USENIX

THE ADVANCED COMPUTING SYSTEMS ASSOCIATION

Supporting H.323 Video and Voice
in an Enterprise Network

Randal Abler and Gail Wells
Georgia Institute of Technology

Supporting H.323 Video and Voice in an Enterprise Network

Randal Abler, Gail Wells
Communications Systems Center
School of Electrical and Computer Engineering
Georgia Institute of Technology
Randal.Abler@csc.gatech.edu
Gail.Wells@csc.gatech.edu

Abstract

H.323 is a relatively new standard for video and voice transmission that specifies using IP packets as the transport. This opens the possibility of adding an inexpensive camera to a modern desktop, and allowing two way video-conferencing between any offices so equipped. While the H.323 specification addresses LAN networks, can H.323 be used in a WAN environment? What characteristics are necessary to support H.323 in the LAN and WAN networks, and what is the impact of H.323 on other traffic in the network? This paper attempts to outline the impact of running H.323 on a network and lay out some guidelines that should be useful for accommodating H.323 on both local and wide area networks.

Motivation

This paper describes results from ongoing work to examine and potentially deploy H.323 in Georgia, done in conjunction with the State of Georgia Department of Administrative Services (DOAS), and with the support of the National Science Foundation under grant NCR-9613986. Currently Georgia operates an independent H.320 video network known as GSAMS, based on dedicated leased lines. This network currently contains about 350 nodes throughout the State of Georgia. H.320 is a commonly deployed video-conferencing standard based on switched circuit such as ISDN connections, or dedicated leased lines such as T1 circuits. Unless technology such as ATM is deployed, these connections cannot simultaneously support any other type of use while a H.320 video call is in session. Therefore expensive T1 circuits are dedicated to occasional video calls. Since H.323 uses IP networking as its transport, it can co-exist with other uses of the telecommunications circuits.

DOAS also operates a data network that supports IP, IPX and SNA for sites across the State of Georgia, with IP connectivity to the Internet. Multiple state agencies use this network for a wide variety of tasks. Therefore a tentative infrastructure for supporting H.323 already exists. It is inevitable that some H.323 use will occur across this network unless explicit steps are taken to restrict it.

Rather than taking a restrictive approach, we are investigating the feasibility of deploying H.323 over data network, and may migrate some of the H.320 sites to H.323. A dedicated test setup involving two locations, interconnected with commercial Frame Relay and a private ATM network, is being used to both technically evaluate the options and to provides hands on experience for network designers and operators, as well as demonstration capability for interested customers.

H.323 Overview

H.323 [H.323] is an ITU standard that provides for the transport of real-time voice and video over IP networks. H.323 specifies H.261, and optionally H.263 as the video compression technique, similar to H.320. H.323 can be used for voice only transport as in *Voice-Over-IP*, or *Internet Telephony*. (This does not mean that all Internet Telephony applications are H.323 compliant.) H.323 defines clients, Gateways, and Gatekeepers. A client is any device capable of sending H.323 video or audio. T.120 extensions allow data sharing over an H.323 connection. The application sharing in Microsoft's NetMeeting, which allows a remote user to see the contents of any application window, is a good example of how T.120 can be used. File transfers can also be implemented.

A Gateway is a device that translates H.323 into some other voice/video transport. Two common types of gateways are H.323<->Analog Telephone line (POTS), and H.323<->H.320. H.320 is an older standard for video conferencing. It is often implemented over ISDN, or sometimes over leased lines. In using a gateway for an outgoing call, the outgoing call is placed to the Gateway's IP address, and additionally a phone number (for POTS or ISDN) is passed to the gateway. The Gateway attempts to call the designated phone number, and connects audio, and video if appropriate. In an incoming call, the gateway can either be hardwired to contact a specific IP address, or provide

some mechanism of selecting an address. In an incoming POTS call, for example, a secondary dial tone may be used which acts like an "extension", or a PBX style interface into the public telephone network may allow direct mapping of phone numbers to H.323 client IP addresses.

Gatekeepers allow network administrators to place limits on H.323 clients. This can be used to limit the maximum bandwidth a client can use, or completely disable clients. H.323 clients broadcast for a local Gatekeeper, and if none is found, enable themselves. If a gatekeeper is found, the client will check with the gatekeeper for appropriate restrictions. Gatekeepers are appropriate for networks which need the control, but the administrative burden may be prohibitive in a small or sparsely used H.323 network.

Normal H.323 clients are point-to-point devices, i.e., you call one person. To have a conference with multiple people, an multi-point conferencing unit (MCU) is necessary. The MCU allows multiple people to participate. MCU's may operate in different modes. In a voice-activated mode, the image of whomever is talking is transmitted to the other participants. In "continuos presence" mode, participants receive a split screen image, often split into quarters, allowing four remote sites to be displayed.

MCU's, gatekeepers, and gateways implementations may be combined into one product. MCU's and gatekeepers can be implemented as software only additions to a computer. For production use a dedicated computer, typically running Microsoft NT, is preferred. Gateways can be implemented as dedicated devices or as add in card(s) in a computer, PBX, or router.

While H.323 is targeted at video-conferencing, there are other video over IP techniques. H.323 is essentially a point to point protocol, much as the existing telephone network. If your goal is to broadcast a meeting, or to have some form of video library available on demand over an IP network, then other solutions are better suited. A widely used product suite to address these needs in the Real-Networks products. These products use a proprietary encoding. Several other products exist which offer similar capabilities. If broadcast of video content is used within your network, then it is worth looking into enabling multicast IP routing. In a non-multicast IP network, a separate video stream exists from the server to each client. Multicast IP allows multiple clients to monitor one common video packet stream, reducing the load on the server and saving network bandwidth.

