Preparing Your IP network for High Definition Video Conferencing

White Paper

Global Services

April 2007

POLYCOM®

TOGETHER, GREAT THINGS HAPPEN.

Preparing Your IP network for High Definition Video Conferencing

Table of Contents

POLYCOM®

TOGETHER, GREAT THINGS HAPPEN.

Preparing Your IP network for High Definition Video Conferencing

1.0 Overview

Leading enterprises today recognize the business efficiency and productivity that comes from the use of video conferencing as a key communications tool. Using video to quickly connect executives and teams across geographic boundaries provides a closer working relationship, saves time and travel, and allows enterprises to use diverse talent from different parts of the globe on joint tasks.

The recent introduction of High Definition (HD) Video Conferencing has increased the quality of the video experience substantially. However, deploying high definition video provides unique challenges to the IP network supporting it. This white paper gives an overview of the impact of HD video conferencing on a converged IP network, and suggests approaches for managing that impact to support a high quality video conferencing service.

An understanding of IP network design and deployment is helpful in understanding this guide, as is a general knowledge of IP network deployment (switching, routing, bandwidths, error mechanisms, etc.).

2.0 Video Conferencing Bandwidth Demand

The most significant difference between traditional H.323 video conferencing and High Definition (HD) video conferencing is the increased bandwidth demand. Whereas a traditional video conferencing connection might use 384Kbps or 512Kbps of transport bandwidth, the HD systems can use as much as 4 Mbps of audio and video transport. To understand the network impact, IP overhead has to be added onto these values.

Table 1 below shows typical transport rates for video conferencing and for HD video conferencing. The second column of the table shows the demand placed on the network using Ethernet technology and the third column shows bandwidth used on an ATM link.

Table 1 - Video Conferencing Bandwidth Rates

	Rate	Ethernet	ATM
Video Conferencing	192K	230K	240K
	384K	460K	480K
	512K	614K	640K
	768K	920K	960K
High Definition Video Conferencing	1024K	1.2M	1.3M
	1472K	1.8M	1.9M
	1920K	2.3M	2.4M
	3840K	4.6M	4.8M
	4096K	4.9M	5.1M

The values shown in the Ethernet and ATM columns of Table 1 are higher than the video conferencing rate because they include the overhead of the IP protocol, and in the case of ATM the further overhead of ATM cells. These larger bandwidth values should be used to understand the impact of HD video conferencing on the bandwidth of WAN and LAN links in a converged network.

POLYCOM
TOGETHER, GREAT THINGS HAPPEN.

Bandwidth and QoS

Bandwidth use is an integral part of QoS. Sufficient bandwidth must be in place on each link to carry the expected real-time traffic. So the first question is what is the expected traffic? It is important to analyze expected demand so that proper bandwidth planning can be done to support video conferencing on the network links.

Call density and patterns must be estimated based on expected usage. If video conferencing will primarily be taking place from video conferencing rooms, then demand can be estimated by making assumptions about room utilization. Call destinations will have to be estimated using knowledge about the business and likely call patterns for users. Create a spreadsheet that estimates the amount of simultaneous videoconferences from each major location to other locations during the busiest hours of the day in the enterprise network.

Table 2 and Figure 1 below show the results of such an analysis. In this example an enterprise with eight offices is connected by a common service provider with an MPLS mesh connection. A demand spreadsheet was created to estimate the number of simultaneous video conferencing calls to each site during the busy hours of the business (Table 2.) Note that the busy hour for a particular office may not coincide with the busy hour for another office depending on time zones and the nature of the business.

Table 2 - HD Video Conferencing Demand Example

	Atlanta	*Chicago*	*Dallas*	*Phoenix*	*San Jose*	*Boston*	*London*	*Tokyo*
Atlanta	0	1	1	1	1	1	1	0
Chicago	1	0	1	1	1	1	1	0
Dallas	1	1	0	1	1	1	1	1
Phoenix	1	1	1	0	1	1	1	1
San Jose	1	1	1	1	0	1	1	1
Boston	1	1	1	1	1	0	1	0
London	1	1	1	1	1	1	0	0
Tokyo	0	0	1	1	1	0	0	0
Total	6	6	7	7	7	6	6	3

The total count of HD video conferencing calls was then multiplied by 1920 Kbps plus 20% overhead to generate the video conferencing demand shown in Figure 1. For this example, the MPLS link to each office must be sufficiently large to support these levels of video conferencing, and simultaneously support the data (and possibly voice) traffic of the office as well.

