to Cisco Virtual Workspace (VXI) Smart Solution include:

- Endpoint connection to the virtual machine over the display protocol of choice
- Virtual machine connection to Active Directory, VMware vCenter, DNS, NetBios, and DHCP servers (you should create object groups representing all these infrastructure services)
- Virtual machine connection to the virtual desktop broker
- Virtual machine connection to the Internet (depending on the enterprise's Internet policy)
- Virtual machine connection to application servers (you should create object groups representing application servers that are outside the system)
- Virtual machine connection to unified communications components (virtual machine to virtual machine for audio, and virtual machine to Cisco Unified Communications Manager and Cisco Unified Presence for presence and unified communications control traffic)

All attempts by any HVD to directly communicate with any other HVD in the zone or in a different zone are blocked by default, and should also be logged. To log these events, create an explicit any-any deny policy that is logged. The implicit deny policy in Cisco VSG does not allow logging. Place the VMware vCenter and virtual desktop brokers in a separate zone to allow better control and also to simplify change management.

Table 21 lists ports that need to be open in the Cisco VSG firewall to allow basic communication from the HVD and applications, such as

Table 21 Ports That Need to Be Open

Ports	Usage	Applied VSG Zone
67	Bootstrap Protocol (BootP): UDP	HVD Zone - Required during bootup of Microsoft Windows based VMs
137	NetBios Name service	HVD Zone - Required during bootup of Microsoft Windows based VMs
4172	PCoIP	HVD to External, HVD to Broker Zones
80	Internet	HVD zone - depending on enterprise policy
5222, 526	Cisco Unified Presence Server (if used) for Presence and instant messaging: Extensible Messaging and Presence Protocol (XMPP)	HVD and unified communications zone
5060,5061	Jabber - SIP	HVD and unified communications zone
69	Jabber/CUPS: TFTP	HVD and unified communications zone

Note Cisco VSG and Cisco Virtual WAAS (vWAAS) are both virtual appliances that require the Cisco Nexus 1000V digital virtual switch to function. Both these virtual appliances need to intercept traffic on the Cisco Nexus 1000V and are not supported simultaneously in this release. As a workaround, if Cisco VSG is installed in the data center, Cisco WAAS appliances can be used. Cisco WAAS appliances in the data center typically intercept traffic at the data center edge. If Cisco vWAAS is installed in the data center, you should use Cisco ASA to segment and implement a firewall for the HVD traffic at the data center edge.

Securing the Virtual Desktop within the Data Center

Antivirus for Desktop Virtualization

Antivirus protection is required as users work on their virtual desktops and access unsecure areas of the network or the Internet. Traditional antivirus and antimalware software does not scale well in a dense virtual desktop environment. Any virus definition updates, on-access scans, scheduled disk scans, boot-up scans, and so forth consume significant memory, computing, and I/O resources and can severely affect the virtual machine density that can be supported. This problem is critical to address because the viability of the entire virtual desktop environment depends on achieving high densities of desktops per server while maintaining a secure virtual desktop environment.

A highly optimized and dedicated virus-scan server cluster performs the resource-intensive work of scanning the virtual desktop files, thereby significantly increasing the virtual machine density that can be achieved. The basic concept is that any new file that needs to be scanned is sent to a central virus-scan server, which scans the file only if it has not been scanned previously as a result of a request from another virtual desktop. If a previous scan of the file has already occurred, the client makes use of a cached scan result rather than triggering a new full scan of the file.

Summary

Whether you are using an existing design adapted for deployment of the Cisco Virtual Workspace (VXI) Smart Solution or a completely new green-field installation, security requires a comprehensive end-to-end approach. With an understanding of the threats present throughout and the countermeasures available to address those threats, the Cisco Virtual Workspace (VXI) Smart Solution can help ensure a secure environment from the endpoint to the data center.

Scaling and High Availability

Introduction

In an Enterprise network, performance and capacity planning for desktop virtualization (DV) services has three fundamental dimensions:

- Capacity planning of compute and storage needs
- Capacity planning of network, network services, application and other infrastructure necessary to support desktop virtualization
- WAN capacity planning for delivering the service to remote sites

This chapter explores the tools, methods, and design considerations for deploying and scaling an end-to-end Cisco Virtual Workspace (VXI) Smart Solution. Testing performed across the end-to-end system forms the basis for the recommendations and other guidance provided in this chapter. Detailed results from the testing are available in a separate document published along with this document: Cisco VIRTUAL Workspace (vxi) Smart Solution 2.7 Performance and Capacity Results Guide for VMware.

Note that the test results and the related design guidance are based on a user workload that includes Cisco collaboration applications and antivirus software as a standard part of the workload. You will see throughout this chapter that the workload is a critical factor in any data or design guidance for capacity planning a desktop virtualization deployment. The user workload is effectively what the average user does on a desktop and can vary between enterprises and between groups within an enterprise. Therefore, you can use the information here to guide you in your capacity planning, but you should carefully consider your own workloads and environments and make the necessary adjustments for your own deployments. This chapter is not intended as a comprehensive guide for scaling every Cisco and third-party component used in the Cisco Virtual Workspace (VXI) Smart Solution. For this guidance, read the documentation for the individual products.

Capacity Planning - Compute and Storage

When migrating users from physical desktops to virtual desktops, you need to have a good understanding of the user base and the resource requirements for CPU, memory, disk space and disk I/O requirements. Many enterprises transitioning to desktop virtualization also see it as an opportunity to migrate to Microsoft Windows 7, so understanding the implications of such a transition is also important.

A crucial step in the process is grouping the users being migrated according to common application, workload, and desktop (dedicated or personalized) needs so that capacity planning for computing and storage needs can be performed at the group level. Grouping enables the administrator to develop a base profile for each user group based on common factors such as the same operating system, applications, and use patterns. The base profile should also include resource metrics such as CPU, memory, network bandwidth, and disk I/O utilization that are representative of the group by collecting data over a period

of time from a statistically significant number of users from within that group. The base profile can then be used to estimate the computing and storage needs of the user group in a virtualized environment. If disparate user groups are not estimated separately, the enterprise could end up with wasted capacity, or if capacity is underestimated, resource constraints could result in a poor user experience. Therefore, from an overall resource estimation and capacity planning perspective, administrators should group users with similar workloads and environments rather than across the entire user base being migrated, particularly in large desktop virtualization deployments.

The next several sections explore the steps involved in planning for the computing and storage needs of an end-to-end Cisco Virtual Workspace (VXI) Smart Solution deployment. The approach taken can be summarized as follows:

- Develop a profile of the users in their physical desktop environments that includes a workload profile and resource needs based on CPU, memory, storage, and bandwidth utilization.
- Group users based on common factors (workload profile, OS type, etc.) that affect computing and storage needs, including desktop factors that become important with virtualization (for example, pooled or dedicated resources and persistent or non-persistent memory). The common factors that define the group will form the base profile for the group.
- Using the base profile, estimate the resource requirements for each user group in a virtualized environment. These requirements should factor in optimizations and other best practices that can improve resource utilization and performance of the virtual desktop.
- Estimate the per-server capacity in terms of the number of virtual desktops based on the estimated resource needs. This estimate will depend on the server and its capabilities. Identify the storage capacity and I/O performance required to support these users; this data can be used in sizing the overall storage needs of the virtual desktop deployment. This estimate should take into account factors such as the effect of peak use, login storms, and outages and server downtime.
- Validate the single-server estimations using a workload profile that is representative of the particular user group. Perform this validation either in the lab or in a live pilot deployment with a small subset of users.
- Extrapolate from the single-server data to determine the overall computing and storage hardware needs for the larger group.
- Repeat the process for other groups in the overall deployment.

User Workload Profile

A base profile used in capacity planning should include the workload profile in terms of the applications, the activities within those applications, and a use pattern that is representative of the user group. The group's workload profile can then be used to classify the user group into one of the generic workload profiles commonly used by vendors to characterize users with similar computing, storage, and networking needs. The generic profiles form the basis for performance and scalability results used in capacity planning, and benchmarking tools used in server and storage performance testing use these profiles to generate a corresponding workload. The intensity of the workload can significantly affect the scale and capacity results, so the workload profile is critical to any data used for estimating computing and storage resource requirements. The workload profile is also critical if any testing is performed in the customer environment to validate the resource estimations because the workload profiles used by load-generation tools should closely match the user group's workload profile to accurately size the environment. Workload profiles thus play a critical role in any data used for capacity planning, and therefore in the accuracy of the sizing estimation used to determine the server and storage needs of the desktop virtualization system being deployed. Because it is difficult to accurately reflect the workload generated by users in different enterprise environments using a generic workload profile, any capacity estimation based on data using a generic workload profile should be adjusted to accommodate any

differences. A more conservative estimation should be implemented, and resource utilization should be monitored closely over a period of time to make any adjustments needed to the estimated server and storage needs.

Note that the workload used in the Cisco Virtual Workspace (VXI) Smart Solution for scaling and performance benchmarking includes the following applications that are typically not seen in similar benchmarking tests from other vendors:

- Cisco Collaboration applications: Cisco Jabber for Windows or Cisco Unified Personal Communicator in desk-phone control mode is used in the workload.
- Antivirus applications: Antivirus software is part of every desktop and contributes to the workload because scanning is in effect while the user workload is being processed. Antivirus software does have a measurable effect on both server and storage performance.

For more information about the workload used in the Cisco Virtual Workspace (VXI) Smart Solution for scaling and performance capacity testing, see the Workload Considerations section of this chapter.

Resource Utilization in Current Environment

An important factor for estimating resource requirements in a virtualized environment is the resource utilization in the current physical desktop environment. Therefore, for a given target user group being migrated to virtual desktops, it is important to have a full understanding of the current environment and to characterize the resource utilization in terms of the following metrics:

- Average and peak CPU utilization
- Average and peak memory utilization
- Storage
  - Per-user storage capacity
  - I/O operations per second (IOPS), broken down by read vs. write I/O. Ratio of each is critical in determining storage sizing requirements
  - I/O Throughput (in bytes per second)
- Bandwidth utilization on the LAN

Administrators should monitor the use pattern of the target user group and determine the average and peak utilization for each of these in the environment. Monitoring should factor in potential variations in use pattern based on geographical location, when users log on for the day including shift transitions for environments that work in shifts, timing of backups, virus scans, and similar activities.

The resource utilization of the average physical desktop user can be determined as follows:

- **CPU utilization**: Use tools such as Microsoft Windows Perfmon to collect average and peak CPU utilization from physical desktops in the target user group being migrated. Collect this data from a statistically significant number of desktops over a period of time while there is significant workload. You can then use a statistical average from the collected data to determine the peak and average CPU utilization for the group.
- **Memory utilization**: Also collect data about memory utilization on a physical desktop using the same tools as for CPU utilization, or a similar tool. As with CPU, you should analyze the data collected over a period of time from a significant number of desktops to determine the statistical averages of the group in terms of peak and average memory utilization. This data will be used to determine the memory needs of the group when using a virtualized desktop.

- **Storage capacity and performance (IOPS and throughput) of the physical desktop**: You can also determine IOPS and throughput from the physical desktop using the same tools as for collecting CPU and memory utilization information. Determine the peak and average data for the group in the same way as for CPU and memory utilization. This data will be used to determine the storage requirement of a virtualized desktop in that group.

Once the administrator has characterized the average and peak resource utilization for the group using a physical desktop, the process of estimating the compute, storage, and networking needs for migrating to desktop virtualization can begin.

Estimating Resource Requirements in a Virtualized Environment

To accurately estimate the resource requirements in a virtualized environment, several factors must be considered. In this section, we will take a closer look at three of these factors, namely CPU, memory, and storage. The data gathered in the previous section in terms of CPU, memory and storage can be used to estimate the number of virtual desktops a given server in the data center can support. Virtualization does introduce additional factors so the above resource requirements may need to be adjusted before estimating server capacity. Capacity of a single Server capacity can now be used to estimate hardware resources or servers necessary for a large-scale deployment.

Estimating CPU

To estimate the CPU resources needed in a virtualized environment, you can use the data from the physical desktops as illustrated in the example below. Consider the following scenario:

- Average CPU utilization of the physical desktops in the target user group = 8%
- Physical desktops are using a dual-core 2GHz processor
- VMware recommends using a guard band of 10 to 25 percent to handle the following:
  - Virtualization overhead
  - Peak CPU utilization
  - Overhead associated with the display protocol processing
  - Spikes in CPU utilization

Based on the above, the average CPU requirement of each desktop is 8% of 2x2GHz = 320GHz. With a conservative guard band of 25 percent, the average CPU requirement for each desktop = 400MHz.

You can now start sizing your server requirements using the average CPU per desktop and the compute capabilities of the server chosen for your deployment. Processor info for different Cisco UCS servers that can be used for hosting virtual desktops are shown in the tables below.

A comprehensive list of Cisco UCS server models can be found here:

- **B-series**: http://www.cisco.com/en/US/products/ps10280/prod_models_comparison.html
- **C-series**: http://www.cisco.com/en/US/products/ps10493/prod_models_comparison.html#~tab-a

Table 22 Cisco UCS B-Series Blade Servers - Models and Processor Info

Server Model	Processor
Cisco UCS B200 M3 Blade Server	Two 8- core Inte ®Xeon E5-2600 series processors

Server Model	Processor
Cisco UCS B230 M2 Blade Server	Two 10-core Intel® Xeon® processor E7-2800 series processors
Cisco UCS B250 M2 Extended Memory Blade Server	Two 6-core Intel® Xeon 5600 series processors

Table 23 Cisco UCS C-Series Rack Mount Servers - Models and Processor Info

Server Model	Processor
Cisco UCS C220 M3 Rack Server	Two 8-core Intel® Xeon® E5-2600 series processors
Cisco UCS C240 M3 Rack Server	Two 8-core Intel® Xeon® E5-2600 series processors
Cisco UCS C260 M2 Rack Server	Two 10- core Intel® Xeon® E7-2800 series processors

Note Each server model supports different processor types and speeds, though only one is shown per server in the above tables

For additional information about the Cisco UCS server chassis and blade servers, please refer to the Virtualized Data Centerchapter of this document.

Using computing power as the only criterion, you can calculate the number of desktop virtual machines on a given blade server as shown here. For example, the number of virtual desktops that a Cisco UCS B250 M2 server can support would be:

```
Total compute power = 2 socket x 6 core a 3.33GHz = 39.96 GHz
Average CPU utilization of desktop = ~400MHz
Number of virtual desktops per server = 39.96GHz/(400MHz+Overhead) = ~90 desktops
```

Note This estimate is theoretical, based on a single factor (CPU). A number of other factors need to be considered to determine the actual number of virtual desktops that can be supported on a given server blade. Please use actual data from testing when performing capacity planning for specific customer deployments.

The single server sizing estimation above could be lower or higher in your deployment depending on the average utilization of the physical desktops used. Similarly, the Cisco UCS server model chosen for a given desktop virtualization deployment can also affect the number due to differences in the computing capabilities of the different servers available on the Cisco Unified Computing System™. As new processors are released for each server model, improvements in processor designs can further increase the number of desktops supported in which case an estimation based on compute power is only a starting point as higher densities could be supported.

If computing power is the only criterion used, the above estimation for a single Cisco UCS blade server can be extrapolated to determine the overall Cisco UCS server needs of the deployment. However, a similar exercise using memory and storage is necessary to determine the limiting factor for your environment before this estimation can be used deployment wide. A number of other factors also have to be considered before an estimation can be considered final for a given server blade.

Estimating Memory

To estimate the overall memory requirements in a virtualized environment, use the same methodology used for estimating CPU. The memory estimate for a single virtual desktop can be calculated from the statistical average determined from the physical desktops, as shown in the following example.

- Average memory utilization for the physical desktops in the target user group is approximately 1GB.
- The transparent page sharing (TPS) feature on the VMware hypervisor can significantly reduce the memory footprint, particularly in desktop virtualization deployments, in which the OS and applications data loaded into memory from different desktop virtual machines on the same host may have a lot in common. However, since TPS is a mechanism for over committing memory, it is not factored into the calculation here; it is a deployment decision for the administrator to consider as a part of the overall desktop virtualization rollout.
- To accommodate additional memory demands due to spikes in memory utilization or additional applications, a 25 percent increase in the estimate is used. The aggregate memory requirement is therefore approximately 1.25GB = ~1.3G.
- The memory requirement for a virtualized desktop, along with the physical memory resource available on the blade server chosen for the deployment, can be used to estimate the number of virtualized desktops that can be supported on a given blade.

The memory capacity on the various Cisco UCS server models is listed in tables below. Please refer to the following documents for a more comprehensive and up-to-date list:

- **B-series**: http://www.cisco.com/en/US/products/ps10280/prod_models_comparison.html
- **C-series**: http://www.cisco.com/en/US/products/ps10493/prod_models_comparison.html#~tab-T

Table 24 Cisco UCS B-Series Blade Servers - Memory Capacity

Server Model	Maximum Memory Supported
Cisco UCS B200 M3 Blade Server	384 GB
Cisco UCS B230 M2 Blade Server	512 GB
Cisco UCS B250 M2 Extended Memory Blade Server	384 GB

Table 25 Cisco UCS C-Series Rack Mount Servers - Memory Capacity

Server Model	Maximum Memory Supported
Cisco UCS C220 M3 Rack Server	256 GB
Cisco UCS C240 M3 Rack Server	384 GB
Cisco UCS C260 M2 Rack Server	1 TB

For additional information about the Cisco UCS server chassis and blade servers, please refer to the Virtualized Data Center chapter in this design guide.

