White Paper

Cisco Virtual Workspace (VXI) Smart Solution: Address the Collaboration Challenges of Virtualization

What You Will Learn

Organizations that need to reduce operating costs, increase business agility, maintain strict data compliance laws, and address evolving workspace requirements such as bring-your-own-device (BYOD) initiatives often consider deploying virtualized desktop infrastructure (VDI). However, many of these organizations do not progress from the pilot to the deployment phase because of poor user experience with collaborative applications involving digital interactive media, often referred to as rich media.

The purpose of this document is to discuss rich-media challenges related to collaborative applications in a virtual desktop environment and the architecture that the Cisco® Virtual Workspace (VXI) Smart Solution (formerly called Cisco Virtualization Experience Infrastructure [Cisco VXI™]) provides to address these challenges. This document explains the factors that degrade the user experience and describes the unique Cisco solution that creates a positive user experience. This document discusses the following topics:

- Use case showing a video call with and without VDI
- Architectural challenges created by media when VDI is used
- Impact of rich media on a VDI deployment
- How the Cisco Virtual Workspace (VXI) Smart Solution overcomes these challenges

Target Audience

This document is intended to assist solution architects, sales engineers, field engineers, and consultants in the planning, design, and deployment of the Cisco Virtual Workspace (VXI) Smart Solution. This document assumes the reader has an architectural understanding of the Cisco Virtual Workspace (VXI) Smart Solution system and has reviewed the Cisco Virtual Workspace (VXI) Smart Solution Release 2.7 Cisco Validated Designs based on Citrix XenDesktop, Citrix XenApp and VMware View.

Use Case: A Video Call

To illustrate the challenges of rich media in VDI, this section compares the data flow in a traditional desktop environment with that in a virtual desktop environment. The use case focuses on the media flow, which is the main challenge. Later sections provide more details about call signaling and setup.

Video Call Using a Traditional Desktop

For desktop users running a soft client, a point-to-point video call results in the following data flow (Figure 1):

1. PC 1 captures audio and video data from the

4. PC 2 decompresses the audio and video

microphone and camera.
2. PC 1 compresses the audio and video data and sends it to PC 2.
3. The network expedites delay-sensitive traffic.

stream.
5. PC 2 renders the video to the screen.

Note: Steps 1 through 5 are repeated to send data from PC 2 to PC 1.

Figure 1 Real-Time Audio and Video Data Flow Between PC 1 and PC 2

In this video call, data is always compressed before it is pushed on the network. The media uses a defined protocol, Real-Time Transport Protocol (RTP), which enables the network to identify the traffic as media and apply quality-of-service (QoS) policy to improve the user experience.

Video Call With VDI

With a hosted virtual desktop (HVD), point-to-point communication results in the following data flow (Figure 2):

1. End-client 1 captures audio and video from the microphone and camera.
2. End-client 1 sends **uncompressed** audio and video data as part of the display protocol to the data center.
3. Virtual desktop 1 compresses the audio and video data.
4. Virtual desktop 1 sends the compressed audio and video data to virtual desktop 2.
5. Virtual desktop 2 decompress the audio and video data.
6. Virtual desktop 2 renders the video as a screen image and **sends it as screen data** through the display protocol to end-client 2. Audio is sent **uncompressed** within the display protocol.
7. End-client 2 displays the video on the screen and plays the audio on the speakers.

Note: Steps 1 through 7 are repeated to send data from end-client 2 to end-client 1.