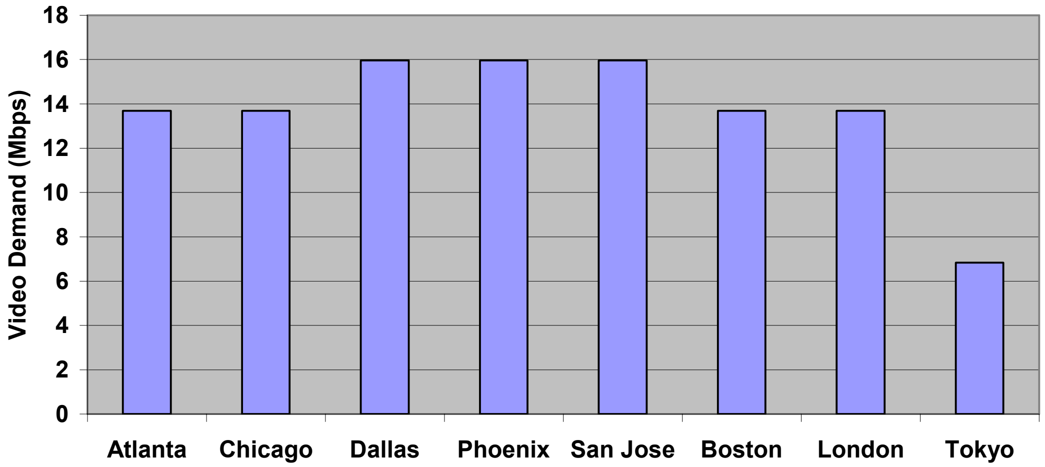

Figure 1 - Video Conferencing Bandwidth Demand

Bridge (MCU) Bandwidth Demand

Key infrastructure components of the video conferencing system need special consideration. First consider the video conferencing bridge (or MCU). If six video conferencing endpoints are engaged in a conference call, all six endpoints have established a full duplex connection to the bridge. The bridge network connection must be able to sustain the maximum number of endpoints that will be in all simultaneous conference calls. Thus the bridge should be placed near the core of the network where bandwidth is more plentiful. Furthermore, the bridge should be placed in the facility where the highest percentage of conference call users reside to minimize the WAN traffic required to support these conference calls.

Each client that connects to the bridge will have a traffic stream flowing from the client to the bridge at the bandwidth negotiated for that video conference. If each client has negotiated a 1.9 Mbps bandwidth call, and there are 6 clients, the bridge will be supporting 1.9 Mbps x 6 or 11.5 Mbps of traffic. When we add the 20% additional bandwidth required for IP packet overhead, this now comes to 13.8 Mbps.

Some video conferencing endpoints also support a built-in multipoint conferencing mode. If a video conferencing endpoint is acting as a bridge for a small conference, there will be a proportionate increase in the bandwidth to that client. A 4-person conference using one of the 4 clients as a bridge will generate three full duplex streams to the client acting as a bridge. The other three clients will see a single full-duplex stream.

3.0 Available Bandwidth

Once the bandwidth demand has been calculated, an evaluation of existing network bandwidth and utilization is required to determine if there are sufficient resources to support the new real-time load. Each link of the network needs to have sufficient bandwidth to support the voice and video traffic expected, plus the existing data applications that use those same connections.

Although this sounds like a daunting task, in practice it usually means evaluating the wide area network links, the backbone connections of the bridge, and client connections where there may be 10Mbps Ethernet or shared Ethernet connections. Often much of the infrastructure of an enterprise does not need detailed bandwidth analysis, just these key elements.

Client connections should all be upgraded to 100 Mbps full duplex if possible. If the video conferencing endpoint does not support full-duplex operation, it is preferable to run at 100Mbps half-duplex. If the endpoint supports full duplex, but does not support 100 Mbps, it is preferable to run at 10 Mbps full duplex.