Using memory as the single criteria for sizing the hardware needs in a Cisco Virtual Workspace (VXI) Smart Solution deployment, you can calculate the number of desktop virtual machines that a blade server can support as shown here. For example, the number of virtual desktops that a Cisco UCS B250 M2 server can support would be:

```
Memory Capacity = 192GB
Average memory requirement for a virtualized desktop = ~1.3GB
Number of virtual desktops per server= 192G/1.3G = 147 desktops
```

As with CPU estimation, the estimated number of virtual desktops on a single server may be lower or higher, depending on the data gathered from the physical desktops and the model of Cisco UCS server selected for the deployment. Also, this data can be used to extrapolate the total number of servers needed for the deployment if memory is determined to be the limiting factor.

Note that the memory utilization from a physical desktop used in the preceding calculation can vary depending on the guest OS and applications deployed in a given environment. An alternative but also a less accurate method of estimating the number of virtual desktops is to use the minimum recommendation from Microsoft for per-virtual machine memory utilization, as shown in Table 26. However, for Microsoft Windows XP, the minimum recommendation shown here should be increased to 512 MB to accommodate Microsoft Office applications and other applications that may be running on the desktop. The memory configuration used in the Cisco Virtual Workspace (VXI) Smart Solution is also provided as an example. The memory configuration was sufficient to provide a good user experience for the workload profile validated in the end-to-end system.

Table 26 Memory Configuration

Microsoft Windows OS	Minimum (Microsoft) Memory Requirement	Memory Configuration in Cisco Virtual Workspace (VXI) Smart Solution for a Cisco Knowledge Worker + Profile
Microsoft Windows XP with Service Pack 3	256 MB	1 GB
Microsoft Windows 7 32b	1 GB	1.5 GB/ 2 GB
Microsoft Windows 7 64b	2 GB	2 GB

Estimating Storage

For storage, the average IOPS and throughput data collected from monitoring the physical desktops can be used as the storage requirements for the virtualized desktops. For example, if the average IOPS is 5 and the average throughput is 115 kbps, then the same IOPS and throughput values should be expected when the desktop is virtualized. IOPs may be lower in Virtual Machine than physical desktop, when optimization is applied (unnecessary services disabled, defragmenter/superfetch disabled). VDI Pilot pool with optimized Golden Image will be good reflection of IOPs requirements. For a desktop virtualization deployment, the factors summarized here can also have a significant effect and should be considered when sizing storage needs. For example, IOPS can peak when:

- **Users are powering on**: When users come in at the beginning of the workday and start powering on their virtual desktops, IOPS and throughput will peak, a situation referred to as a boot storm.
- **Users are logging on**: Though the virtual desktops do not need to be powered on, there can be peaks in storage I/O as users are logging on in the morning to start their work. This situation is referred to as a login storm.
- **Other activities occur**: Activities such as antivirus scan and backups can cause storage performance requirements to spike.

Some applications specific to a customer environment can cause similar spikes in storage I/O. All these factors must be taken into account when designing the storage environment for a virtual desktop deployment.

Another aspect that needs to be considered when sizing storage needs is the disk space allocated to a virtual desktop. You can calculate this space by adding the storage requirements required for each of the following items:

- Operating system and base set of applications

- Page and swap files and temporary files created by the OS and applications
- Page and swap files created by a VMware ESX and ESXi host for every virtual machine deployed on the host (equals the memory allocated for the virtual machine)
- Microsoft Windows profile (user settings such as desktop wallpaper)
- User data (equivalent to the My Documents folder in Microsoft Windows)

For an example of the storage allocation to use for virtual desktop machines, see Table 27. Note that deploying desktop virtualization using View linked cloned desktops with a provisioning server will minimize the per-desktop disk space necessary for windows and applications as large groups of desktop pools can share the same master virtual machine image. As a result, only the delta between that and the available disk space on the master virtual machine may need to be allocated on a per-desktop basis by using deployment models mentioned above. However this delta disk size can also be minimized by refreshing the OS disk periodically or by using non-persistent desktops - so a number of options exist in this regard as well.

Table 27 Storage Allocation for Desktop V

Guest OS on virtual desktop	Minimum Disk Space Microsoft Windows and Applications	Microsoft Windows Page File and Temporary Files	Hypervisor Swap File	Microsoft Windows User Profiles	User Data
Microsoft Windows XP	10 GB	3 GB	1 GB	2 GB	5 GB
Microsoft Windows 7 (32-bit)	16 GB	4 GB	1.5 GB	2 GB	5 GB
Microsoft Windows 7 (64-bit)	20 GB	4 GB	2 GB	2 GB	5 GB

Estimating Server Capacity

As stated before, several factors can influence the performance and scalability of a server. The estimation for the number of virtual desktops on a given server can yield a different number if each factor is considered independently. For this reason, the estimations performed in the Estimating CPU and Estimating Memory sections earlier in this chapter for a Cisco UCS B250 M2 server are theoretical exercises. However, the data (summarized in Table 28) can aid in finding the limiting factor for a given server, as well as provide the initial virtual machine density to target if testing is performed to validate the estimation using the specific workload for that environment.

Table 28 Estimated Capacity

Factor Used to Determine Capacity	Average Value for a Virtualized Desktop	Server Capacity
(Theoretical for Cisco UCS B250 M2)		
CPU	400 MHz	90
Memory	1.3 GB	147

Note The estimates in Table 28 are not the actual capacity of the server. They are theoretical estimations based on the CPU and memory utilization assumptions for the user group in a given environment.

The theoretical estimation shows that the limiting factor for this deployment is the compute capacity However, actual scalability and performance results on a Cisco UCS B250 M2 in the Cisco Virtual Workspace (VXI) Smart Solution using the Cisco Knowledge Worker+ workload shows the results to be different from the ones in the table above. For actual data from the testing, see the Summary of Results section later in the chapter.

The reason for the difference between the theoretical and estimated data can be attributed to workload used in the theoretical estimation vs. the one used in testing. For any desktop virtualization deployment, workload is one of the most critical factors for accurately estimating the virtual desktop capacity of a given server. Therefore, it is critical that you use a workload that closely matches the user group's workload when validating the theoretical server sizing estimation.

In the Cisco Virtual Workspace (VXI) Smart Solution, single server scalability testing was done for a number of deployment profiles that reflect how customers would deploy and use the system. The testing was done using test tools located in the campus network that initiate VDI sessions to the virtual desktops hosted in the data center where the session would span the following:

- Campus network that consists of access, distribution and core network layers built using catalyst 3500, 4500 and 6500 series layer 2 and layer3 switches
- Data center network, also with a core, aggregation and access layer made up of Cisco Nexus 5000 and
- Cisco Nexus 7000 series built in accordance with Cisco validated data center infrastructure design
- Data center Services aggregation layer where firewalls are used to control all traffic entering and leaving the data center. Firewalling is also used at the access layer within the data center to control traffic between virtual desktops from all other services and application infrastructure residing in the data center. Redundant ACEs are used for load balancing all connections to application servers used for providing desktop virtualization services such as View connection servers
- Data Center that consists Cisco UCS 5108 B-series chassis with B250 M2 and B200 M2 blade servers connected to Cisco Nexus 1000 series access switches. Both NAS and SAN based storage were used depending on the needs of the deployment profiles tested.

For details on the results and data from the single server testing done in the Cisco Virtual Workspace (VXI) Smart Solution using a Cisco KW+ workload - please see the Summary of Results section of this chapter.

In addition to the workload, there are a number of other factors apart from CPU, memory, and storage utilization can influence a server's scale and capacity numbers for a virtual desktop deployment as discussed in the next few sections.

Design Considerations - Compute

In this section, other factors that impact the scalability of deployment are addressed.

Hypervisor Considerations

Hypervisor resource consumption should be closely monitored for CPU, memory and storage performance. Memory consumption on a hypervisor should be monitored to see if any memory ballooning could be an indication that the host is in need of additional memory. Resources should also be monitored at the cluster level so that additional resources can be added if needed.

Memory Considerations

The transparent page sharing (TPS) feature available on ESXi hypervisor can significantly reduce the memory footprint, particularly in desktop virtualization deployments where the OS and applications data may have a lot in common across different desktop virtual machines. TPS uses a background process to monitor the contents of memory and evaluates the data being loaded to determine if it is the same as what is already in memory. If it is the same, the virtual machine attempting to load the duplicate data will be redirected to existing content in memory, thereby enabling memory sharing. TPS can be thought of as memory de-duplication feature and is enabled by default. For more information about this feature, see VMware document: http://www.VMware.com/resources/techresources/531

In a desktop virtualization environment, transparent memory sharing enables a server to accommodate a larger number of virtual desktops on a single blade, at least from a memory perspective though this may not be a limiting factor.

Since TPS uses redundancy to share and over commit memory between virtual machines running on a host, the workloads on these virtual machines should be as similar as possible. To optimize the effect of TPS in a virtual desktop deployment, you should group virtualized desktops of users with similar workloads, such as the same guest OS (Microsoft Windows 7 and Windows XP) and applications (Microsoft Office and antivirus applications), on the same host to optimize the effect of TPS.

TPS behaves differently on the newer hardware-assisted virtualization processors, such as the Intel Nehalem and Westmere processors that are used on Cisco UCS servers. The newer processors use memory pages that are 9KB in size and improve performance by 10 to 20 percent. TPS operates on 4-KB pages to eliminate duplicate data. With these newer processors, TPS is not in effect until the available memory reaches a minimum and there is a need to over commit memory. A background process is still monitoring and scanning the memory pages to determine when TPS takes effect. See the following VMware Knowledge Base articles for more information about TPS with the newer hardware assisted virtualization processors:

- TPS in Hardware Memory Management Unit (MMU) Systems: http://kb.VMware.com/kb/1021095
- TPS Is Not Utilized Under Normal Workloads on Intel Xeon 5500 Series CPUs: http://kb.VMware.com/kb/1020524

Also note that VMware studies have shown that TPS does not have any effect on the performance of the host and therefore recommends the use of this feature. Please contact VMware for additional information about TPS.

Power Management Policy

The power management policy for the virtual desktops in a desktop virtualization environment can have affect the host resources. VMware View recommends that the virtual desktop be put in a suspended state if it is not in use. The suspended state is an optimal configuration that enhances the user experience while reducing resource (CPU and memory) use. If all virtual machines are left powered on, the host resources cannot be used by other virtual machines on the same server. With persistent desktops, the virtual machine can be immediately suspended when the user logs off.

High-Availability (HA) Considerations - Compute

For a desktop virtualization deployment of significant scale, high availability of the virtual desktop is a concern for most administrators. Virtual desktop machines, as a best practice, are typically deployed in clusters so that a pool of hosts are available to the cluster. With clustering, VMware can distribute the virtual desktops across the pool of resources using VMware DRS to increase the resources available to the virtual machine. The clusters are also used to implement specific high-availability features such as VMware High Availability (HA), DRS, Fault Tolerance, and vMotion. However, deploying servers in a cluster can change and potentially limit the maximums that VMware supports. The supported limits are available through VMware configuration maximums and are available when new releases change the supported limits. Table 29 lists some of the data relevant to sizing a desktop virtualization deployment. Table 29 should be reviewed for planning any large-scale deployment of Cisco Virtual Workspace (VXI) Smart Solution. For a complete set of configuration maximums, refer to VMware documentation.

Table 29 Configurations Maximums

Limits	VMware vSphere 4.0 Update	VMware vSphere4.1	VMware vSphere 5.0
Number of virtual machines per host	320	320	512
Number of vCPUs per core/host	25/-	25/-	25*/2048
Hosts per high-availability cluster	32	32	32
Number of virtual machines per cluster	-	3000	3000
Number of virtual machines per host with 8 or fewer in the high-availability cluster	160	N/A	N/A
Number of virtual machines per host with more than 8 hosts in a high-availability cluster	40	N/A	N/A
Number of hosts per VMware vCenter server	1000	1000	1000
Number of hosts per data center	100	400	500

* See VMware vSphere5.0 Configuration Maximums Guide for more details: http://www.vmware.com/pdf/vsphere5/r50/vsphere-50-configuration-maximums.pdf

General Considerations

In this section, we take a look at factors that impact multiple aspects of the system, starting with compute and storage.

Guest OS Optimizations

Microsoft Windows can be optimized to improve the performance and scalability of virtualized desktops as outlined below.

- Optimize the Microsoft Windows virtual machine file system for optimal I/O performance by disabling the last-access-time updates process in NTFS. Microsoft Windows will update files with the last access update time when an application opens that file, and disabling this option will reduce the IOPS occurring within the file system.
- Enable Microsoft Windows Best Performance especially for Microsoft Windows 7 deployments it can provide performance important both from a compute perspective and from a WAN BW utilization perspective
- Refresh the OS disk periodically. In VMware View deployments, linked clones maintain a read-write file for storing temporary files or other changes that need to be made to the original OS disk on the parent virtual machine. This file is a diff file, and with time, this file can grow in size and become as large as the main OS disk. To keep this read-write file on every cloned desktop from consuming a large amount of storage, refresh the OS disk periodically to clean up the delta file and reset it to its original state, when the cloned desktop was first created.
- Prevent antivirus software from scanning the main OS disk that each virtual machine uses since it is deployed as a read-only disk with antivirus checks run against it before it was deemed as the golden master for use by the VMware View pool of virtual desktops. This step can help increase storage performance particularly in a large virtual desktop deployment.

A number of windows optimizations can be enabled on virtualized desktops to improve performance. Table 30 shows some of the main optimizations implemented in the Cisco Virtual Workspace (VXI) Smart Solution for Microsoft Windows XP and Windows 7.

Table 30 Guest OS Optimizations

Disable Microsoft Windows Hibernation
Disable Microsoft Windows Defender (N/A on XP)
Disable Microsoft Feeds synchronization
Disable Microsoft Windows Scheduled Disk Fragmentation (N/A on XP)
Disable Microsoft Windows Registry Backup (N/A to XP)
For PCoIP, set the power options for Display to off
Disable mouse pointer shadow
Antivirus scan on write only
Disable Pre-fetch/Superfetch Service (N/A on XP)
Disable Microsoft Windows Diagnostic Policy Service (N/A to XP)
Enable "No automatic updates"
Disable System Restore (since refresh can be done by composer)
Disable paging of the Microsoft Windows OS itself

Disable unwanted services
Turn off unnecessary sounds at startup and shutdown
Disable indexing services
Delete all background wallpapers
Disable screen

Design Considerations - Storage

Linked Clones

Desktop virtualization architectures solution from VMware used in the Cisco Virtual Workspace (VXI) Smart Solution reduce the overall storage needs as follows:

VMware View uses Linked Clone technology through which a parent virtual machine's virtual disk is used as the main OS disk for all clones created through the linked clone process. This feature prevents each cloned desktop from needing its own OS disk, thereby reducing the overall storage capacity needed for the deployment. See Virtualized Data Center chapter for more information on deploying virtual desktops using VMware View.

Using linked clones with VMware View greatly reduces the aggregate storage capacity necessary for migrating to a virtualized environment. Although the cost of the shared storage is significantly higher than that for using separate disks on laptops and desktops, VMware View can reduce overall storage costs due to the ability to share the same OS disks among many desktop virtual machines.

Operating System Disk

In VMware View deployments, the OS disk refers to the parent virtual machine's virtual disk on which the guest OS (Microsoft Windows XP or Windows 7) and applications (Microsoft Office) are installed. This OS disk is read by all desktops in the pool with VMware View deployments, resulting in significant storage savings, since a single OS disk can be used by a large number of desktops without each having to maintain its own OS disk. Ideally, this disk should be read-only for both storage and operation efficiency, but it can be used as a read-write disk to store the following types of typical Microsoft Windows desktop data

- Microsoft Windows profile data
- Temporary files, including page files
- User data

For better storage and operation efficiency, the OS disk should be kept as a read-only disk, and the data listed here should be redirected to another location as follows.

- Microsoft Windows profile data can also be redirected to a Microsoft Windows share or to another virtual disk dedicated for this purpose, so that the data can be saved in the event that the OS disk is updated.
- Temporary files can also be redirected to a non-persistent disk so that the data can be flushed to reduce storage use. A separate location on the SAN or on a transient volume on network-attached storage (NAS) can be used.
- User data that is typically saved in the My Documents folder should be redirected to a Microsoft Windows share or to a separate disk.