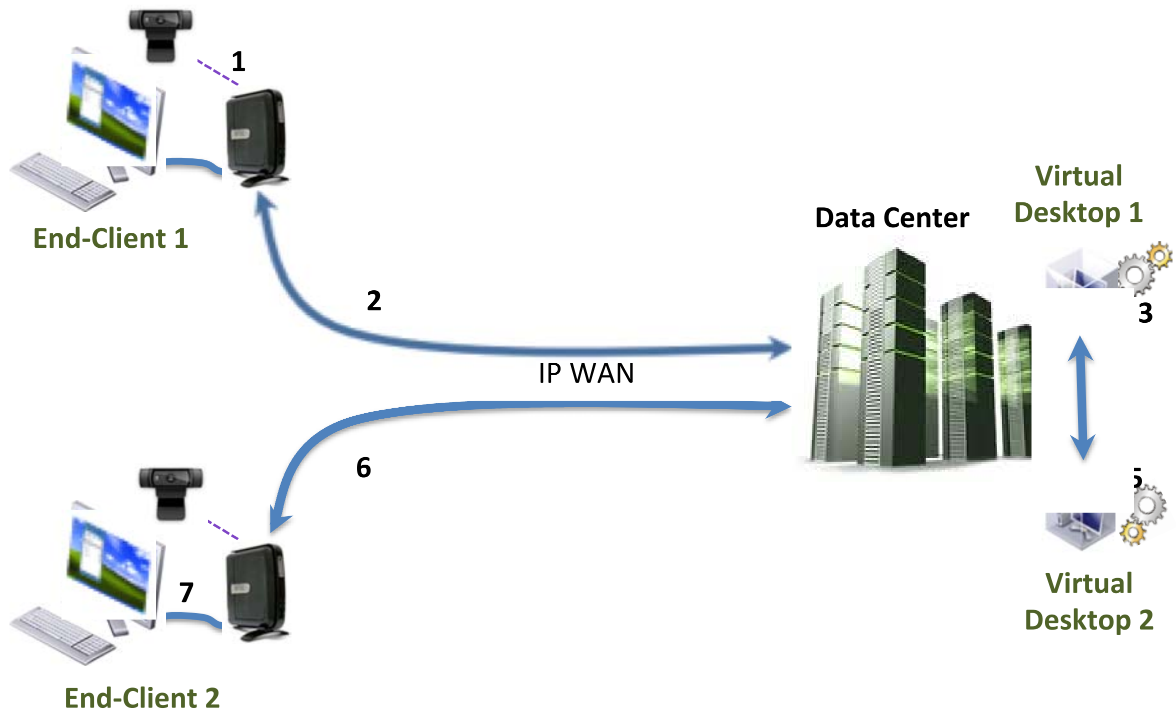

Figure 2 Real-Time Audio and Video Data Flow Between End-Client 1 and End-Client 2

In this video call, the media flows not directly between the endpoints, but rather as USB data into the data center. There it is reencoded and sent as RTP data between the two hosted desktops.

The challenges of this use case are described in the next section.

Architectural Challenges of Media in VDI

The video call use case illustrates the following three main challenges of media in VDI: suboptimal routing, encapsulation and encryption of VDI protocols, and media reprocessing.

Suboptimal Routing

In a video call between traditional desktops, media is routed optimally from one endpoint to the other.

In HVDs, all application and OS processing is performed in the data center. The end clients collect audio data from the audio port or a USB port and video data from a USB port and redirect it to the application running in the data center. This approach effectively creates a hub-and-spoke network (Figure 3).