3.1 Converged Network Links

Converged network links are those where both data traffic and real-time (voice or video) traffic are being supported concurrently. There are two parameters to consider when evaluating the WAN links. First, the expected voice and video (real-time) load should never exceed 33% of the link capacity. Priority-based QoS mechanisms begin to lose their effectiveness at this level. Running with more than 33% high-priority traffic means that the traffic starts to compete with itself, and reliable delivery is compromised.[1,2]

The second parameter is the total bandwidth utilization of the link, including the real-time components and the data components. It is straight forward to determine the bandwidth demand of the real-time applications, but determining the needs of data applications is much more difficult. Data applications are very bursty, and when many of those applications are aggregated on a link their profile is still very bursty. Data applications depend on bandwidth overhead to get good performance. If the bandwidth of a link is limited to the average consumption of the data applications, the applications themselves slow down, creating user frustration and reduced productivity.

3.2 Dedicated Network Links

Dedicated network links carry only real-time traffic. With these links, the 33% real-time traffic limit does not apply because all the traffic is well-behaved real-time traffic. For links dedicated to voice traffic only, very high utilizations are possible. For traffic that includes video conferencing, which is burstier than voice, a limit of 70% utilization should be observed. High speed links (100 Mbps and higher) can be utilized up to 80%, since the number of streams is much higher and the burstiness of an individual stream has less impact on the link.

This increased utilization does not apply to the classes of service provided by an MPLS provider, unless a separate access link is used. The access link is usually the critical link for QoS, because it represents the lowest speed link in the end-to-end path of the video conferencing. If a single access link is used to support both real-time and data traffic, the 33% limitation applies. If a separate link is deployed dedicated to voice and/or video conferencing, the higher utilization values mentioned in the previous paragraph can be used.

---

[1] "Implementing QoS Solutions for H.323 Video Conferencing over IP", Cisco Systems, http://www.cisco.com/warp/public/105/video-qos.html

[2] "Economics of QoS on WAN Access Lines", Bartlett , Sevcik & Moore, October 2004, Business Communications Review Magazine

4.0 Demand Management

If the network analysis determines that there is insufficient bandwidth on critical links, the enterprise has a few options to resolve the conflict:

- Bandwidth upgrade
- Reduce voice or video conferencing demand
- Compression / Application Acceleration Appliances

Bandwidth Upgrade - A bandwidth upgrade is always possible and may be the only solution if insufficient bandwidth is available to carry the required voice or video conferencing load.

Limit Conferencing Demand - The second option is to limit the video conferencing demand. This can be done in a number of ways. First, the bandwidth used by video conferencing calls can be limited. Better HD video quality can be obtained at 4 Mbps but quite good quality can be obtained at 2 Mbps, and even at 1 Mbps. Some testing of the different video quality levels may reveal that a lower bandwidth is sufficient for those offices where limited bandwidth is available.

A second way to reduce demand is to manage call volume so that a limited number of calls can occur simultaneously across each link. If a remote office has three video conferencing units, but the bandwidth of the link can only support two simultaneous calls, a scheduling policy can be put in place to insure that only two systems are being used concurrently. The simplest case of this policy is to insure that the remote office only has the number of video conferencing endpoints that the link can support.

The voice or video conferencing gatekeeper can also be used to help manage bandwidth utilization. The gatekeeper can be assigned a maximum bandwidth available between groups of endpoints, which relate to the topology of the network. The gatekeeper will then only allow calls across that link up to the available real-time bandwidth allocated to that link. The bandwidth value given to the gatekeeper is the maximum amount of real-time traffic allowed on that link, not the link capacity. Once the link utilization reaches this maximum amount, the gatekeeper will refuse additional call requests.