Thin Compared to Thick Provisioning

Thin provisioning is a way to conserve storage resources and increase storage utilization in a virtualized environment. With thick provisioning, when a virtual machine is deployed, the virtual disk associated with the virtual machine is given its full allocation of storage regardless of whether it uses it, resulting in wasted space. With thin provisioning, this inefficiency is reduced by allocating storage resources only when the virtual machine needs them. Therefore, a virtual desktop running Microsoft Windows 7 with a 20-GB disk will not have 20 GB of disk space reserved on the storage system (SAN or NAS), though Microsoft Windows and applications running on the desktop will operate as if it has the full 20 GB of space allocated to it. On the back end, VMware ESX and ESXi hide the actual state of the storage allocation and allocate the space to the desktop only as and when it needs it. The dynamic allocation of storage is performed in chunks, with the unit size of a chunk defined when the data store is created. This specific type of thin provisioning is referred to as a VMware virtual disk and is supported with both file-based (NFS) and block-based (SAN) data stores. For more information about VMware thin provisioning, please refer to the following VMware Knowledge Base article:

VMware KB: Using Thin Provisioned Disks with Virtual Machines:

http://kb.VMware.com/kb/1005418

Therefore, thin provisioning enables the efficient use of the underlying storage resources and improves the scalability of the aggregate storage capacity by over committing the storage. In a virtual desktop deployment, this approach results in a higher number of virtual desktops that can be supported with the given storage capacity. For this reason, thin provisioning of VMware's virtual disk is recommended in the Cisco Virtual Workspace (VXI) Smart Solution, though you should note that thin provisioning does have some effect on the CPU performance of the host.

In the Cisco Virtual Workspace (VXI) Smart Solution, all VMware View pools were validated with thin provisioning enabled on the parent virtual machine used to create the VMware View desktop pools. In each case, the parent virtual machine's virtual disk was residing on either EMC's Fibre Channel attached SAN storage or NetApp's NAS (Network File System [NFS]) storage.

In addition to VMware's virtual disk thin provisioning, storage vendors such as NetApp and EMC offer thin provisioning at the storage level that further improves the storage efficiency gained by VMware thin provisioning. With storage thin provisioning, the actual state of the storage allocation is hidden from VMware ESX and ESXi by the storage system.

In Cisco Virtual Workspace (VXI) Smart Solution, both virtual disk and storage thin provisioning can be deployed in a complementary fashion to optimize storage utilization. All validation with NetApp in the Cisco Virtual Workspace (VXI) Smart Solution was performed with both virtual disk and storage thin provisioning in place.

Note Since thin provisioning is an over allocation of the storage resources, you should carefully monitor the state of the thin-provisioned disk so that additional storage can be added to the data store before a lack of space causes problems.

Storage Optimization Technologies

Storage costs are typically the largest capital expense in a virtual desktop deployment and reducing these costs are critical for eliminating this initial barrier. The problem is further compounded if the storage sizing was not done correctly requiring more investment in storage to maintain good user experience. To meet the I/O demand, a common approach that large storage array vendors take is to add more physical drives to the array. However, this results in wasted capacity as the drives often exceed the storage capacity needed by the user base. To minimize cost and reduce waste, several vendors offer storage optimization technologies that significantly offloads the IOPS going to the back end storage array by

using caching technologies deployed closer to the desktops. Storage Optimization strive to reduce the IOPS performance that the storage array needs to support by serving these I/O requests from a local DRAM or SSD cache residing on the same server or centrally deployed between the servers and the back-end storage system. Vendors can offer solutions that provide read I/O offload, write I/O offload and storage capacity reduction through technologies such as de-duplication and compression. However, the solutions offered can vary so it is important to understand the specific storage problem a given vendor solution addresses. In Cisco Virtual Workspace (VXI) Smart Solution, several storage optimization solutions have been evaluated and validated to fully understand the benefits they may provide. One such solution is a hypervisor-based technology from VMware known as View Storage Accelerator (VSA). VSA leverages a caching feature in VMware ESXI 5.0 that uses memory from Cisco UCS servers for maintaining a hypervisor based cache. By using server memory as its cache, VSA improves user experience, while reducing storage costs by offloading the read IOPS. Read IOPS are typically high in a virtual desktop deployment during desktop boot up, login, application launch and antivirus scans. For a more detailed understanding of VSA and the benefits seen from the testing done in the Cisco Virtual Workspace (VXI) Smart Solution, please refer to the Cisco Virtual Workspace (VXI) Smart Solution 2.7 Performance and Capacity Results Guide for VMware.

Storage Footprint Reduction

Storage vendors support technologies that offer economies of scale for sizing the storage needs of a desktop virtualization deployment. Technologies include data de-duplication to increase storage efficiency by eliminating redundant information in the data being stored. This feature can be used for primary file systems and end-user file data in VMware and other virtualized environments. If the duplicate data is from different virtualized desktop virtual machines, the data is stored only once and the metadata associated with it is changed so that both virtual machines have access to the data. As with thin provisioning, de-duplication can provide significant storage efficiencies, improving desktop virtualization scalability since the existing storage can now support a larger number of desktop virtualization desktops. Therefore, enabling de-duplication, particularly in large desktop virtualization deployments, is highly recommended.

Please refer to EMC and NetApp documentation for more information about using de-duplication in a desktop virtualization environment.

Partition Alignment

Microsoft Windows file system partitions running on virtualized desktops should be aligned with the underlying storage partitions. This alignment can improve storage performance by reducing overall I/O latency while increasing storage throughput. This alignment is currently needed only with Microsoft Windows XP because Microsoft Windows 7 (32-bit and 64-bit) automatically provides this alignment. The problem occurs because Microsoft Windows writes 63 blocks of metadata directly at the beginning of the drive, resulting in misalignment of the first partition created on the disk. As a result, the drives may need to read an extra block of data unnecessarily, causing additional IOPS on the drive. To address the misalignment problem, an aligned partition is created on the drive that aligns with the storage system used. Both block-based and file-based storage systems can benefit from this alignment. In the Cisco Virtual Workspace (VXI) Smart Solution, a 64-KB aligned partition was created on the parent virtual machine of the desktop pools to align with EMC's SAN and NetApp's NAS storage. Please refer to Microsoft and VMware's documentation for information about implementation.

Storage Network Considerations

Jumbo frames

Virtual desktop deployments using IP-based storage should enable jumbo frames to increase storage bandwidth utilization and improve I/O response times. Jumbo frames increase the maximum transmission unit (MTU) for Ethernet frames used to transport IP traffic in data center LAN networks. Enabling jumbo frames increases the Ethernet MTU from the default value of 1518 bytes to 9000 bytes typically and should be enabled on every link between the server hosting the virtual desktops and the IP storage it uses. Jumbo frames not only improve overall throughput, but also reduce the CPU burden on the host for large file transfers.

Separation of storage network

Storage traffic should be physically (ideal) or logically separated from other network traffic using VLANs. Cisco Unified Computing System architecture supports two host bus adapters (HBAs) dedicated to storage if Fibre Channel-attached SAN storage is used.

Similar physical separation is recommended for IP-based storage, such as storage of NFS and Small Computer System over IP (iSCSI) traffic. A separate VMkernel port and VLAN should be used for IP storage traffic using a dedicated uplink port on the host. This type of physical separation of the IP storage traffic from other IP traffic is possible using the latest converged network adapters (CNAs) on the Cisco Unified Computing System, which support 128 virtual uplink ports. If an uplink cannot be dedicated, the separate VLAN used for storage will provide the logical isolation. The physical isolation should be extended into the data center network by using dedicated ports or switches at the access layer where Cisco Nexus 5000 Series Switches are typically deployed. If the storage traffic extends into the aggregation layer of the data center network, Cisco Nexus 7000 Series Switches that are typically deployed at this layer support physical separation through the virtual data center (VDC).

Port Channels

To increase the aggregate uplink bandwidth without sacrificing availability, Port Channels can be used between the LAN uplink ports on the host and the access layer switch (Cisco Nexus 5000 Series). This approach is important if the deployment uses IP-based storage since it significantly increases the LAN bandwidth required.

Multipathing

For both block-based SAN storage and IP-based NAS storage, multipathing can be used to create load-balanced but redundant paths between the host and the storage it uses. VMware learns the various physical paths associated with the storage device, and it uses a path selection scheme to determine the path a given I/O request should take. The three options on VMware for selecting the path to the storage device are fixed, most recently used, and round-robin. Round-robin should be used to load balance I/O traffic across multiple physical paths. Because of the performance improvements the multipathing provides through enhanced storage resiliency, virtual desktop deployments should enable multipathing if the storage vendors support it. Both EMC (Fibre Channel SAN) and NetApp (NFS), included in the Cisco Virtual Workspace (VXI) Smart Solution support this capability.

Validating Capacity Estimates

The next step in the overall resource planning process is to validate the capacity estimations based on factors such as CPU, memory, and storage as well as factors outlined in the previous section. On the basis of the baseline performance data from the physical desktop and the theoretical estimation for the number of virtual desktops that can be rolled out on the Cisco UCS blade server chosen for the deployment, perform performance characterization and validation in a virtualized environment to determine the following:

- Average CPU utilization of the server
- Memory utilization of the server
- Storage IOPS generated by the server
- Network bandwidth utilization of the server
- Application response times

Workload Considerations

To validate capacity estimates for a virtual desktop deployment, one option is to roll out the service to a pilot group and validate the resource estimation with actual users. Alternatively, the data regarding user activities, applications used, and use patterns collected from the physical desktops can be used to define a workload specific to that environment. For testing, the custom workload can then be automated to simulate the user workload, or it can be mapped to one of the generic profiles commonly used by workload generation tools used in scalability and performance testing.

The workload defines the applications that a person actively uses, such as word processing, presentation, and other office applications, but it can also include background activities such as backups and antivirus scans. Workloads for users within a company will vary depending on a person's job or functional role and may differ according to the organization structure (sales, marketing, manufacturing, etc.). Workloads can also vary based on the time of day, particularly if the desktop virtualization users are geographically dispersed. Background activities that begin at specific times, such as backups and antivirus scans, can also increase workloads.

Table 31 shows a very high level categorization of a user's profile based on their desktop usage patterns. Most workload generation tools will have workloads to reflect these profiles - however, other than the profile names and possibly the applications being used, the actual load they generate for a given profile will still be significantly different from one tool to another in that they may not yield the same performance results. From a capacity planning perspective, it is important to understand the details of the workload to assess how closely it matches your environment. It is probably best to assume a difference and tack on an X% performance hit when using that data for your environment - choice of X depends on how closely a given tool's workload and desktop environment used for testing matches that of yours. The preferable option still is to use feedback from your own production environment to validate the compute and storage estimations made when using performance data resulting from the use of a workload generation tool.

Table 31 User Workload Profiles

User Profile	Description
Task Worker	One application open at a time Limited printing Limited mouse usage Primarily text editing.
Knowledge Worker	Multiple applications open at a time Variety of applications Graphical applications with multimedia and use of USB peripherals.
Power User	Multiple applications open at a time Graphical and/or computational intensive applications User may need administrative rights

Cisco Knowledge Worker+ Profile

In the Cisco Virtual Workspace (VXI) Smart Solution outlined in this document, a variation of the first two user profiles was used for testing and will be referred to as the Cisco Knowledge Worker+ (KW+) profile. The details of this profile are outlined in Table 32 below.

Table 32 Cisco Knowledge Worker+ Profile

Applications in the Workload Profile	Activities in an Application
Start Cisco Jabber for Windows (version 9.1.3) and keep it running.	1. Check to see if Cisco Jabber application is installed. 2. Cisco Jabber is installed, so launch Cisco Jabber. 3. Wait for login screen and sign in the user. 4. Change presence status at a rate of 8 status changes sent per user per hour. 5. Send and receive instant message at a rate of 5 instant messages per user per hour.
Start Internet Explorer 9 and keep it running.	1. Start application. 2. Open webpage with Adobe Flash video. 3. Run the video (30 seconds long). 4. Close application.Open 3 more webpages.
Start Microsoft Word (Microsoft Office 2010) and close it.	1. Start application. 2. Open an existing file. 3. Navigate to the last page. 4. Insert a page. 5. Write a paragraph. 6. Save Microsoft Word document. 7. Close application.

Applications in the Workload Profile	Activities in an Application
Start Microsoft Outlook (Microsoft Office 2010) and close it.	1. Start application. 2. Wait until all folders are up-to-date. 3. Perform send and receive operation. 4. Clean inbox by deleting existing email. 5. Send email (4 messages sent). 6. Wait to receive email (13 messages received). 7. Delete email. 8. Read email. 9. Close application.
Start Microsoft Excel (Microsoft Office 2010) and close it.	1. Start application. 2. Open an existing file. 3. Set zoom level to 100%. 4. Page up 10 times (once). 5. Page down 10 times (once). 6. Save document. 7. Close application.
Start Microsoft PowerPoint (Microsoft Office 2010) and close it.	1. Start application. 2. Open an existing file. 3. Play slideshow. 4. Close application.

Applications in the Workload Profile	Activities in an Application
Start Adobe Acrobat and close it.	1. Start application. 2. Open an existing file. 3. Navigate to page 50. 4. Set zoom level to 75%. 5. Zoom up 5 times and zoom down 5 times. 6. Close application.

- Optimized antivirus softwarefrom a leading vendor is always running on the desktop
- The workload profile outlined above is for the latest version of the Cisco VXI KW+ workload. The latest version uses Cisco Jabber for Windows while earlier versions used Cisco Unified Personal Communicator for the Cisco collaboration application in the workload. There can be significant changes from one workload version to another so please contact the Cisco Virtual Workspace solution team for more information regarding the workload.
- All applications except for Cisco Jabber or Cisco Unified Personal Communicator in deskphone mode and Outlook are randomized across multiple iterations of the workload loop. The workload profile outlined above is for the latest version of the Cisco VXI KW+ workload. The latest version uses Cisco Jabber for Windows while earlier versions used Cisco Unified Personal Communicator for the Cisco collaboration application in the workload. There can be significant changes from one workload version to another so please contact the Cisco Virtual Workspace solution team for more information regarding the workload.
- Cisco Jabber or Cisco Unified Personal Communicator in deskphone mode, Outlook and Internet Explorer once it is launched is always running for a given iteration of the workload loop while the workload is exercising other applications.
- A random timer is used in each step above that pauses between 7 and 11 seconds to simulate user "think" time

Antivirus Considerations

Antivirus agent is installed in the golden and as the workload is executed by the simulation tool, this will trigger AV to scan on the reads and writes associated with a given activity, resulting in additional load on the compute and storage resources. For this reason, the scan policy used could change the load on the server and storage resources and impact the performance results used for capacity planning.

Validation Methodology and Results

This section provides an overview of the various scalability and performance testing done in the Cisco Virtual Workspace (VXI) Smart Solution. Testing covered here, primarily looks at the compute and storage aspects of the end-to-end system. Enterprises, in planning for their deployment, can use this info to estimate the server and storage capacity needs of their deployment - however, adjustments should be made to account for any differences in user workload or other areas of the deployment. The actual results from the testing will be part of the Cisco VIRTUAL Workspace (vxi) Smart Solution 2.7 Performance and Capacity Results Guide for VMware that will serve as an addendum to this document.

Validation Methodology

In this section we take a look at the validation methodology used for the scalability and performance testing results documented in the Cisco VIRTUAL Workspace (vxi) Smart Solution 2.7 Performance and Capacity Results Guide for VMware. All testing is done across the end-to-end Cisco Virtual Workspace (VXI) Smart Solution. For performance testing, workload generation tool from Scapa Test Technologies is used to initiate a large number of user sessions. This tool is used for all scale, performance and other characterization type testing.

Workload Profile: Cisco Knowledge Worker+

As stated earlier, the workload used is a critical factor for any performance related characterization done in a desktop virtualization environment. All the test results presented here used the Cisco Knowledge Worker (KW)+ workload unless it is stated otherwise. An overview of this profile is provided in the Workload Considerations section of this chapter.

Success Criteria

For all single server testing, the objective is to determine the virtual desktop density that can be supported on a given model of the server for the specified deployment profile based on a Cisco KW+ profile as defined above. The success criteria used in each case is as follows:

- Good User Experience based on application response times
- CPU Utilization of 80% and/or 90%
- Memory Utilization of 90% with no ballooning (ESXi), swapping but Transparent Page Sharing up to 20% is allowed

Application Response Times

Table 33 summarizes the average application response times used as the success criteria in the Cisco Virtual Workspace (VXI) Smart Solution. All testing was done using Test and Performance Platform (TPP) from Scapa Technologies. On each virtual desktop hosted on the Cisco UCS server, Scapa load generation tool will initiate a VDI session and then initiate activities defined in the workload profile to generate a workload on each desktop. Applications in the workload (except for Cisco Jabber for Windows) are launched and closed in each iteration of the workload loop. Therefore the average response times measured (shown below) for a given application is a combination of the response times measured for that application across all HVDs running on a server as well as the response times across multiple iterations of the workload running on each HVD. The success criteria was derived from a combination of testing done on physical desktops and HVD with these applications and measuring the response times. For each test, the response times measured are compared against the success criteria defined below in order for the test to pass. It is also important to note that Scapa measures the response times from an user/endpoint perspective and not from the hosted virtual desktop in the data center.

Table 33 Success Criteria

Applications	Success Criteria for Maximum Acceptable Startup Times
Cisco Jabber for Windows	5s
Outlook	5s (** see note)
Excel	5s

Applications	Success Criteria for Maximum Acceptable Startup Times
PowerPoint	5s
Acrobat	5s
Internet Explorer	5s
Word	5s (** see note)
** In some earlier tests a 10s test success criteria was used.	