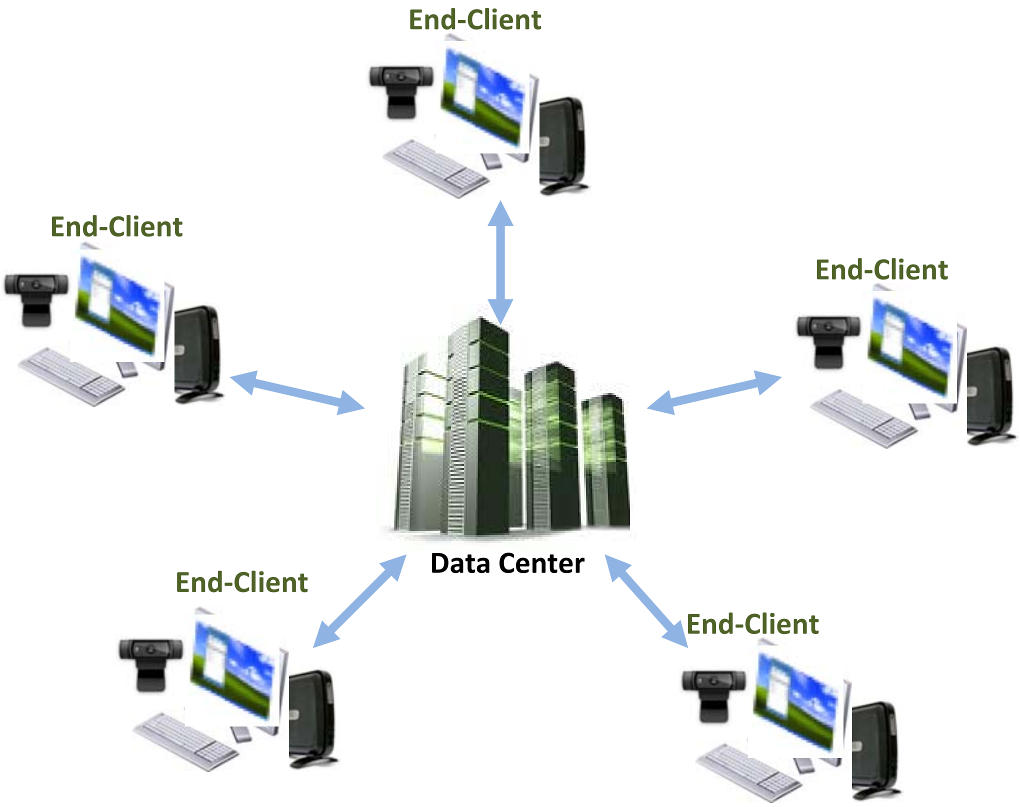

Figure 3 Hub-and-Spoke Network

This form of communication is appropriate for the desktop itself because 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 Session Initiation Protocol (SIP) to help ensure that even though call signaling may use a central point, media flows optimally from one endpoint to another. When data passes through a central point instead of being routed optimally through the network, the flow is referred to as hairpinning.

Some techniques are available to optimize the flow of media within the VDI session, but these do not fix this fundamental problem.

Encapsulation and Encryption of VDI Protocols

With a traditional desktop, applications run on the computer and establish independent network connections, often using well-defined ports and protocols. This process 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. That desktop is then viewed and controlled remotely by the user through a single virtual display protocol (PC over IP

[PCoIP], Citrix Independent Computing Architecture [ICA], Microsoft Remote Display Protocol [RDP], Virtual Network Computing [VNC], or others). These protocols encapsulate:

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

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

Media Reprocessing

In a video call between traditional desktops, media is encoded once by one of the endpoints, and decoded once by the other endpoint.

In a video call between HVDs, 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 data. The data is then sent to the other video application, which may be in the same data center. There the process is reversed as the data is decoded, reencoded as USB data, and sent to the other VDI endpoint.

This decoding and encoding significantly increases the CPU load of the hosted desktop and affects both the user experience and the scalability of the hardware hosting the virtual desktops.

Specific Rich-Media Challenges

The architectural details of suboptimal routing, encapsulation and encryption of VDI protocols, and media reprocessing lead to a number of challenges related to rich media, such as increased latency, increased bandwidth consumption, loss of network optimization, and increased CPU load on the VDI infrastructure.

Increased Latency

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

Consider a scenario in which two end clients are situated in the same building in San Jose, California, and the data center is situated in a different geographical location: Boston, Massachusetts. The hub-and-spoke architecture will require all real-time audio and video data to travel from end-client 1 (San Jose) to the data center (Boston) and then from the data center (Boston) to end-client 2 (San Jose).

In addition, reprocessing the media on the hosted desktop creates both CPU load and additional latency. These two together have a significant impact on the user experience.

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

Optimizing the flow of audio and video within the VDI session offers very little improvement because the fundamental problems of suboptimal routing and media reprocessing remain.

Increased Bandwidth Consumption

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

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

For example, consider a phone call between two employees in a branch office. Instead of flowing entirely within the local network, the call media will be sent over the WAN link to the data center and then back, doubling the amount of data and moving 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 as compressed RTP data.

The combined effects can make the data transfer rate in a virtual desktop environment almost 100 times greater than in a traditional desktop environment (see Figure 5 later in this document).

This scenario is a problem for individual users in normal use, and the effect is multiplied when many people make calls at the same time: for instance, during 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 most needed.

Loss of Network Optimization

When rich-media applications are deployed, the network is designed to provide 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 traffic is combined in this way, the network cannot recognize and differentiate the traffic sufficiently to provide the optimum QoS needed for rich-media flows.

Such lack of differentiation also reduces the capability of the network administrator to monitor and troubleshoot the network. For example, whereas before administrators could estimate, measure, and thereafter adapt their networks for rich-media applications, they can no longer do so because they have no visibility into what constitutes the rich-media traffic in a VDI session.