Compression and Application Acceleration – One more option is to reduce the existing data traffic. A new class of data appliances is available that use various tricks to both reduce data traffic and increase application performance simultaneously. These appliances use compression, caching, TCP termination, transparent turns reduction and other techniques to accomplish their goals. There is a bit of work to determine which approach best suits the data streams employed for each situation, but these appliances can often make room on the link so that video conferencing or voice traffic can be introduced without requiring a bandwidth upgrade.[3]

---

[3] Further reading on Application Acceleration: "Field Guide to Application Delivery Systems", Sevcik & Wetzel, September 2006, Report on NetForecast website

5.0 Real Time Traffic

All video conferencing traffic is real-time traffic, and needs to be given proper Quality of Service (QoS) support both in the local area network (LAN) as well as the Wide Area Network (WAN). High Definition Video Conferencing traffic has the same needs as standard video conferencing traffic in this respect, just with higher bandwidths. For a detailed discussion of real-time traffic support, the reader is encouraged to obtain a copy of "Supporting Real-time Traffic, Preparing Your IP Network for Video Conferencing", a separate white paper from Polycom.

6.0 WAN Vendors and Technologies

Connecting enterprise locations is often done with the help of a Wide Area Network (WAN) service provider. The service provider links become an integral part of the enterprise network, and of the real-time traffic support.

HD Video Conferencing requires real-time traffic support at high data rates for a successful implementation. It is critical that the specifications for the wide area network connection meet the requirements of the expected HD video conferencing demand both for real-time support (low loss, low latency, low jitter) as well as for the bandwidth required.

Many service providers today support this type of demand using Multi Protocol Label Switching (MPLS) technology. MPLS allows the service provider to configure the appropriate bandwidth and to offer classes of service to support the needs of high bandwidth real-time flows. Polycom recommends using an MPLS WAN service to support HD video conferencing streams.

Metro Ethernet service providers offer a layer-2 technology to support high bandwidth flows. Metro Ethernet providers offer service within a metropolitan area, but often not over the longer geographic distances such as across the US or international connections. Layer-2 connectivity can successfully support HD video conferencing if the bandwidth and QoS parameters are properly specified.

Video conferencing traffic should be carried on either MPLS or Layer 2 technologies at a class 4 priority. This translates into an AF41 marking for DiffServ environments (e.g. MPLS) or an IEEE 802.1p marking of 4 for layer 2 environments. [4]

The WAN connections should be verified after installation to ensure that packets are being marked correctly, and that they are being given the appropriate priority through the WAN. Synthetic network test tools are useful for verifying the QoS deployment.

7.0 Real-Time over VPNs or the Internet

Many small to medium sized enterprises today are taking advantage of Virtual Private Networks (VPNs) to connect their geographically distributed offices. VPNs create an encrypted tunnel through the public Internet. The advantage of a VPN is that the cost is often much less than a

[4] "Configuration Guidelines for DiffServ Service Classes", Babiarz, Chan, Baker, IETF, http://tools.ietf.org/html/draft-ietf-tsvwg-diffserv-service-classes-02

dedicated connection. Enterprise VPNs come in two flavors, those that connect two offices through a single WAN provider, and those that use the open Internet, so they may use more than one service provider and their associated peering points.

Carrying real-time traffic through these VPNs is as risky because there is usually no QoS capability offered in the VPN connection. Quality may be good when a single service provider is providing connectivity at both ends, but again no guarantees about bandwidth, loss or jitter are available. Some enterprises use this approach because the value of a voice call or video conference to a remote manufacturing plant or development center justifies the risk, and because the users can be tolerant of failures. If the quality expectation is high, such as support of management staff meetings, sales updates, presenting to clients and other high visibility uses, than the risk of quality degradation and call failure may be too high to use a VPN.

Using the Internet for real-time traffic carries the same risks as a VPN, with less control. When connecting to another party via the Internet, multiple carriers may be involved, and the user has no control over how the call is routed. Hot-potato routing algorithms often cause traffic flowing one direction will take a different route than traffic flowing in the reverse direction. Educational and research institutions have had some luck using the Internet where they have very high bandwidth connections, but the risk of having a poor quality connection is high.

8.0 Conclusions

Deploying high definition video conferencing creates a new and different challenge for the IP-network team. A successful deployment requires careful attention to the requirements of real-time traffic. If each of the steps outlined in this document are addressed and then incorporated into the daily operations of the network, the enterprise can not only have a successful deployment, but also maintain a high quality service over the IP-network through the inevitable changes in applications, locations, and the network itself.