Performance Metrics

The following aspects of the server performance are measured for each deployment profile tested. For ESXi, esxtop is used to measure these metrics using a 5s polling interval. Storage statistics from NetApp and EMC are included where possible.

- Average CPU Utilization
- Average Memory Utilization
- Storage
  - IOPS
  - I/O Bandwidth
  - I/O Latency
- Network Bandwidth Utilization

Summary of Results

Scalability and performance data for capacity planning based on testing done in the Cisco Virtual Workspace (VXI) Smart Solution can be found in the Cisco Virtual Workspace (VXI) Smart Solution 2.7 Performance and Capacity Results Guide for VMware. A high level summary of the deployment profiles characterized from a single server perspective is provided in Table 34.

Table 34 Results Summary - Compute & Storage

Objective	Server Model	Storage	Desktop Virtualization Profile	HVD Profile
Scalability and performance characterization of Cisco UCS B200M3 server with VMware View (Vblock)	Cisco UCS B200 M3 with 384G of memory	VSPEX (EMC VNX 5500) - Fibre Channel	VMware View 5.1 on VMware ESXi5.0U1	Microsoft Windows 7 32-bit with 2 GB of memory and 20 GB disk; Persistent
Scalability and performance characterization of Cisco UCS B230M2 server (Vblock)	Cisco UCS B230 M2 with 256G of memory	VSPEX (EMC VNX 5500) - Fibre Channel	VMware View 5.0 on VMware ESXi 5.0	Microsoft Windows 7 32b with 1.5G of memory and 20G disk; Persistent

Objective	Server Model	Storage	Desktop Virtualization Profile	HVD Profile
Scalability and performance characterization of Cisco UCS B250M2 server (Vblock)	Cisco UCS B250 M2 with 192G of memory	VSPEX (EMC VNX 5500) - Fibre Channel	VMware View 4.6 on VMware ESXi 4.1	Microsoft Windows 7 32b with 1.5G of memory and 20G disk; Persistent
Impact of CPU utilization counter change and vSphere 5.0 changes on scale and performance of Cisco UCS B250 M2 (Vblock)	Cisco UCS B-250 M2 with 192G of memory	VSPEX (EMC VNX 5500) - Fibre Channel	VMware View 5.0 on VMware ESXi 5.0	Microsoft Windows 7 32b with 1.5G of memory and 20G disk; Persistent
Storage Optimization with VMware's View Storage Accelerator	Cisco UCS B200 M3 with 384G of memory	VSPEX (EMC VNX 5500) - Fibre Channel	VMware View 5.1 on VMware ESXi5.0U1	Microsoft Windows 7 32b with 1.5G of memory and 20G disk; Persistent

Application Characterization

Goal of application characterization is to characterize the performance of the application as a standalone application in the workload, running across all user desktops deployed on a server. An Enterprise looking to deploy a Cisco Collaboration application for their virtual desktop users can use the performance data to understand the incremental impact of that application as more and more users start using them concurrently. The impact is to the server resources, which can change the number of users that can be deployed on that same server once the new application is rolled out. Table 35 shows a summary of the different applications characterization efforts done in the Cisco Virtual Workspace (VXI) Smart Solution.

Table 35 Results Summary - Applications

Objective	Server Model	Storage	Desktop Virtualization Profile	HVD Profile
Scale and Performance characterization of Cisco Jabber for Windows with VMware View	Cisco UCS B200 M3 with 384 GB of memory	VSPEX (EMC VNX 5500)	VMware View 5.1 on ESXi 5.1	Microsoft Windows 7 32-bit with 2 GB of memory
Scale and Performance characterization of Cisco Contact Center - CTIOS Agent	Cisco UCS B230 M2 with 256 GB of memory	NFS on NetApp FAS 3170	N/A - See test profile for more detail.	Microsoft Windows 7 32-bit with 2 GB of memory

Network Services

There are a number of network services that are enabled in the Cisco Virtual Workspace (VXI) Smart Solution that provide critical functionality and optimizations needed for a virtual desktop deployment. These network services are as follows:

- Cisco Nexus 1000V virtual switch or Cisco Nexus 1010 Appliance - access layer switch for virtual desktops in data center and gateway to network services using vPath
- Virtual Security Gateway (VSG) - enables access layer security for virtual desktops in data center
- Application Control Engine (Cisco ACE) - enables load balancing and SSL Offloading of connection setups to connection broker
- Adaptive Security Appliance (Cisco ASA) appliance, blade, virtualized - data center aggregation layer security
- Wide Area Application Services (Cisco WAAS) - Reduces BW consumption and improve user experience for users in branch sites through WAN optimization

Cisco Nexus 1000V

Cisco Virtual Workspace (VXI) Smart Solution leverages Cisco Nexus 1000V to provide a number of key services while functioning as the access layer switch for virtual desktop, infrastructure (DV, Collaboration & Productivity Applications, Directory and Network Services) and management VMs needed for the deployment. A single Cisco Nexus 1000V can be used, particularly in small deployments, for both user and non-user VMs. However, as you get to larger, 500+ user deployments, using separate Cisco Nexus 1000V switches is recommended for a number of reasons as outlined below.

- Infrastructure and Management VMs are likely to be deployed on separate servers from that of user desktops, for optimal use of server resources and due to differences in server model selected for user vs. non-user VMs. If the design calls for these servers to be in different data centers due to administrative policies or other organizational needs, separate virtual switches would be required anyway since a Cisco Nexus 1000V cannot span data centers. Also, since Cisco Nexus 1000V is licensed on a physical CPU socket basis, there are no licensing advantages in using same or different Cisco Nexus 1000V when user desktops and non-user desktops are on different servers.
- User desktops are also likely to have similar network, security and monitoring policies that would have to be defined through similar port profiles on every switch with users on it. They may also have common administrative, operational and high availability policies within the organization. Grouping the users and dedicating one or more Cisco Nexus 1000V switches for user desktops can minimize configuration changes and reduce the overall administrative and operational burden.
- Desktop virtualization, by definition, moves the user's desktops into the data center where critical infrastructure and other server VMs reside. To minimize the impact of a potential security threat, it may be best to provide as much isolation as possible between these server and user VMs.
- Lastly, infrastructure VMs are likely to have multiple layers of security and isolation that are broadly applied to both virtual desktop users in the data center and physical desktop users outside the data center, which can make any benefit of co-locating them on the same virtual Ethernet module or switch, non-existent.

When deploying virtual desktops on a Cisco Nexus 1000V, a number of features are recommended. The scalability of these features, in addition to the platform itself, is important to consider from a capacity planning perspective. A high level summary of the relevant features is shown in the table below.

Table 36 Cisco Nexus 1000V features in Cisco Virtual Workspace (VXI) Smart Solution

Cisco Nexus 1000V –OR-Cisco Nexus 1010	Features Most Relevant to Cisco Virtual Workspace (VXI) Smart Solution
Security Features	DHCP Snooping, Dynamic ARP Inspection, IP Source Guard, Port Security, ACL, Virtual Service Domains Using vPath - VSG
Network Features	VLAN segmentation, 802.1Q tagging, QoS marking (COS & DSCP)Using vPath or WCCP to Cisco WAAS
HA features	HA for virtual desktops and/or associated infrastructure VMs through network VMotion that enables mobility of security and network policies and maintenance of connection state when vMotion occurs
Troubleshooting	Enables network features and visibility to troubleshoot virtual machine traffic through features such as per-virtual machine interface stats, SPAN/ERSPAN, CDP, NetFlow

Above features are same for Cisco Nexus 1010, the appliance version of Cisco Nexus 1000V. For a more comprehensive discussion on the services and functions that a Cisco Nexus 1000V or Cisco Nexus 1010V provides in the Cisco Virtual Workspace (VXI) Smart Solution, please refer to the Securing Cisco Virtual Workspace chapter of this document.

To scale a virtual desktop deployment using Cisco Nexus 1000V or Cisco Nexus 1010 appliance for user desktops, both the overall scalability of the platform and that of the features recommended must be well understood. The next three tables summarizes the Cisco Nexus 1000V's product and feature level scalability limits based on the most recent software version used in system level validation of the Cisco Virtual Workspace (VXI) Smart Solution.

Table 37 Cisco Nexus 1000V Platform Scale Limits

Cisco Nexus 1000V	Limits	Notes
Number of virtual switch chassis per vCenter	Multiple	
Can Cisco Nexus 1000V span multiple data centers?	Yes	It can span multiple physical data centers but not virtual data centers within vcenter
Number of Virtual Ethernet Modules (VEM) in a Virtual Switch Chassis	64	VEM is the Ethernet module of a switch and runs on each Cisco UCS server where the virtual machines reside; 1 VEM per host or Cisco UCS server
Number of Virtual Services Modules (VSM) in a Virtual Switch Chassis	2	Second VSM is for redundancy

Cisco Nexus 1000V	Limits	Notes
Number of virtual modules in a Virtual Switch Chassis	66	64 VEM + 2 VSM; 64 single-width or dual-width Cisco UCS servers can be connected to a VSM, forming a single 1000V virtual switch chassis; Servers can be connected in L2 or L3 mode
Number of vEthernet ports per Virtual Ethernet Module	216	Each Cisco UCS server can have 200 VMs connected via vEthernet ports to a single VEM in the virtual switch chassis; For a Knowledge Worker workload, this is still well above the max density that can be supported by Cisco UCS servers today, including Cisco UCS B200M3. For lighter workloads, this could be a limiting factor
Number of vEthernet ports per 1000V virtual switch chassis	2000	
Number of Port Profiles per Virtual Switch Chassis	2000	
Number of vEthernet ports per port profile	1000	
Number of MACs per VLAN per VEM	4000	16,000 per VE

Table 38 Cisco Nexus 1010 Platform Scale Limits

Cisco Nexus 1010 Appliance	Notes
Scales up to 4 VSM and 256 VEM	64 VEM per VSM or virtual switch chassis

Table 39 Scalability of Features on Cisco Nexus 1000V for a virtual desktop deployment

Cisco Nexus 1000V or Cisco Nexus 1010 Appliance	Limit
Port Security per host/virtual switch chassis	216/2000
ACL interfaces per host/virtual switch chassis	256/2048
QoS interfaces per host/virtual switch chassis	256/2048
vSD interfaces per host/virtual switch chassis	214/2048
VSD per host/virtual switch chassis	6/64
VLANs per virtual switch chassis	2048
SPAN/ERSPAN	64
NetFlow	256
CDP - enabled globally by default so it applies to all ports on the virtual switch chassis	N/A

Note Please refer to the product documentation on cisco.com for additional info and for the most up to date scalability and performance numbers.

High Availability

Cisco Nexus 1000V should be deployed 2 VSMs where possible to provide high availability from a switch supervisor perspective. However even if a VSM fails, the Cisco Nexus 1000V's virtual ethernet modules will continue to forward traffic though no configuration changes can be made. Redundant VSMs should be deployed on separate hosts. If using Cisco Nexus 1010 appliance, it should be deployed as a pair to provide redundancy.

Cisco Virtual Security Gateway (VSG)

With desktop virtualization and deployment of user desktops in the same Enterprise data center as virtualized servers, protecting the virtualized environment and enforcing access layer firewalling of user desktops are a critical requirement for Enterprises. Cisco's Virtual Security Gateway (VSG) was designed to address this need and is key component in the overall security architecture. A more detailed discussion of VSG can be found in the Securing Cisco Virtual Workspace (VXI) Smart Solution chapter of this document.

Two key functions provided by VSG in a virtual desktop deployment include:

- Secure segmentation of user desktops and data into zones
- Securing per-user-desktop virtual machine traffic and providing inter-virtual machine firewalling in a scalable manner

VSG is tightly coupled with Cisco Nexus 1000V to provide performance accelerated network security services. Cisco Nexus 1000V, through its vPath feature, provides intelligent traffic steering that redirects traffic from user desktop VMs to a VSG services virtual machine that makes a policy determination based on predefined policies. Cisco Nexus 1000V learns and caches the policy decisions from the initial flows of a given user virtual machine. Policy enforcement for all subsequent flows are now done at Cisco Nexus 1000V based on the cached policies, thereby offloading VSG from any further policy enforcement. This feature greatly enhances the scalability of secure virtual desktop solution and should be considered when doing capacity planning for a deployment.

When planning for a virtual desktop deployment with Cisco Nexus 1000V and VSG, scale limits for the number of VSGs and Cisco Nexus 1000Vs that can be managed together using Cisco VNMC, number of zones per VSG, number of rules and policies needed per zone and number of user desktop VMs that can be supported should be well understood for the software releases being deployed. Performance limits in terms of the number of connection/second and maximum number of concurrent connections needed should also be well understood for the release being deployed. For the software releases validated in the Cisco Virtual Workspace (VXI) Smart Solution, Table 40 highlights the scalability and performance limits of VSG.

Table 40 Scalability and Performance Limits of VSG

Metric	Limit
Number of VNMCs per vCenter	2
Maximum Number of Cisco Nexus 1000V VSMs supported per VNMC	4

Metric	Limit
Maximum Number of VSGs per VNMC	128
Number of Cisco Nexus 1000V VEMs supported per VSG/VNMC	12/12
Number of Zones supported per VSG/VNMC	32/4096
Number of VMs supported per VSG/VNMC	300/800-1000
Number of policy rules supported per VSG/VNMC	1024/8192
Maximum Number of Concurrent Connections per VSG	256k
Maximum Number of Connections per second per VSG	4096
Maximum Number of connections supported per VSG	256

Note Please refer to the product documentation on cisco.com for additional info and for the most up to date scalability and performance numbers.

Additional virtualized firewall features, available natively on Cisco Nexus 1000V, are also needed to secure desktop virtual machine traffic - scalability limits for these are provided in the Cisco Nexus 1000V section above. VSG Deployment guides available on cisco.com with each software release are also a very good source of information for product design and deployment details.

High Availability

Cisco VSG can be deployed in a highly redundant manner in an Active/Standby configuration. Standby VSG stays in sync with Active VSG to provide stateful failover that minimizes traffic disruption in the event of a failure. To maximize redundancy, Active and Standby VSGs should be deployed on separate server or hosts.

Cisco Application Control Engine (ACE)

Enterprises require availability and performance from their applications and with the drive towards virtualization and data center consolidation, they also want to reduce the number of servers for a given application or service by eliminating idle server resources and maximizing the performance of the remaining servers. This must be done in a secure manner that protects the data center and the applications being delivered. Cisco ACE enables an Enterprise to meet these objectives while accelerating data center consolidation efforts and delivery of new services, while reducing operational costs. Cisco ACE is an integral part of the Cisco Virtual Workspace (VXI) Smart Solution that provides the following network services for a virtual desktop deployment:

- Load balancing of traffic to virtual desktop connection brokers
- Offloading of SSL termination from Connection brokers

Cisco ACE, on the same physical platform, can be virtualized to provide similar services to applications other than desktop virtualization thereby enabling the Enterprise to further scale their data center investment and achieve greater operational efficiencies.

Load balancing service provided by Cisco ACE in Cisco Virtual Workspace (VXI) Smart Solution is a network layer function, and not an application level feature that some application vendors provide. Cisco ACE will load balance traffic to as many connection brokers as needed for the deployment. If one connection broker fails, traffic will be distributed across the remaining connection brokers. This ensures the availability of the connection brokers while maximizing the available resources. It is important to note that traffic destined to the brokers are during session establishment in most deployments. Once the session is up, the broker is no longer in the communication path; the user's client device talks directly with the IP of the virtual desktop. See Virtualization Aware Network chapter for additional info on deploying Cisco ACE in a Cisco Virtual Workspace (VXI) Smart Solution.

SSL Offloading is another way to achieve higher levels of scale from the connection brokers by offloading the SSL termination/origination function to Cisco ACE. All virtual desktop session traffic is encrypted and in this case, the Cisco ACE will offload only the launch of session traffic but not once the session is established.

When planning for a scalable virtual desktop deployment, Cisco ACE should be a fundamental part of that design. Cisco ACE is available both as an appliance and as a module for the catalyst 6500. Scalability and Performance Limits of the Cisco ACE module and Appliance that are relevant for a virtual desktop deployment are provided in Table 41 below.

Table 41 Scalability and Performance Limits of Cisco ACE

	Cisco ACE Module	Cisco ACE Appliance	Notes
Load balancing: Number of concurrent L4 connections (Unproxied)	4,000,000	1,000,000	
Load balancing: Number of concurrent L7 connections (Proxied)	512,000	128,000	
SSL Offloading: Number of concurrent SSL connections	100,000	100,000	Subset of L7 proxied connections
Number of virtual contexts	251	21	1 Admin context and remaining User contexts
VLANs	4000	4000	

Note Please refer to the product documentation on cisco.com for additional info and for the most up to date scalability and performance numbers.

Cisco Adaptive Security Appliance (ASA)

Cisco's Adaptive Security Appliance (ASA), discussed in depth in the Securing Cisco Virtual Workspace (VXI) Smart Solution chapter of this document, is a fundamental component of Cisco's Secure Borderless Network Architecture and an integral part of the Cisco Virtual Workspace (VXI) Smart Solution. Cisco ASA provides a number of key security functions that are fundamental to most desktop virtualization deployments. These functions can be summarized as Firewall and Security Gateway Services and address the following security use cases for desktop virtualization.