Inability to Provide Call Admission Control

Although provisioning the right amount of bandwidth in the network for rich-media sessions is critical, call admission control also must be applied 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 Call Admission Control (Locations-CAC) in Cisco Unified Communications 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.

Inability to Use the Voice VLAN

Voice VLANs are recommended for a variety of reasons, such as to prevent broadcast storms for voice devices, to provide a trusted means of providing QoS for rich-media traffic, and to allow multiple devices to be daisy-chained and connected to the network on a single port. These benefits are lost if the media is encapsulated within the VDI stream.

Loss of Emergency Calling Capabilities

In North America, Enhanced 911 services, and equivalent services in other countries, enable emergency operators to:

- Locate the user based on the calling number
- Call back the emergency caller if disconnection occurs

If the entire softphone is running in the data center, then emergency services will be unavailable or will incorrectly identify the location of the caller as the data center.

Increased CPU Load in the Data Center

Processing the media for the call in the data center increases the CPU load on the data center servers because the HVD has to:

- Decode USB audio and video from the endpoint
- Encode that media as RTP data for the call
- Decode incoming media from RTP data
- Encode incoming media as USB data to be sent to the endpoint

Media encoding and decoding is one of the most CPU-intensive tasks. For a call between traditional endpoints, only one encode and one decode operation occurs. The hosted desktop, however, has to perform more operations, significantly increasing its CPU utilization. Furthermore, while a traditional desktop generally has excess CPU capacity, hosted desktops are not generally designed with latent capacity. A large CPU processing requirement from one user can affect both the performance for that user and that for other users 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 voice-over-IP (VoIP) deployments. In many cases, these tools are already in place 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 they are also either mixed directly into the VDI desktop image generated in the data center or tunneled through the VDI protocol connection as a virtual stream. In both cases, network management tools cannot differentiate among VDI desktop image packets, multimedia packets, or even USB redirection packets. Typically, in VDI all these streams are intertwined and delivered to the endpoint through a single IP connection.

Approaches like Citrix Multi-Stream ICA provide some benefits for network monitoring and management tools; for example, Multi-Stream ICA creates four different IP connections, which can have differentiated treatment in the network. Each of these four connections may still have multiple streams, which are mixed and tunneled through the single connection. The network still cannot recognize different streams in the same connection, nor can it apply different treatment to each stream. Additionally, Multi-Stream ICA does nothing to address or ameliorate the hairpinning problems of VDI protocols.

Media in VDI sessions makes 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

Cisco Unified Survivable Remote Site Telephony (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 or phone application is running in the data center, SRST in the branch office does not help. If the WAN link fails, then employees in the branch office lose not only their desktops, but also their rich-media capabilities.

Device Variety and Mobility

Advances in technology during the past several years, especially in computing and mobility, present new possibilities for workers and allow them to be more productive away from the office. Yesterday's uniform computing devices (Microsoft Windows–based laptops and desktops) provided and managed by the IT department are giving way to a variety of alternatives such as tablets and smartphones that are likely to be owned and managed by the user. 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 access applications such as Citrix Receiver or VMware View Client.

Although these access and collaboration applications are available on a variety of platforms, careful design is required to maintain the UI and QoS user experience both on the campus and in remote locations over both wired and wireless networks.

General User-Experience and Management Challenges

The problems discussed so far add up to a seriously flawed user experience, especially at times of peak load when optimization is most needed.

These problems also impair the ability of administrators to manage their networks and communications systems. Existing tools, techniques, and experience are no longer applicable, and new tools are not available.

An architecture is needed that provides a user experience equivalent to that with a traditional desktop, while preserving IT investment in the existing network.

Test Results

In a system test, on an idle virtual desktop, CPU and network utilization averaged 6 percent and 20 Kbps, respectively, as illustrated in Figure 4.

CPU and network utilization increases many times when point-to-point communication occurs on a virtual desktop. As illustrated in Figure 5, in the test, CPU utilization increased to 51 percent, and network utilization increased to 44 Mbps.