- **Firewall**:

  - Securing all traffic entering the Enterprise data center. This includes users accessing their virtual desktops or applications hosted in the data center, using a range of access devices and from various locations.
  - Segregation of hosted virtual desktop traffic from other critical application, management & infrastructure services residing in the same data center.

- **Security Gateway**:
  - Termination of Site-to-Site (IPsec/SSL) VPN sessions or Cisco AnyConnect/Clientless VPN user sessions from teleworker and mobile users.

In the next sections, we take a closer look at some factors to consider in capacity planning for a virtual desktop deployment with Cisco ASA.

Cisco ASA series comes in a wide range of form factors, performance levels, security services and provide a comprehensive security solution for the Enterprise, all managed through a single management interface (Cisco Security Manager). In the Cisco Virtual Workspace (VXI) Smart Solution, Cisco ASA is recommended in the Enterprise data center as a high performance and scalable firewall solution to meet the needs of a virtual desktop deployment. Cisco ASA is also needed to terminate VPN connections at the Internet edge, located in the data center or another location depending on the Enterprise design. If Cisco ASA is already deployed in the data center to provide security services for non-virtual desktop traffic, capacity planning for desktop virtualization should consider leveraging the same Cisco ASA for virtual desktop traffic. In the Enterprise data center, Cisco recommends deploying a Cisco ASA 5585-X as it is specifically designed to meet the scalability and performance needs of today's growing Enterprise data centers. Table 42 below shows the scalability and performance limits of this Cisco ASA for the security gateway and firewall services needed in Cisco Virtual Workspace (VXI) Smart Solution. For more details on this model, please refer to this document:

http://www.cisco.com/en/US/prod/collateral/vpndevc/ps6032/ps6094/ps6120/design_guide_c22-624431.html

Table 42 Scalability and Performance Limits of Cisco ASA 5585-X

Cisco ASA 5585-X with SSP 60	Limit
Maximum Firewall Throughput	20 Gbps
Maximum Firewall Connections	10 million
Maximum Firewall Connections/Second	350,000
Maximum Site-to-Site VPN sessions	10,000
Maximum Cisco AnyConnect or Clientless VPN User Sessions	10,000
Security Contexts	250
VLANS	1000

Note Please refer to the product documentation on cisco.com for similar info on other Cisco ASA models and for the most up to date scalability and performance numbers.

A comparison of the scalability and performance limits of all Cisco ASA models can be found here:

http://www.cisco.com/en/US/products/ps6120/prod_models_comparison.html#%7Etab-c

Dependingon the network design, Firewall and Security Gateway Services functions can be combined on a single platform with Cisco ASA operating in both roles. Cisco ASA's role can also be expanded to include other advanced security services such as IPS (intrusion Protection System), Unified Communication Security Services and Content Security services that meet multiple security requirements beyond that of desktop virtualization. Further, consolidating multiple security services on a Cisco ASA lowers the overall costs for the Enterprise by reducing management and operational complexity. Enterprise can also reduce hardware costs that would otherwise be required if services were deployed on different hardware platforms. Combining services can be done in a highly scalable manner on Cisco ASA and should be factored into your capacity planning and selection of security platform for desktop virtualization.

Power consumption and rack space requirements are an important consideration for scaling today's data center networks. Cisco ASA 5585-X can meet the scalability and performance needs of Enterprise data centers in a 2 RU platform using only ~800W of power.

In summary, Cisco ASA provides the following benefits that should be factored into your capacity planning effort for a scalable and secure virtual desktop deployment hosted in data center.

- Cisco ASA can provide the security services needed for a virtual desktop deployment but it can also meet the larger security needs of the overall Enterprise Data Center
- Cisco ASA can provide multiple security services on a single platform or multiple platforms as needed, all managed through a single management interface
- Cisco ASA meets the scalability and performance of today's data center environments in a 2 RU platform and using only ~800W of power
- Cisco ASA meets the scalability and performance of today's data center environments with a Firewall performance of 10million connections and 350,000 connections/second and a VPN performance of 10,000 Site-to-Site (IPsec/SSL), Cisco AnyConnect or Client VPN user sessions
- Consolidation of multiple security services without incurring additional hardware costs
- Consolidation of security services that simplifies DC and management complexity while reducing operational costs

High Availability

All models of Cisco ASA except for one low-end platform, support high availability in the form of Active/Standby or Active/Active failure configurations. The latter enables load sharing among the redundant devices but is only available in multiple context mode while Active/Standby configuration can be used with both single and multiple context modes. Cisco ASA provides sub-second and stateful/stateless failovers in either configuration to minimize impact to user sessions.

Cisco ASA also provides an interface redundancy feature whereby a link failure results in an interface-level failover on the same Cisco ASA rather than a device failover to the redundant Cisco ASA. This feature will minimize any service interruptions that occur with device failover. Most inspection engine states are preserved with an interface-level failover versus a device failover. This along with other benefits is highly recommended in virtual desktop deployments.

Cisco Wide Area Application Services (WAAS)

Desktop virtualization, by definition, moves a user's desktop from the user's physical location to a virtual desktop hosted in the Enterprise data center network. In order to access the desktop, the user needs network connectivity from his physical location (Campus, Branch, Internet) to the Enterprise data center hosting the user's desktop. Instead of a local desktop, the user now uses a keyboard, mouse and display attached to a client device/endpoint(zero, thin or thick) to access the desktop and use it as one would a physical desktop - but across the Enterprise network. This requires additional network bandwidth that

needs to be planned for from a capacity planning perspective as applications (e.g. Microsoft Word) that used to run locally on the physical desktop without the need for any network bandwidth, now requires bandwidth to the use the same application. The user 's interactions and experience with that application that was once a matter of local desktop resources, can be noticeably impacted if there is network congestion, packet loss or other network impairments. The display protocols used for transporting the user's keyboard and mouse interactions to the virtual desktop and the remote desktop view in the reverse direction to the user's display, are well optimized to minimize the bandwidth consumption. However, there is still a need for further bandwidth optimization, particularly in today's environment, where an ever-increasing demand for multiple types of video and other collaboration applications are becoming a standard requirement including virtual desktop deployments. Maintaining user experience in this environment, particularly for a branch user while minimizing bandwidth requirements, can be challenging.

In branch office deployments, Cisco WAAS can provide network optimization services that greatly reduce the bandwidth requirements, while improving user experience. For a scalable branch deployment, Cisco recommends deploying Cisco WAAS at the Enterprise edge (Branch, Campus/DC Edge). For more details on the design and deployment of Cisco WAAS in a virtual desktop deployment, please refer to the Virtualization Aware Network chapter of this document.

From a capacity planning perspective, the overall size and scope of the desktop virtualization deployment in the branches will determine the scalability requirements for Cisco WAAS. Information needs to include the total number of virtual desktops across the WAN, the average bandwidth for the virtual desktop sessions, number of branch sites and the deployment size of each branch. This data will determine the selection of Cisco WAAS platform, both at the Enterprise head-end and at the branch side. Cisco WAAS comes in a number of form factors that include software and hardware based (Cisco SRE Modules) add-ons to Cisco WAN ISR routers and as physical (WAE) and virtual (Cisco vWAAS) appliances. For a more comprehensive list of platform options and capabilities, see the following cisco.com document:

http://www.cisco.com/en/US/products/ps5680/Products_Sub_Category_Home.html#%7Eall-prod

The deployment data can be used to size the Cisco WAAS platforms needed for the different branch sites. The tables below show the scalability and performance for different Cisco WAAS configurations that can be used in a virtual desktop deployment. Note that Cisco WAAS performance is rated based on 2 metrics: number of TCP connections and WAN BW that it can optimize. The number of TCP connections can be estimated from the number of users and connections per session required for a given number of concurrent virtual desktop sessions. Similarly the BW needs of the deployment can be estimated from the number of users at the branch and their per-session BW needs. To ensure a successful deployment, both metrics need to be evaluated to stay within the performance limits of that platform configuration. For additional information on capacity planning for a desktop virtualization deployment, see Network - WAN Capacity Planning section later in this chapter.

Table 43 Scalability and Performance Limits of Cisco vWAAS

Model	Optimized TCP Connections	WAN BW	Virtual Cores	Memory	Hard Disk
Cisco vWAAS-750	750	8 Mbps	2 vCPU	4 GB	250GB
Cisco vWAAS-6000	6000	90 Mbps	4 vCPU	8 GB	500 GB
Cisco vWAAS-12000	12000	310 Mbps	4 vCPU	12 GB	750 GB

Table 44 Scalability and Performance Limits of Cisco WAAS Appliances

Model	Optimized TCP Connections	Drive	RAID	Memory	WAN BW
WAE-674-8GB	6000	600 GB	RAID 5	8 GB	90 Mbps
WAE-7341	12000	900 GB	RAID 5	8 GB	310 Mbps
WAE-7371	50000	1500 GB	RAID 5	24 GB	1 Gbp

Note Please refer to the product documentation on cisco.com for similar info on other Cisco WAAS models and for the most up to date scalability and performance numbers.

Network - WAN Capacity Planning

This section looks at the enterprise WAN and the factors that need to be considered to determine the number of desktop virtualization users that can be deployed at a branch site with a finite amount of bandwidth. In simple terms, the number of desktop virtualization users that a WAN link of a given bandwidth can support is:

Bandwidth of WAN link /Bandwidth of single desktop virtualization session

However, in desktop virtualization environments, this type of sizing is a challenge due to a number of variables that can affect the amount of bandwidth in a single desktop virtualization session. In voice over IP (VoIP) environment, in which a voice call is a voice call—that is, there is a predictable, smooth, and a constant 80 kbps per call for a given codec—desktop virtualization session traffic is a complete variable, as a user's network traffic is today. In a voice environment, if a branch location needs to support

10 simultaneous calls, 10 x 80 kbps = 800 kbps of traffic is needed

For desktop virtualization, the session traffic flow can vary widely between two different users, based on what they are doing on their desktops—this is just one of several variables that can vary the bandwidth requirements for a session. Also, every action that a user initiates (mouse clicks, keyboard actions) and the response to that action is transported across the network and is part of the session traffic between the virtual desktop hosted in the data center and the client that the end user is using. Unlike physical desktops where many applications can be launched and used without network access (e.g. editing a local copy of a Microsoft Word document), using a virtual desktop implies that all activities on that desktop require network access and therefore network bandwidth. For a given virtual desktop session, there is a main display session that is either TCP (RDP) or UDP (PCoIP) based for transporting the desktop display to the end user. There could also be additional TCP sessions associated with the main display session to provide features and experience equivalent to that of a physical desktop such as multimedia redirection [MMR] and USB redirection traffic.

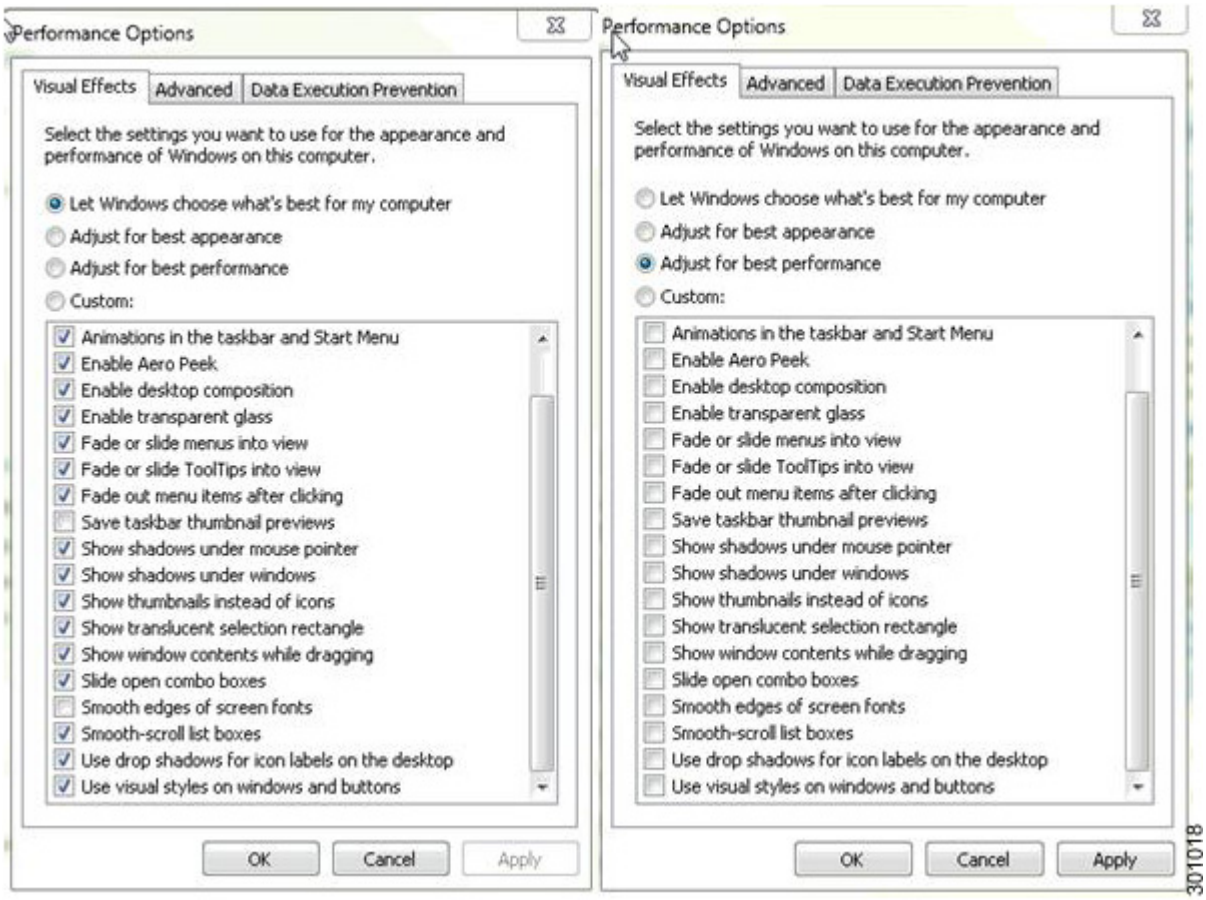
- Transport protocol used
- Type of encryption used
- Level of compression it provides
- Display rendering - local or remote
- How the protocol adapts to changes in network conditions (Available Bandwidth, Latency, Jitter)

These aspects of the protocol behavior can have a direct bearing on your WAN capacity planning in a number of ways.

In addition to the display protocol itself, there are a number of other factors that can impact network bandwidth and therefore WAN sizing as outlined below

- Number of monitors, screen resolution and color depth can all have an impact on bandwidth.
- Applications provided on the desktop (Microsoft Office, Softphone, Web browser, Instant Messaging) and how the user uses these applications as well as other activities on the desktops greatly vary the per desktop or per user load on the network and this variability can have a significant impact that makes network sizing a bit challenging. For this reason, any theoretical exercise using data from Cisco or other vendors should be validated by monitoring your environment and calibrating the sizing data through pilots or other testing that takes into account traffic patterns, use cases and other parameters that may be unique to your Enterprise.
- The guest OS used on the virtual desktops, namely Microsoft Windows XP or Windows 7 (32-bit and 64-bit) can also have an impact on the network bandwidth needs due to differences in the desktop experience they offer or other OS specific changes that impact the session behavior or experience. One such example is the appearance and performance setting on Microsoft Windows 7 desktops - see Figure 49 Based on the testing done in the Cisco Virtual Workspace (VXI) Smart Solution, if the default option of `Let Windows choose what's best for my computer' (default) is changed to Best Performance, a bandwidth savings of ~30% is possible. Since this change impacts GUI and other nice to have graphics features such as transparency without any impact to functionality, it is worth considering for VDI deployments across the WAN. It should also reduce the CPU processing needs of a desktop which in turn can increase the number of desktops that can be hosted on a given Cisco UCS server. For this reason, all validation in the Cisco Virtual Workspace (VXI) Smart Solution is with Windows Best Performance enabled. In short, due to video content being very bandwidth intensive, with stringent loss, jitter, and latency requirements, QoS and the bandwidth allocation for video requires careful consideration and planning with the use case and associated traffic patterns well understood.

Figure 49 Windows Appearance and Performance Setting



Video

Another important consideration is whether streaming video needs to be supported in virtual desktop. If so there are a number of considerations that impact capacity planning as well the overall user experience. Video in general is very bandwidth intensive with stringent loss, jitter and latency requirements and so video deployment in branch environments require very careful consideration and this challenge exists regardless of desktop virtualization. In VMware View deployments, the video can be transported using Multimedia Redirection (MMR) which is a separate TCP session from that of the main display session which could be UDP based (PCoIP) or TCP based (RDP). This provides the administrator with the ability to provide not only provide QoS that has a direct impact on the overall user experience at the branch but also allows WAN optimization to be used to reduce the bandwidth requirements across the WAN. However there are some caveats that are worth mentioning. First, MMR cannot be used to transport flash video. Second Cisco WAAS does not optimize flash video as of now and lastly, VMware View does not have support for MMR when the guest OS running on the HVD is Microsoft Windows 7.