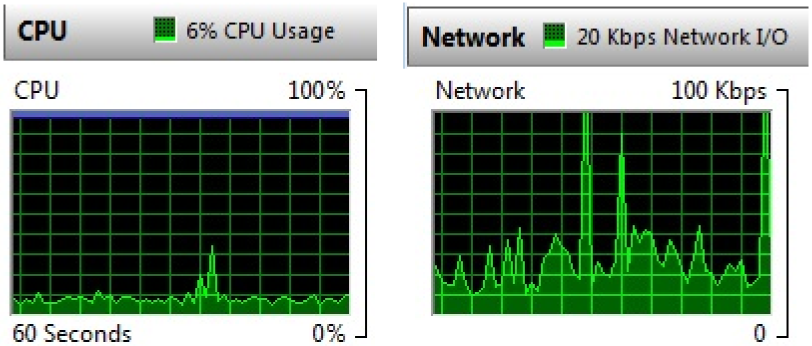

Figure 4 Resource Monitor for an Idle Virtual Desktop: CPU Utilization = 6 Percent and Network Utilization = 20 Kbps

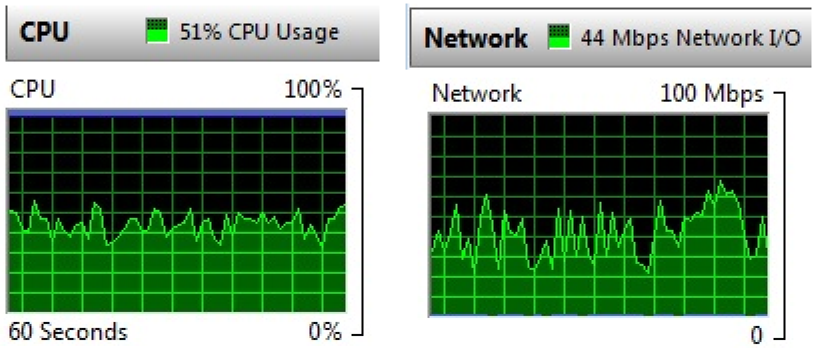

Figure 5 Resource Monitor for a Virtual Desktop During a Point-to-Point Call: CPU Utilization = 51 Percent and Network Utilization = 44 Mbps

Cisco Virtual Workspace (VXI) Smart Solution

The Cisco Virtual Workspace (VXI) Smart Solution adds intelligence in the end clients as well as in the data center. The 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 an excellent experience for the user. Interaction with other applications works as if everything were running on the virtual desktop.

This approach addresses the problems discussed earlier in a way that maintains the value of the existing investment in network and communications while merging transparently with the virtual desktop deployment.

The voice and video capabilities in virtual environments are delivered through the Cisco Jabber™ platform enabled by the Cisco Virtualization Experience Media Engine (VXME).

The Cisco Jabber application is a unified communications client that provides voice and video communications capabilities as well as other services. It runs on several different platforms and provides the same user experience on every platform. It can be used in soft-phone mode as a standalone application that terminates media or in desk-phone control mode to control the user's Cisco IP Phone using computer telephony integration (CTI). In desk-phone control mode, the Cisco Unified IP Phone originates and terminates media.

Cisco VXME is an add-on software module that is currently supported in the Cisco Virtualization Client (VXC) 6215 thin client, and will be supported on third-party Linux and Microsoft Windows thin clients and Microsoft Windows PCs. The Cisco Jabber application controls Cisco VXME when the Cisco Jabber application runs on a virtual desktop that is accessed through a thin client or Microsoft Windows PC equipped with Cisco VXME. Under Cisco Jabber control, Cisco VXME registers with Cisco Unified Communications Manager and originates and terminates media. In these virtualized scenarios, Cisco VXME acts as an extension of the Cisco Jabber application. The Cisco Jabber application automatically recognizes that it is running on a virtual desktop, and if Cisco VXME is installed and available on the endpoint, it establishes a communication channel with Cisco VXME to offload voice and video processing to the endpoint.

Figure 6 shows the interaction between the Cisco Jabber application and Cisco VXME.

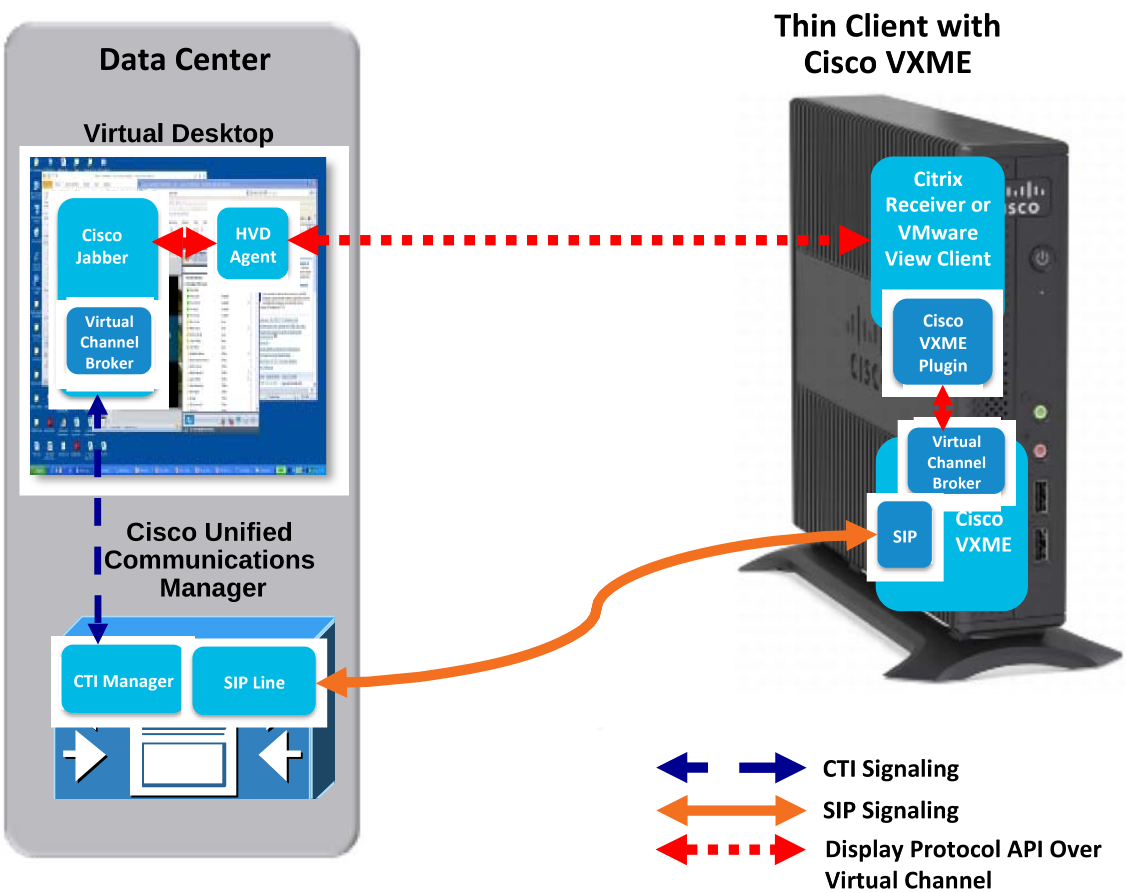

Figure 6 Cisco Jabber Application and Cisco VXME on a Thin Client

Next-Generation Unified Device with VDI and Collaboration

The thin-client solution provides the next-generation virtual workspace, with desktop, phone, and dedicated video capabilities combined into one unified device.

The following steps show how the Cisco Jabber application establishes an interactive session:

1. The Cisco Jabber application establishes a communications channel with Cisco VXME running in the thin client. Cisco VXME registers with Cisco Unified Communications Manager.
2. User 1 initiates a call through the Cisco Jabber interface running in virtual desktop 1.
3. The Cisco Jabber application in virtual desktop 1 connects with Cisco Unified Communications Manager.

4. Cisco Unified Communications Manager establishes a call between the originating and terminating thin-client endpoints.
5. Call media flows directly between the two Cisco VXME instances in the thin clients.