Printing

Printing in a desktop virtualization environment can also affect the per-session traffic and therefore WAN capacity needed to support it depending on the type of printing solution deployed. Enterprises can deploy the print server in their data center and have the print traffic traverse the WAN network to a local printer at the branch site either USB attached or otherwise. The USB print traffic is transported on a

separate channel from that of the main display session in VMware environments, which helps in providing network level QoS that can be important for overall user experience. From a WAN capacity planning perspective, WAN optimization can provide significant to reduce the bandwidth needs associated with printing. There are two deployment options that can be used with Cisco WAAS in a desktop virtualization environment - one is to use the Cisco WAAS Print Application Optimizer (PAO) feature (available as of Cisco WAAS 4.1) to accelerate Microsoft Windows printing and the second is to deploy a print server on the branch router on the Cisco WAAS appliance itself. In the first case, PAO optimizes the Windows printing protocol CIFS/MSRPC by removing the chattiness of the protocol across the WAN through metadata caching, delayed closing of printer handles, asynchronous handling of print data, and so on- in addition to the transport level optimization that is fundamentally provides for TCP based protocols. In the second case, by deploying the Cisco WAAS as a print server, the spooled print traffic does not have to traverse the WAN network. See the Virtualization Aware Network chapter for more details on the deployment options for printing in a desktop virtualization environment, including the benefits that Cisco WAAS can bring to your deployment.

WAN Optimization

Another important consideration when sizing WAN links is whether to use WAN optimization. Cisco's Cisco WAAS is an application performance optimization solution designed to reduce bandwidth requirements, thereby enabling a larger number of users to be deployed across a given WAN link. It also reduces latency by locally caching the application traffic, which serves to improve user experience. In desktop virtualization deployments, there are two key benefits in deploying Cisco WAAS as outlined below:

- The inherent nature of desktop virtualization where the desktop events are remotely displayed across the network results in a significant increase in the network load that would not exist with physical desktops and the costs associated with upgrading the WAN links to migrate branch users from physical desktops to virtual desktops could be significantly reduced by deploying WAN optimization. Therefore, from a capacity planning and a TCO perspective, this is highly recommended. Various models of Cisco WAAS are available and should help in closely matching it to the needs of your environment. Additionally, Cisco WAAS can be used to optimize bandwidth for non desktop virtualization traffic, which can be particularly important for branches with both physical and virtual desktops users. Please refer to the Virtualization Aware Network chapter of this document for more information on deploying Cisco WAAS in desktop virtualization environments.
- Though less critical to WAN sizing, Cisco WAAS also has the benefit of reducing the latency by locally caching some of the application traffic. Also, by optimizing the application traffic and reducing the bandwidth needs, it reduces the likelihood of congestion related performance issues that can impact user experience.

Note Cisco WAAS, currently can optimize TCP based traffic, be it display protocol traffic or traffic outside the display protocol. It cannot optimize UDP based PCoIP traffic.

Estimating Network Bandwidth

Based on the discussion so far, it should be clear that there are number of variables at play with desktop virtualization that can impact the per-session bandwidth utilization and therefore network capacity planning for the overall branch deployment. Since the applications, workloads and use patterns can all vary the bandwidth requirements, the sizing of the WAN link is not a trivial effort, and any sizing estimations must be validated in the customer environment.

Due to the highly variable nature of desktop virtualization traffic, there are many factors that can impact the per-session bandwidth utilization and therefore network capacity for the overall branch deployment in an Enterprise. One approach to determining the bandwidth requirements for a WAN link is to characterize the bandwidth needs using a load generation tool and a generic workload profile (Task Worker, Knowledge Worker) that best matches your environment. See Workload Considerations section earlier in this chapter for more info on the generic profiles and the workload used for network characterization in the Cisco Virtual Workspace (VXI) Smart Solution. However, any estimation based on generic workloads should be adjusted to account for any variability in your environment or one should calibrate the workload used for estimation using a workload that is more representative of your environment. If Cisco WAAS is used for WAN optimization, some testing in a pilot environment should be considered due to the variability of the traffic usage pattern in your environment that would be difficult to simulate using test tools.

To aid in the sizing process, validation was done in the Cisco Virtual Workspace (VXI) Smart Solution to determine bandwidth sizing data with a detailed per-application analysis of the peak and average utilization used by the more common applications in a Knowledge Worker workload. These and other network characterization testing done in the Cisco Virtual Workspace (VXI) Smart Solution are summarized in Table 45 and can be starting point for guiding the sizing of your WAN network for desktop virtualization.

Network Characterization Results

Data to guide your WAN capacity planning based on testing done in the Cisco Virtual Workspace (VXI) Smart Solution can be found in the Cisco VIRTUAL Workspace (vxi) Smart Solution 2.7 Performance and Capacity Results Guide for VMware. A high level summary of the areas covered in the Cisco VIRTUAL Workspace (vxi) Smart Solution 2.7 Performance and Capacity Results Guide for VMware are provided in the Table 50 below.

All of the testing was done from branch sites across an end-to-end Cisco network based on Cisco Virtual Workspace (VXI) Smart Solution architecture outlined in earlier chapters of this document. As stated earlier, the data presented here is based on a Cisco Knowledge Worker+ workload - however it provides valuable insight into the factors that an Enterprise will need to consider for WAN deployments with limited bandwidth. For a complete description of the Cisco KW+ workload, please refer to the Workload Considerations section earlier in this chapter.

Table 45 Results Summary - Network

Objective	Server Model	Wan Link	DV Profile	HVD Profile
Understanding the bandwidth (BW) characteristics of a Cisco KW+ workload	Cisco UCS B200 M2 with 96G memory	T1 with 80ms Latency	VMware View 4.5 on ESXi 4.1; PCoIP and RDP	Microsoft Windows 7 32b with 1.5G of memory and 20G disk; Persistent
Understanding the bandwidth characteristics of a video-only workload	Cisco UCS B200 M2 with 96G memory	T1 with 80ms Latency	VMware View 4.5 on ESXi 4.1; PCoIP and RDP	Microsoft Windows 7 32b with 1.5G of memory and 20G disk; Persistent

Objective	Server Model	Wan Link	DV Profile	HVD Profile
Impact of display protocol adaptiveness on server/compute performance at scale	Cisco UCS B200 M2 with 96G memory	T1 with 80ms Latency	VMware View 4.5 on ESXi 4.1; PCoIP and RDP	Microsoft Windows 7 32b with 1.5G of memory and 20G disk; Persistent
Impact of Cisco WAAS Optimization on WAN deployments with View RDP	Cisco UCS B200 M2 with 96G memory	T1 with 80ms Latency	VMware View 4.5 on ESXi 4.1; PCoIP and RDP	Microsoft Windows 7 32b with 1.5G of memory and 20G disk; Persistent

Key Takeaways

In summary, some key takeaways based on the network characterization testing done in the Cisco Virtual Workspace (VXI) Smart Solution are as follows:

- Minimum bandwidth required for PCoIP and RDP with the specified workload is 320kbps and 1.28Mbps respectively. The peak bandwidth consumed by the same workload is 3.6Mbps for PCoIP and its greater than 2Mbps for RDP. This data can be used in sizing WAN links and for enabling QoS polices on these links.
- Certain functions or features within an application may cause peak bandwidth consumption though the application as a whole may not consume as much. For example, slide show mode in PowerPoint has the highest BW impact in the specified workload.
- Rich media application used in the Cisco KW+ workload, namely Cisco Unified Personal Communicator 8.5 in deskphone mode does not have a significant BW impact however PowerPoint and Outlook are the biggest bandwidth consumers in the specified workload.
- Cisco WAAS optimization for RDP increased the number of users with good UE from 1 to 15 with 90% optimization. If customers can achieve even 60% optimization with Cisco WAAS, it would still be significantly higher than without Cisco WAAS.

Management and Operations

Overview

An end-to-end Cisco Virtual Workspace (VXI) Smart Solution deployment requires a comprehensive management architecture that can provision, monitor, and troubleshoot the service for a large number of users on a continuous basis. Several applications and tools are available to assist the administrator. While there is no one single application that can control all the aspects of Cisco Virtual Workspace (VXI) Smart Solution, a clear understanding of the capabilities and positioning of these separate tools is critical for their effective usage.

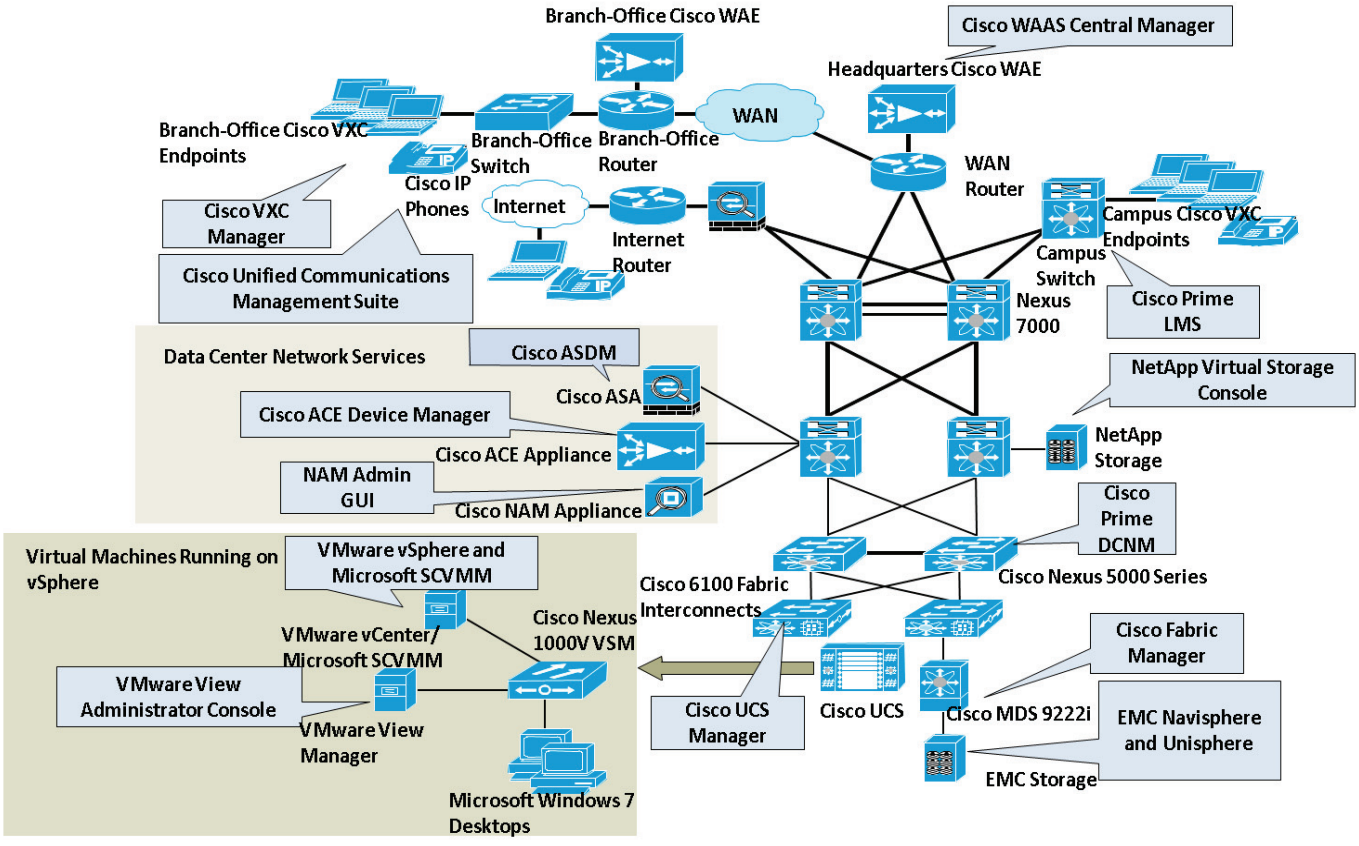
This chapter contains an overview of the management functions needed for an efficient Cisco Virtual Workspace (VXI) Smart Solution environment. It also describes the management aspects of user profiles, applications, and desktops; highlights applications available for system management; and discusses some commonly encountered situations and resolutions. The chapter describes tested and

recommended tools; discusses their effective utilization; and provides general design guidance regarding deployment. Before actual deployment and usage, readers should review the detailed documentation provided by the tool vendor.

Management Functions

Figure 50 provides an overview the various management applications that have been verified with the Cisco Virtual Workspace (VXI) Smart Solution architecture.

Figure 50 Management Tools in Cisco Virtual Workspace (VXI) Smart Solution



Cisco Virtual Workspace (VXI) Smart Solution management functions can be broadly divided into the following functional areas:

- **Provisioning and configuration**: Provision end users, virtual desktops, and endpoints using batch provisioning tools and templates. Vendor-provided APIs (XML) can be used for automation and self-service provisioning. These tools include software image and application management on endpoints and virtual desktops. Proper configuration helps ensure that the proper policy is applied for access, applications, etc.

- **Device monitoring**: Monitor the status of every element in real time and obtain diagnostics. Simple Network Management Protocol (SNMP), syslog, XML-based monitoring and HTTP-based interfaces are used to manage devices. This function also includes inventory and asset management (hardware and software) of endpoints and virtual desktops
- **Quality-of-service (QoS) monitoring**: Monitor and troubleshoot the status and quality of experience (QoE) of user sessions. This function includes the use of packet capture and monitoring tools such as Cisco Network Analysis Module (NAM), Cisco NetFlow, and Wireshark to monitor a session. It also enables the desktop virtualization administrator to remotely access the endpoint and virtual desktop to observe performance and collect bandwidth and latency measurements. These tools also can measure computing, memory, storage, and network utilization in real time to identify bottlenecks or causes of service degradation. Session details records for a virtual desktop session can indicate connection failures and quality problems
- **Statistics collection and reporting**: Collect quality and resource use measurements and generate reports useful for operations, infrastructure optimization, and capacity planning. Measurements include session volume, service availability, session quality, session detail records, resource utilization, and capacity across the system. The reports can be used for billing purposes and for service-level management

Each of these functions is then applied to Cisco Virtual Workspace (VXI) Smart Solution core building blocks, which are classified into the following categories:

- Data center (computing and storage, load balancing, and switching)
- Network (power, Cisco Wide Area Application Services [WAAS], QoS, and security)
- Endpoints and users (Cisco Virtualization Experience Client [Cisco VXC] and user profiles)
- Unified communications

Design Considerations for Management Tools

- **Scalability**: Management tools should be scalable for a large number of endpoints through a centralized control point. The use of polling, SNMP traps, and syslog messages is crucial for monitoring a large-scale deployment. GUI-based tools are helpful for navigating complex deployments and viewing detailed reports. Many device management tools provide an API or command-line interface (CLI) for automating workflows. This feature is particularly useful when adding a large number of users, desktops, or endpoints.
- **High availability**: Management applications should be deployed in redundant configurations (primary and secondary servers). Configurations and databases should be backed up periodically. When possible, deploy management application servers on virtual machines, to use resources efficiently and capitalize on high-availability features provided by the hypervisor infrastructure, such as virtual machine migration (VMware vMotion), VMware Distributed Resource Scheduling (DRS), and VMware Fault Tolerance (FT). In general, the best place to locate management servers is in the data center with other critical resources.
- **Traffic separation**: As a general guideline, a separate IP network for management traffic is recommended. For example, a dedicated IP subnet and VLAN for remote-access, SNMP, syslog, and FTP traffic could be managed out of band, where practical. This approach helps ensure that end-user and administrative traffic flows do not compete for, or interfere with, available bandwidth, and that remote access to a device is not compromised when the device needs to be reset or provisioned. Traffic separation also mitigates threats to network security and availability that could be introduced when end-user and administrative traffic share the same interface. The isolation of management and user traffic is especially important in the data center, where virtualization of the desktops concentrates user traffic in an infrastructure traditionally used for server and management traffic

loads. The isolation of management traffic may not be practical in some parts of the system. For example, an endpoint usually has a single Ethernet port, so it must use this port for both end-user and administrative traffic. Also, it may not be practical to set up a separate out-of-band network dedicated to management traffic across a WAN due to cost and network address conservation concerns. In these cases, keep in mind that certain types of management traffic (SNMP polling) can consume substantial bandwidth and should be scheduled appropriately, monitored, and possibly rate limited using standard QoS techniques.

Each management tool uses a specific set of protocols to communicate with devices. Refer to the vendor documentation for a complete list of protocols and ports used by each tool. Make sure that these ports are open on all intermediary routers, switches, and firewalls.

Cisco Virtual Workspace Management Tool Summary

Data Center and Applications

The main data center components that need to be managed are the computing servers, hypervisor, virtual machines, storage, switching fabric, connection managers, and servers that provide network services. The management tasks include the provisioning of end users, virtual desktops, desktop pools, hypervisor, and storage, as well as the monitoring of sessions and use of resources (computing, memory, storage, and network). Table 46 summarizes the components and management tools.

Table 46 Data Center and Applications - Management Tools

Product	Management Tool	Description	Product Documentation Link
VMware ESX, ESXi and virtual machines	VMware vCenter and vSphere client	Use VMware ESX and ESXi hypervisor manager to create and manage virtual machines	VMware vSphere documentation
VMware View Manager 5.0	VMware View Administrator Console	Create virtual desktop pools, and grant user privileges, and monitor sessions.	VMware View documentation
EMC Unified Storage	EMC Unisphere Management Suite	Provision and monitor the SAN-based storage array	EMC Unisphere Management Suite
NetApp FAS 3170	NetApp Virtual Storage Console	Provision and manage NetApp Unified Storage arrays	NetApp Virtual Storage Console
Cisco UCS B-Series Blade Servers	Cisco UCS Manager	Provision and monitor the Cisco UCS B-Series Blade Servers	Cisco UCS Manager

Product	Management Tool	Description	Product Documentation Link
Virtual desktops (guest OS)	Standard enterprise desktop and OS management tools (Microsoft Systems Management Server [SMS], and System Center Configuration Manager [SCCM])	Provision and monitor the virtual desktops.	Microsoft Systems Center
Microsoft Active Directory (AD), Domain Name System (DNS), and Dynamic Host Control Protocol (DHCP)	Standard enterprise management tools	Manage end user profiles and perform authentication of user sessions. Provide DHCP services to endpoints.	Microsoft Active Directory and Network services

Network Infrastructure

The main elements in the network infrastructure (LAN, WAN, and SAN) that need to be managed are the switches, routers, WAN acceleration devices, load balancers, security gateways, and network analyzers. These components span the data center, enterprise core, WAN, and branch offices and provide both data and storage connectivity. The management tasks include provisioning and monitoring these elements, monitoring a desktop virtualization session, and reporting the status of the network.

Table 47 Network Infrastructure - Management Tools

Product	Management Tool	Description	Product Documentation Link
Cisco Network Analysis Module (NAM)	Cisco NAM Admin GUI	Provision and monitor Cisco NAM	Cisco NAM documentation
Cisco Application Control Engine (Cisco ACE)	Cisco ACE Device Manager	Provision and monitor Cisco ACE	Cisco ACE documentation
Cisco Wide Area Application Service (Cisco WAAS)	Cisco WAAS Central Manager	Provision and monitor Cisco WAAS appliances, generate aggregate reports on optimization	Cisco WAAS documentation
Cisco Adaptive Security Appliance (Cisco ASA)	Cisco Adaptive Security Device Manager	Provision and monitor Cisco ASA	Cisco ASA documentation
Cisco AnyConnect VPN Client	Cisco AnyConnect Profile Editor	Provision Cisco AnyConnect VPN Client	Cisco AnyConnect Documentation
Cisco MDS 9000	Cisco Fabric Manager	Provision and monitor the Cisco MDS 9000 Family SAN switches	Cisco Fabric manager documentation

Product	Management Tool	Description	Product Documentation Link
Cisco Nexus 7000/Cisco Nexus 5000/ Cisco Nexus 2000	Cisco Data Center Network Manager	Provision and monitor the Cisco Nexus 7000, Cisco Nexus 5000, Cisco Nexus 2000 Series and Cisco MDS 9000 family switches.	Cisco DCNM documentation
Cisco Nexus 1000v	Cisco NX-OS CLI, Cisco DCNM and vCenter	Provision and monitor the Cisco Nexus 1000v series	Cisco Nexus 1000v documentation
Cisco Catalyst® 6500, 4500, 4900, and 3560 Series Switches and Cisco 3900 and 2900 Series Integrated Services Routers (ISRs)	Cisco Prime LAN Management Solution (LMS)	Provision and monitor Cisco routers and switches	Cisco Prime documentation
Cisco security policy on routers, switches, and security appliances	Cisco Security Manager	Provision and monitor security policies on Cisco routers, switches, and security appliances	Cisco Security Manager documentation
Cisco IOS Netflow	Cisco Netflow Collector	Collect, analyze, and report on netflow data collected from routers and switches	Cisco Netflow Collector documentation
Cisco EnergyWise	Management application compatible with Cisco EnergyWise platform	Monitor and control power consumption on switches and endpoint devices	Cisco EnergyWise documentation

Virtualization Experience Clients

The desktop virtualization endpoints that need to be managed are the thin clients, the Cisco VXC 6215. The management of endpoints includes tasks to provision and monitor endpoints, update images, perform asset management, measure the quality of the user experience, and remotely access the endpoints. Use of enterprise desktop management tools such as Altiris, and other management applications such as Cisco VXC Manager, and Cisco Unified Communications Manager is recommended to manage the desktop virtualization endpoint OS. Table 48 lists management tools for these endpoints.

Table 48 Desktop Virtualization Endpoints - Management Tools

Product	Management Tool	Description	Product Documentation Link
Cisco VXC endpoints (except Cisco VXC 4000)	Cisco VXC Manager	Image, configuration, asset management, shadowing of VXC endpoints	Cisco VXC Manager documentation
(Microsoft SMS and SCCM)		Image, configuration, and asset management of desktop virtualization endpoints	

Unified Communications

The main unified communications elements that need to be managed are Cisco Unified IP Phones, Cisco Jabber clients, and Cisco Unified Communications servers (Cisco Unified Communications Manager, Cisco Unified Presence, and Cisco Unity® devices). The management of UC elements includes tasks to provision and monitor the IP phones, and Cisco Jabber clients, Cisco Unity accounts, and servers. The Cisco Unified Communication servers include an embedded web-based GUI management interface that can be used to complete the basic provisioning and monitoring tasks. The Cisco Unified Communications Management suite of tools is recommended for advanced management features, scalability, and performance.

Table 49 Unified Communications - Management Tools

Product	Management Tool	Description	Product Documentation Link
Cisco IP phones, and Cisco Jabber clients	Cisco Unified Communications Management Suite	Cisco Unified Operations Manager, Provisioning Manager, Service Monitor, and Service Statistics Manager	Cisco UCMS documentation
Cisco Unified Communications Manager, Cisco Unified Presence, and Cisco Unity Connection	Cisco Unified Communications Management Suite	Cisco Unified Operations Manager, Provisioning Manager, Service Monitor, and Service Statistics Manager	Cisco UCMS documentation

Managing Desktops

Desktop provisioning and monitoring is critical to successful management of a Cisco Virtual Workspace (VXI) Smart Solution system. The challenge of Cisco Virtual Workspace (VXI) Smart Solution is to deliver a customized Microsoft Windows desktop to every user without bringing all the PC management issues from the desktop to the data center. The complexity associated with meeting end user expectations for a virtual desktop infrastructure (VDI) desktop experience that matches their desktop PC experience

can lead to delays or failures with the implementation of desktop virtualization initiatives. It is critical to eliminate these impediments to broad-based Cisco Virtual Workspace (VXI) Smart Solution deployment in order to help IT organizations meet their intended strategic objectives and goals.

Virtual desktops and desktop virtualization endpoints can be managed using standard Microsoft Windows software management tools such as Altiris. These tools can be used to push out agent updates, OS fixes, security updates, and applications. The Microsoft Windows desktop user profiles can be managed and applied to desktops using tools such as VMware View Persona. These desktop personalization tools improve login times and prevent the corruption of user profiles.

Desktop Provisioning

By using a combination (one or more) of three key virtualization technologies in a best-practices framework, Cisco Virtual Workspace (VXI) Smart Solution deployment total cost of ownership (TCO; Operating Expenses [OpEx] and capital expenditure [CapEx]) may be substantially reduced and productivity gains can be achieved. These three technologies are User Virtualization, Application Virtualization and Desktop Layering. These three key virtualization technologies permit IT organizations to deliver desktops replicating the traditional end user PC experience whilst managing their data center desktop infrastructure using pooled, non-persistent HVD images.

- **User Virtualization** - Decouples the end-user OS and application settings from the underlying desktop image, stores each user's digital persona in a centralized management framework, and delivers the user's settings on-demand to the desktop running a generic Windows OS image
- **Application Virtualization** - permits Windows OS image independent packaging of desktop applications and on-demand delivery of those applications to sand-boxed containers in Windows desktops from a single centrally managed application repository. App Virtualization supports two delivery methods: centrally hosted or streamed to the desktop
- **Desktop Layering** - permits abstracted installation of operating system, application, and user persona binaries into a specific end user's HVD instance such that the binaries are all stored and managed independently of the OS, but appear to be installed directly into the Windows desktop at runtime.

When these technologies are combined, all of those characteristics that make a particular end user's desktop unique to that end user have been abstracted away from the Windows OS and may be managed as independent entities. The decoupling of the user persona from a specific Windows OS instance permits that persona to be applied to any Windows OS instance at any time. Decoupling applications through either virtualization or layering allows patches or upgrades to be applied to only a single virtualized or layered app instance, which will then benefit all users of the application the next time they launch it. In addition, the decoupling of user personas and applications from the Windows OS permits a very stripped-down, `vanilla' base golden image to be used in a pooled, non-persistent Cisco Virtual Workspace (VXI) Smart Solution model for a large variety of end users. This reduces the storage footprint and facilitates centralized patching and updates to the golden OS image. There will always be a design tradeoff between using persistent and non-persistent desktops in terms of the amount of customization a user is allowed to perform.

VMware View Manager Administrator Console

VMware View Manager is used for managing remote desktop sessions between desktop virtualization endpoints and virtual desktops. It authenticates users initiating sessions and redirects them to a virtual desktop. The VMware View Administrator Console is a web-based application used to deploy and manage desktop pools, control user authentication, monitor desktop use, examine system events, and

perform analysis. VMware View also includes the VMware View Agent running on the virtual desktop and the VMware View Client running on the desktop virtualization endpoint. Note that these are software components that also need to be managed as described here.

Features and Guidelines for Using VMware View Manager Administrator Console in Cisco Virtual Workspace (VXI) Smart Solution

Use the VMware View PowerCLI cmdlets and vdmadmin command line interface (CLI) commands to automate the task of adding a manual pool or monitoring active sessions or desktops. These tools can be used by an administrator who needs to provision a large number of desktops or by an end user who is self-provisioning through a web-based interface. Refer to the VMware View Manager documentation for more information.

You can customize the VMware View Manager configuration by exporting the configuration in LDAP Data Interchange Format (LDIF), making changes using a text editor and then importing it back. This process can also be used to back up the VMware View Manager configuration. The CLI-based commands are vdmimport and vdmexport. Refer to the VMware View Manager documentation for more information.

You should synchronize virtual desktop clocks to a network-based Network Time Protocol (NTP) server. As an alternative, you can synchronize the virtual desktop clocks with the VMware ESXi host, and the host can be synchronized with a central NTP server. This synchronization is accomplished on the virtual desktop using the VMware tools configuration option. Consult the VMware documentation for detailed guidelines for synchronizing the virtual desktop clocks.

Desktop Monitoring and Assessment

An important monitoring and reporting requirement in Cisco Virtual Workspace (VXI) Smart Solution deployments is a single-pane high-level view of the complete deployment. The individual IT departments that participate in a Cisco Virtual Workspace (VXI) Smart Solution deployment (desktop services, data center storage, data center computing, and networking) have tools appropriate to their specific domains, but in the event of an outage, a tool is needed that can be used by an IT help desk worker to quickly diagnose an end-user's desktop problem, identify the root cause, and route a trouble ticket to the right IT department to rapidly resolve the problem. Many end users simply report degraded performance in their HVD, making this support task more challenging. The system performance monitoring and reporting tool monitors and reports critical metrics such as active desktop sessions and session traffic performance problems.

The hypervisor management infrastructure provides tools that can be used for desktop monitoring. Resource (CPU, memory, storage, and network) use on individual desktops can be measured and monitored in real time by using the VMware statistics logging application. Consider using the Intel IOMeter tool and the built-in Microsoft Windows Task Manager, which display resource utilization and performance measurements. These tools are also useful for troubleshooting purposes. VMware vCenter can be used to capture these measurements on a per–virtual machine basis.

Measuring the resource utilization of physical desktops enables administrators to assess the physical desktop infrastructure prior to migration to a virtualized environment. It also enables capacity planning for future resource build-out (computing, memory, storage, and network resources) and post-deployment QoS monitoring on virtual desktops. An audit log of the files being accessed and the applications being invoked on the desktop can also be useful for monitoring desktop use and troubleshooting performance problems. Consider using Microsoft Windows Performance Manager in addition to any other software or application probes that can be installed on the desktop.

Virtual Desktop Assessment

Design Guidelines

A comprehensive pre-deployment assessment of the exiting desktop computing infrastructure; virtual desktop implementation design and capacity planning; and post-deployment optimization and monitoring of virtual desktops are important aspects of a successful implementation of a Cisco Virtual Workspace (VXI) Smart Solution. The best way to assess an organization's readiness for desktop virtualization, and its potential benefits, is to develop a set of realistic and well-defined business drivers and a clear snapshot of the current infrastructure. The tools that perform virtual desktop pre-assessment, design, implementation, monitoring, optimization and capacity planning are key to performing this evaluation. These tools provide the following capabilities: ability to understand existing desktop computing resource usage and user experience, ability to do capacity planning for desktop virtualization deployment, ability to monitor virtual desktop resource usage and user experience, and ability to demonstrate and report on benefits of desktop virtualization. The goal of the assessment is to determine readiness and requirements for desktop virtualization. Determining suitability of a desktop or user for desktop virtualization is part of the design and planning stage of a deployment.

These tools also provide visibility into desktop computing resource usage, application workload, user behavior, and user experience for large desktop deployments. Typical measurements provided by these tools include usage metrics in terms of compute, memory, storage, and network resources as well as desktop user experience metrics in terms of application response times, login times and latency measurements. A detailed hardware and software profile of the desktops is also compiled that includes operating system, hardware profile, applications, user settings, peripherals (printers), and monitor usage. Performing a baseline user experience assessment prior to virtualization is important in assessing the user experience in a virtual desktop deployment.

After data gathering is complete, the assessment tools generate reports that are used to identify the optimal and sub-optimal desktop candidates for desktop virtualization. The goal of the assessment is not a final design, but rather an analysis of the feasibility of a desktop virtualization project given the combination of network infrastructure, user applications, and desktop environment. The relative suitability or fit for virtualization is determined using an objective fitness rating based on composite metric using pre-established thresholds. Assessment tools can also provide estimate of productivity improvement based on reduction in login times, or due to benefits of desktop high availability and replacement time reduction, as well as reduction of application load times.

The assessment tool typically requires the installation of agent software on the desktop for data collection and communication of the data to a central repository for aggregation and analysis. The operation of the agent software is transparent to the end user and should not impact the performance of the desktop. It is important that the ports used for this communication are open on all intermediary network elements. As thousands of desktops can be monitored, it is important to allocate enough storage for the database on the central repository. The agent software can be installed manually on a desktop or be distributed automatically using an LDAP/GPO or other standard tool used for updating desktops (Microsoft SMS). Once data is collection begins, it is collected continuously for the duration of the evaluation (for instance, 30 days). The collected data will allow the creation of a detailed picture of the machines, application, and user inventory of the desktops being monitored.

Some indications of poor desktop performance include:

- Slow User Logins identifies logins taking longer than specified threshold
- Slow Application Load Times identifies applications that took longer than specified number of seconds to load
- Top CPU Consuming Applications shows the applications consuming the most CPU cycles
- Top Memory Consuming Applications shows the applications using the most memory

- Machines with High CPU Usage identifies machines using more than specified threshold percentage of CPU
- Machines with High Memory Usage identifies machines using more than specified threshold percentage of Memory

The assessment and monitoring tool provides real-time display of inspection data, diagnostic reports, and assessment reports. The tool provides summary views of inspection data using graphs and charts to help analyze trends and identify potential problems. These views are useful when examining the health of the Cisco Virtual Workspace (VXI) Smart Solution deployment. When specific measurement thresholds are passed, alerts can be sent via e-mail or RSS feed. Automated programs can be used to poll the RSS feeds and to pass on the data to other systems. It is recommended to schedule a report to be sent via e-mail on daily basis to indicate how many machines are reporting data to ensure that data collection is progressing without any issues.

Deployment, Configuration, and Best Practices

It is recommended to use a dedicated network monitor attached to the virtual switch to monitor virtual desktop network traffic and report on application usage. Proper network planning and provisioning is important in order to allocate the correct bandwidth and to ensure low latency for users accessing their virtual desktops across WAN.

It is recommended to establish a baseline for user experience prior to migration. A post-deployment user experience measurement quantifies the impact to the end user during a deployment or pilot. The assessment will assist in determining how to provision the virtual desktops (resources, images, and applications) and which desktop protocol to use, and which desktop protocol features to enable (peripherals).

Identifying applications that are poor fit for virtualization is key to a successful Cisco Virtual Workspace (VXI) Smart Solution deployment. Ensure that applications currently used operate correctly and without performance impact in a virtualized environment especially when scaled to tens of virtual desktops on a single host machine.

Cisco Advance Services provides a Cisco Virtual Workspace (VXI) Smart Solution planning, design and monitoring service that will conduct the evaluation and generate an assessment report and virtualization design plan. The outputs of the services include Cisco Virtual Workspace (VXI) Smart Solution design documents, operation readiness and implementation plans, self-service guides, PoC report, test plan, and a readiness report. Please consult Cisco AS for further information at

http://www.cisco.com/en/US/products/ps10374/services_segment_service_home.html

Using System Level Tools

This section briefly describes the Microsoft Systems Center that can help provide composite views into the whole deployment.

Microsoft System Center

Microsoft System Center 2012 consists of a family of products. These applications provide tools so that system administrators can deploy, install, configure, monitor, detect, diagnose, and correct problems in the computing environment.