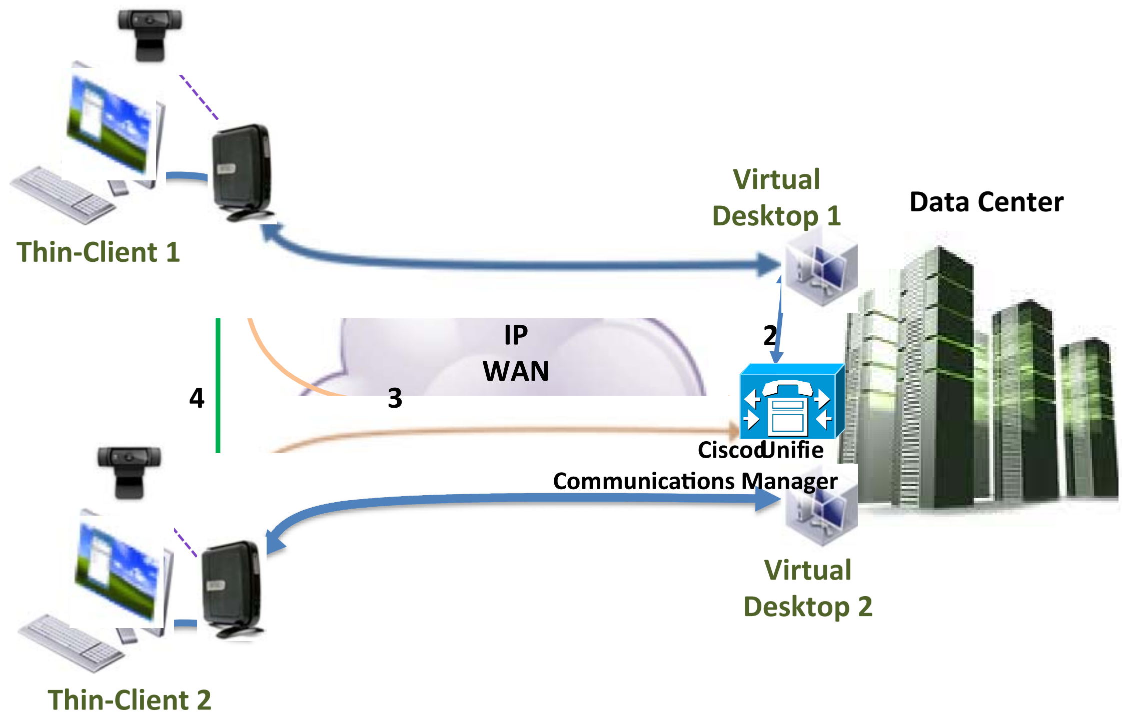

Figure 7 Control and Media Flow between Two Thin Clients

Third-Party Devices

In addition to thin clients, the Cisco Virtual Workspace (VXI) Smart Solution includes endpoints on various platforms. These include desktops and laptops running Microsoft Windows, Apple Macs, Apple iPads, and various other tablets and smartphones that are now becoming popular in organizations.

Users can access their virtual desktops and virtualized applications using such applications as Citrix Receiver and VMware View Client available on these platforms. The Cisco Jabber application is also available on these platforms and provides collaboration capabilities to handle all media directly in the endpoint, avoiding any hairpinning through the data center. For platforms that do not currently support Cisco VXME, users should run the Cisco Jabber application on the endpoints directly.

How the Cisco Solution Addresses the Rich-Media Challenges

The Cisco solution addresses the problems described earlier. In general, the solution enables the technology and tools used for voice and video today to work equally well with unified communications in VDI deployments.

This approach contrasts with attempts to optimize unified communications and media within VDI. Rather than solving the problems, optimization just incrementally improves the existing architecture to mitigate some of the problems to some extent. In addition, this approach requires unique and new sets of tools to provision and manage deployments.

The sections that follow describe the Cisco solution in general terms. Refer to individual product data sheets and the Cisco Validated Designs for the Cisco Virtual Workspace (VXI) Smart Solution referenced later in this document for complete information.

Optimal 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 capability of the network to optimally route traffic, reducing latency.

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

In addition, the total number of encoding and decoding processes required for a video call are 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 the Cisco VXME system, each video and audio stream is encoded only once by the sender and decoded once by the receiver.

Optimal Bandwidth Consumption

Routing media directly from one endpoint to the other instead of hairpinning the traffic through the data center reduces bandwidth consumption. For example, if one Cisco VXME endpoint is used to call another in the same branch office, the media flows between the endpoints in the branch office with no impact on the WAN link.