Cisco has tested and verified the Microsoft System Center Operations Manager (SCOM) with Cisco Virtual Workspace (VXI) Smart Solution. Microsoft SCOM uses an XML-based language to describe the system being managed. The capability to create such descriptions, called management packs, makes

Microsoft SCOM extensible and customizable. Cisco has defined such a pack that can be applied to Microsoft SCOM. Using this pack, Microsoft SCOM can now effectively manage Cisco UCS servers, including chassis and blades, and can track and correlate faults and events, etc.

For more information about this management pack, including instructions for downloading, installing, and using it, see http://developer.cisco.com/web/unifiedcomputing/systemcenter.

This management pack extends the management capabilities of Cisco UCS Manager to Microsoft SCOM. All the hardware information, including processors, memory, etc., normally monitored from Cisco UCS Manager can now be viewed and monitored from Microsoft SCOM as well. All the alerts that Cisco UCS Manager generates are also now available on Microsoft SCOM. These alerts may be acted upon from within Microsoft SCOM, and their status is synchronized between Microsoft SCOM and Cisco UCS Manager. This integration is particularly useful for organizations that already employ Microsoft System Center in the management framework.

The integration of Microsoft System Center with Cisco UCS is a step toward management from a single pane. Along with the Cisco UCS management pack, packs available from other Cisco Virtual Workspace (VXI) Smart Solution component vendors can be integrated into Microsoft MS System Center to provide a complete platform from which to monitor and control all aspects of the Cisco Virtual Workspace (VXI) Smart Solution deployment.

Summary

Management strategy should be well thought out, defined and built into the deployment. There are several tools from Cisco and other vendors that are tested and validated to work with Cisco Virtual Workspace (VXI) Smart Solution. Some tools, such as those for endpoint management, are mandatory, but others are optional, although highly recommended for any large deployment. At this time there is no single tool that can be used to manage all aspects of Cisco Virtual Workspace (VXI) Smart Solution. This limitation mandates that administrators develop a good understanding of the capabilities of available tools and how to correlate their results.

Virtual Workplace References

Virtualized Data Center

Cisco Data Center Design Zone

http://www.cisco.com/en/US/netsol/ns743/networking_solutions_program_home.html

Best Practices in Deploying Cisco Nexus 1000V Series Switches on Cisco UCS B Series Blade Servers

http://www.cisco.com/en/US/prod/collateral/switches/ps9441/ps9902/white_paper_c11-558242.html

Cisco Validated Design – VMWare View on Cisco UCS and EMC Storage

http://www.cisco.com/en/US/docs/solutions/Enterprise/Data_Center/Virtualization/ucs_view_emc.html

EMC FAST VP for Unified Storage

http://www.emc.com/collateral/software/white-papers/h8058-fast-vp-unified-storage-wp.pdf

EMC Clariion, and VNX Fast Cache

http://www.emc.com/collateral/software/white-papers/h8046-clariion-celerra-unified-fast-cache-wp.pdf

EMC VNX Deduplication and Compression

http://www.emc.com/collateral/hardware/white-papers/h8198-vnx-deduplication-compression-wp.pdf

NetApp TR-3298: RAID-DP: NetApp Implementation of RAID Double Parity for Data Protection

http://media.netapp.com/documents/tr-3298.pdf

NetApp TR-3347: FlexClone Volumes: A Thorough Introduction

http://media.netapp.com/documents/tr-3347.pdf

NetApp TR-3437: Storage Best Practices and Resiliency Guide

http://media.netapp.com/documents/tr-3437.pdf

NetApp TR-3505: NetApp Deduplication for FAS, Deployment and Implementation Guide

http://media.netapp.com/documents/tr-3505.pdf

NetApp TR-3563: NetApp Thin Provisioning

http://media.netapp.com/documents/tr-3563.pdf

Modular Data Center Blocks

Cisco Virtual Workspace (VXI) Smart Solution As-Built Reference Guide

http://www.cisco.com/en/US/docs/solutions/Enterprise/Data_Center/Cisco VXI/configuration/VXI_Config_Guide.pdf

Datacenter Deployment Guide

http://www.cisco.com/en/US/docs/solutions/Enterprise/Borderless_Networks/Smart_Business_Architecture/February2012/SBA_Mid_DC_DataCenterDeploymentGuide-February2012.pdf

Auto Smartports on Cisco Catalyst Switches

http://www.cisco.com/en/US/docs/switches/lan/auto_smartports/12.2_55_se/configuration/guide/iosaspcg.pdf

Location Tracking - Cisco Catalyst Switch Configuration

http://www.cisco.com/en/US/docs/switches/lan/catalyst3750/software/release/12.2_50_se/configuration/guide/swlldp.html

WCS - Context-Aware Services configuration guide

http://www.cisco.com/en/US/partner/docs/wireless/mse/3350/7.0/CAS/configuration/guide/CAS_70.html

High-Availability - Cisco Catalyst Switches HA design guide

http://www.cisco.com/en/US/prod/collateral/switches/ps5718/ps4324/prod_white_paper0900aecd805e6a95.html

WAN Deployment Guide

http://www.cisco.com/en/US/solutions/collateral/ns340/ns414/ns742/ns982/c07-610746-02_wanDeploy.pdf

Dynamic Multipoint VPN (DMVPN)

http://www.cisco.com/en/US/products/ps6658/index.html

Cisco WAAS - Technical Overview

http://www.cisco.com/en/US/partner/prod/collateral/contnetw/ps5680/ps6870/white_paper_c11-705319.html

Introduction to the Cisco WAAS Central Manager GUI

http://www.cisco.com/en/US/docs/app_ntwk_services/waas/waas/v421/configuration/guide/intro.html#wp1122501

Configuring Cisco WAAS Traffic Interception

http://www.cisco.com/en/US/docs/app_ntwk_services/waas/waas/v4013/configuration/guide/traffic.html

Configuring Cisco WAAS Application Acceleration

http://www.cisco.com/en/US/docs/app_ntwk_services/waas/waas/v4013/configuration/guide/policy.html

Configuring Cisco WAAS Network Settings

http://www.cisco.com/en/US/docs/app_ntwk_services/waas/waas/v4013/configuration/guide/network.html

Configuring and Managing Cisco WAAS Print Services

http://www.cisco.com/en/US/docs/app_ntwk_services/waas/waas/v4013/configuration/guide/printsvr.html

PfR - Enhancing WAN experience using PfR and Cisco WAAS

http://www.cisco.com/en/US/prod/collateral/iosswrel/ps6537/ps6554/ps6599/ps8787/prod_white_paper0900aecd806c5077.html

Data Center - Network infrastructure

http://www.cisco.com/en/US/docs/solutions/Enterprise/Data_Center/DC_3_0/DC-3_0_IPInfra.html

Campus Network High Availability Design

http://www.cisco.com/en/US/docs/solutions/Enterprise/Campus/HA_campus_DG/hacampusdg.html

Branch/WAN Design

http://www.ciscosystems.com/en/US/docs/solutions/Enterprise/Branch/srlgbrnt.pdf

Cisco ACE - Configuring Network Access

http://www.cisco.com/en/US/docs/app_ntwk_services/data_center_app_services/ace_appliances/vA1_7_/configuration/device_manager/guide/UG_ntwk.html

Cisco ACE - Remote access

http://www.cisco.com/en/US/docs/interfaces_modules/services_modules/ace/v3.00_A1/configuration/administration/guide/access.html

Cisco ACE - Device Manager Guide

http://www.cisco.com/en/US/docs/app_ntwk_services/data_center_app_services/ace_appliances/vA1_7_/configuration/device_manager/guide/dmguigd.html

Cisco ACE - Configuring Session Persistence

http://www.cisco.com/en/US/docs/app_ntwk_services/data_center_app_services/ace_appliances/vA1_7_/configuration/device_manager/guide/UG_lb.html#wp1062118

Cisco ACE - Configuring Health Monitoring

http://www.cisco.com/en/US/docs/app_ntwk_services/data_center_app_services/ace_appliances/vA1_7_/configuration/device_manager/guide/UG_lb.html#wp1045366

Cisco ACE - Balancing Algorithm

http://www.cisco.com/en/US/docs/app_ntwk_services/data_center_app_services/ace_appliances/vA1_7_/configuration/device_manager/guide/UG_lb.html#wp1045366

Cisco ACE - Configuration of Source NAT

http://docwiki.cisco.com/wiki/Basic_Load_Balancing_Using_One_Arm_Mode_with_Source_NAT_on_the_Cisco_Application_Control_Engine_Configuration_Example

Cisco UCS - Cisco UCS Manager Configuration Guide

http://www.cisco.com/en/US/docs/unified_computing/Cisco UCS/sw/gui/config/guide/1.3.1/UCSM_GUI_Configuration_Guide_1_3_1.pdf

Cisco Bring Your Own Device (BYOD) Smart Solution Design Guide

http://www.cisco.com/en/US/docs/solutions/Enterprise/Borderless_Networks/Unified_Access/byoddg.html

Rich Media, Collaboration and User Experience

Cisco Wide Area Application Services (Cisco WAAS) Configuration Guide

http://www.cisco.com/en/US/docs/app_ntwk_services/waas/waas/v421/configuration/guide/cnfgbook.pdf

Cisco Unified SRST Configuration Guide

http://www.cisco.com/en/US/docs/voice_ip_comm/cusrst/admin/srst/configuration/guide/SRST_SysAdmin.pdf

Cisco Unified Communications 9.X SRND

http://www.cisco.com/en/US/docs/voice_ip_comm/cucm/srnd/9x/uc9x.html

Cisco Jabber for Windows Administration Guide

http://www.cisco.com/en/US/docs/voice_ip_comm/jabber/Windows/9_0_1/b_jabber_win_icg.html

Cisco Virtual Workspace (VXI) Smart Solution implements this architecture by leveraging Cisco VXME, which enables Cisco Jabber for virtualized environments.

Cisco SAFE Architecture Guide

http://www.cisco.com/en/US/docs/solutions/Enterprise/Security/SAFE_RG/SAFE_rg.pdf

Cisco Data Center 3.0 Security Guide

http://www.cisco.com/en/US/docs/solutions/Enterprise/Data_Center/DC_3_0/dc_sec_design.pdf

Cisco Identity-Based Network Security Guide

http://www.cisco.com/en/US/solutions/collateral/ns340/ns394/ns171/CiscoIBNS-Technical-Review.pdf

Cisco 3560 Command Reference Guide

http://www.cisco.com/en/US/docs/switches/lan/catalyst3560/software/release/12.2_35_se/command/reference/cli1.html#wp2757193

Cisco Business Ready Teleworker Design Guide

http://www.cisco.com/application/pdf/en/us/guest/netsol/ns107/c649/ccmigration_09186a008073377d.pdf

Cisco ASA 9.x: VPN Access with the Cisco AnyConnect VPN Client Configuration Example

http://www.cisco.com/en/US/products/ps6120/products_configuration_example09186a00808efbd2.shtml

Scaling and High Availability

VMware View 5.0 Documentation

http://www.VMware.com/support/pubs/view_pubs.html

VMware View 5 Performance and Best Practices - Technical White Paper

http://www.VMware.com/files/pdf/view/VMware-View-Performance-Study-Best-Practices-Technical-White-Paper.pdf

VMware KB article on HIMP(HaltingIdleMsecPenalty) Parameter

http://kb.VMware.com/selfservice/microsites/search.do?language=en_US&cmd=displayKC&externalId=1020233

Desktop Virtualizatio with VMware View5

http://www.vmware.com/files/pdf/techpaper/PCoIPvHDXsinglesession03-05-12.pdf

Workload Considerations for Virtual Desktop Reference Architectures

http://www.VMware.com/go/view4rawc&ei=hFXQTr6QDKbjiAKk84ngCw&usg=AFQjCNEIFL

Virtual Reality Check - Phase III and Phase IV http://www.projectvrc.com/white-papers.html Desktop Virtualization with View 4.5 and EMC Storage

http://www.cisco.com/en/US/solutions/ns340/ns414/ns742/ns743/ns993/landing_dcVirt-VM_EMC.html

Desktop Virtualization with View 4.5 and NetApp Storage

http://www.cisco.com/en/US/solutions/ns340/ns414/ns742/ns743/ns993/landing_dcVirt-VM_netapp.html

VMware View on NetApp Deployment Guide

http://media.netapp.com/documents/tr-3770.pdf

Scalability Study for Deploying VMware View on Cisco UCS and EMC Symmetrix V-Max Systems

http://www.cisco.com/en/US/docs/solutions/Enterprise/Data_Center/App_Networking/vdiucswp.html

Enterprise VMware View 5 Network Optimization with PCoIP

http://www.VMware.com/files/pdf/view/VMware-View-5-PCoIP-Network-Optimization-Guide.pdf

Network Design for View

http://www.cisco.com/en/US/solutions/ns340/ns414/ns742/ns743/ns993/landing_dcVirt-vm_view4.html

VMware vSphere 5.0 Configuration Maximums

http://eas-web2.cisco.com/vxi-cvd/vmware/2.6/wp-admin/www.vmware.com/pdf/vsphere5/r50/vsphere-50-configuration-maximums.pdf

Performance Best Practices for VMware vSphere 5.0

http://www.VMware.com/pdf/Perf_Best_Practices_vSphere5.0.pdf

Memory Resource Management in VMware ESX Server

http://www.vmware.com/resources/techresources/531

VMware KB Article on Transparent Page Sharing Memory

http://kb.VMware.com/kb/1021095

Performance Study of VMware vStorage Thin Provisioning

http://www.vmware.com/pdf/vsp_4_thinprov_perf.pdf

VMware KB: Using thin provisioned disks with virtual machines

http://www.VMware.com/pdf/vsp_4_thinprov_perf.pdf

vSphere 5.0 Resource Management

http://pubs.VMware.com/vsphere-50/topic/com.VMware.ICbase/PDF/vsphere-esxi-vcenter-server-50-resource-management-guide.pdf

VMware KB: Using thin provisioned disks with virtual machines

http://kb.VMware.com/kb/1005418

Management and Operations

Cisco UCS GUI Configuration Guide

http://www.cisco.com/en/US/docs/unified_computing/Cisco UCS/sw/gui/config/guide/1.3.1/UCSM_GUI_Configuration_Guide_1_3_1.pdf

Cisco Unified Communications Manager Systems Guide

http://www.cisco.com/en/US/docs/voice_ip_comm/cucm/admin/8_0_2/ccmsys/accm.pdf

EMC Unisphere management suite

http://www.cisco.com/wwworkarea/6662204388087300332/WWStaging/TestJob/Projects/Untitled%201%20Project/Temp/products/detail/software/unisphere.htm#wpxref

NetApp Virtual Storage Console

http://www.cisco.com/wwworkarea/6662204388087300332/WWStaging/TestJob/Projects/Untitled%201%20Project/Temp/us/products/management-software/vsc/virtual-storage-console.html#wpxref

Microsoft Active Directory and Network services

http://www.microsoft.com/

Cisco Network Analysis Module Deployment Guide

http://www.cisco.com/en/US/prod/collateral/modules/ps2706/white_paper_c07-505273.html

Cisco NAM Appliance Documentation (command reference and user guide)

http://www.cisco.com/en/US/partner/products/ps10113/tsd_products_support_series_home.html

Cisco Wide Area Application Services Configuration Guide (Software Version 4.1.1)

http://www.cisco.com/en/US/docs/app_ntwk_services/waas/waas/v411/configuration/guide/cnfg.html

Cisco Adaptive Security Device Manager Configuration Guide

http://www.cisco.com/en/US/products/ps6121/tsd_products_support_configure.html

Cisco ACE 4700 Series Application Control Engine Appliances Documentation

http://www.cisco.com/en/US/products/ps7027/tsd_products_support_series_home.html

Cisco UCS Manager Configuration Guide

http://www.cisco.com/en/US/products/ps10281/products_installation_and_configuration_guides_list.html

Cisco DCNM documentation

http://www.cisco.com/en/US/products/ps9369/tsd_products_support_configure.html

Cisco Fabric manager documentation

http://www.cisco.com/en/US/partner/products/ps10495/tsd_products_support_configure.html

Cisco Nexus 1000v documentation

http://www.cisco.com/en/US/partner/products/ps9902/products_installation_and_configuration_guides_list.html

Cisco Prime documentation

http://www.cisco.com/en/US/partner/products/sw/cscowork/ps4565/tsd_products_support_maintain_and_operate.html

Acronyms

Table 50 defines the acronyms and abbreviations used in this publication.

Table 50 List of Acronyms

Acronym	Expansion
CNA	Converged Network Adapter
FC	Fibre Channel
FCoE	Fibre Channel over Ethernet
HBA	Host Bus Adapter
HVD	Hosted Virtual Desktop
IOPS	I/O per second
iSCI	Small Computer System Interface over IP
LAN	Local Area Network
NFS	Network File Share
PCoIP	PC over IP
RDP	Remote Desktop Protocol
TPS	Transparent Page sharing
vDC	Virtual Data Center
VEM	Virtual Ethernet Module
VM	Virtual Machine
vPC	Virtual Port Channel
VSM	Virtual Switch Module
VoIP	Voice Over IP
WAN	Wide Area Network