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

Restoration of Network Optimization

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

Full Call Admission Control

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

QoS, Voice VLAN, and Medianet Support

Cisco VXME supports voice VLAN segregation. It uses Cisco Discovery Protocol to identify the voice and access VLANs, if they are configured, on the access switch port, and it places media in the voice VLAN and all other traffic in the access VLAN. Cisco VXME obtains configured Differentiated Services Code Point (DSCP) values from Cisco Unified Communications Manager after registration.

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

Enhanced 911 Service

Cisco VXME uses Cisco Discovery Protocol to identify itself to the access switch. This feature enables the network to geographically locate the thin client and the user.

Decreased CPU Load in the Data Center

Only the user-interface elements run in the data center, dramatically reducing CPU use. All media processing is performed on the endpoint, eliminating that load from the HVD used with that endpoint.

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 as well as decreased load.

Restoration 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 the customer site. The Cisco solution enables the existing network and unified communications management tools to be used with a VDI deployment.

Excellent Experience for Users and Administrators

With the features described here, the Cisco solution provides an excellent unified communications and rich-media experience. For IT administrators, the Cisco Virtual Workspace (VXI) Smart Solution simplifies management, using the same tools and techniques that are used to manage the network and unified communications systems today.

The Cisco Virtual Workspace (VXI) Smart Solution also preserves existing investments in equipment, tools, and training. The solution can be placed transparently into an existing Cisco Unified Communications deployment, complete with QoS, Cisco Wide Area Application Services (WAAS), voice VLAN, and Enhanced 911 (E-911) support.

Test Results Using the Cisco Solution

The thin-client solution dramatically reduces CPU and network utilization, as illustrated in Figure 8. During lab testing, CPU utilization dropped to 10 percent, and network utilization dropped to 44 Kbps.

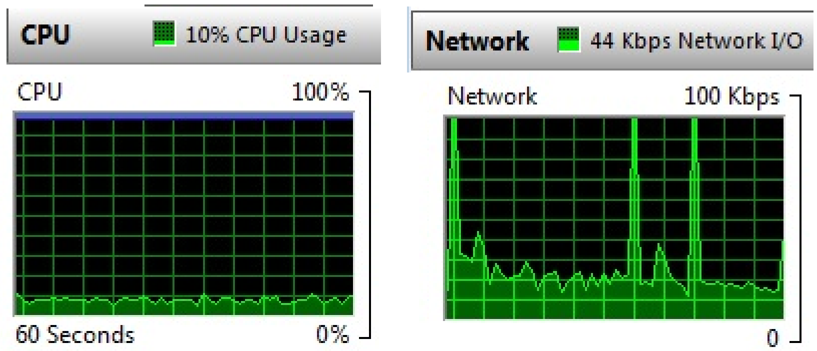

Figure 8 Resource Monitor During a Point-to-Point Call: CPU Utilization = 10 Percent and Network Utilization = 44 Kbps

Table 1 compares CPU and network utilization for a virtual desktop in three states: idle, with a call in a VDI environment, and with a call with the Cisco solution.

Table 1 CPU and Network Utilization for a Virtual Desktop

	Idle	Call in a VDI Environment	Call with Cisco Solution
CPU	6%	51%	10%
Network	20 Kbps	44 Mbps	0.044 Mbps

Conclusion

The Cisco Virtual Workspace (VXI) Smart Solution includes a family of products and proven technology to solve the point-to-point communications challenges in virtual desktop environments. The Cisco Virtual Workspace (VXI) Smart Solution is designed to create an uncompromised user experience and to reduce the load on network and data center resources.

The solution also preserves existing investments in both network and communications infrastructure. Many customers have made investments in not only the infrastructure itself, but also in management tools, diagnostic tools, training, and experienced staff. The solution enables customers to merge successful network and Cisco Unified Communications deployments with VDI.

Cisco, through ongoing innovative developments and partnership with display protocol providers (Citrix, VMware, and Microsoft), has created a virtualization-aware solution that understands user needs and brings the user experience closer to that on a traditional desktop computer.

For More Information

Cisco Virtual Workspace (VXI) Smart Solution: http://www.cisco.com/go/vxi

Cisco VXME: http://www.cisco.com/go/vxc

Cisco Medianet architecture: http://www.cisco.com/go/medianet

Cisco Jabber platform: http://www.cisco.com/go/jabber

Printed in USA 05/12