Test Lab configuration

Figure 1 shows the connectivity of the evaluation and

Figure 1 - Video Test Network
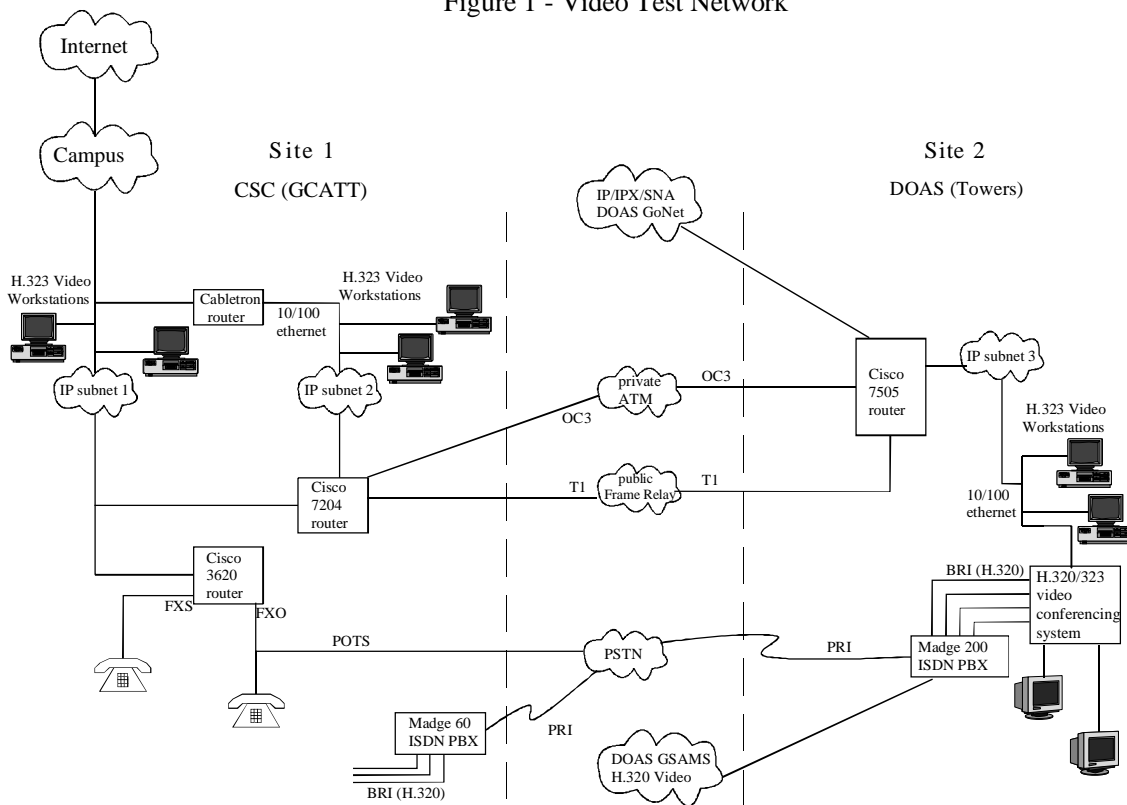
testing facilities. Both sites are located in the Atlanta area, separated by approximately 4 miles. One site is located on the Georgia Tech campus and the other site is located in a state office building in the downtown area. The sites are interconnected with a T1 Frame Relay link, and with an OC3c ATM link which runs via an OC12c trunk. The T1 link is provisioned with multiple virtual circuits to allow parallel testing of multiple transports. This baseline configuration is modified as needed to add network monitors, load generators, and additional network components such as additional routers to create more IP hops.

Each site has three or more PC's dedicated to H.323 video testing. This allows multiple concurrent session to be run. Future testing will allow for the addition of Gatekeepers, Gateways, and MCUs from various vendors. The network connectivity supports testing local area network configurations using shared 10Mbps Ethernet, switched 10Mbps Ethernet, as well as shared and switched 100Mbps Ethernet. The ATM interconnectivity between sites allows testing in a configuration similar to what might be found in a high-speed backbone (campus) network, with multiple router hops and switching points. By using a circuit emulation line over the ATM network, leased line connections from 56kbps to T1 speeds are being tested. Finally, dial-in servers (not shown) allow modem-based and ISDN connections to be tested.

Typical H.323 configuration

Configuring a PC to work as an H.323 station requires both the correct hardware and software. Getting a good working combination is not as simple as might be desired. The computer audio system should support full duplex sound, which may require an upgraded sound system even in modern PCs. Echo cancellation is not good in many situations and often the best

solution is to use headsets instead of speakers to eliminate the feedback. Video inputs to a desktop can use the video standard NTSC into an encoder card, or a parallel port or USB connected camera. The parallel port or USB cameras are easy to connect, but often are more limited in the image size and quality. An NTSC encoder card, plus an NTSC camera is a more expensive combination, but offers better quality, and can be used to connect to a VCR or other NTSC output device. Most of the desktop H.323 solutions rely on the system CPU to implement the audio and video compression and decompression. The CPU speed becomes a limiting factor in sending good quality video. While usable video can be achieved with a slow CPU if a small image is sent, and a low bandwidth is specified, good performance requires the fastest of modern PC's. Slower computers will also have a more noticeable problem with audio delay. Audio delay causes two problems. The H.323 systems we tested do not delay the video and audio the same amount. This causes a lip-sync issue. Also, when the audio delay is noticeable, the dynamics of human conversation become difficult, and two people will often start speaking at once.

Initial tests were done with 200MHz Pentium systems, which could not keep up with medium quality audio and video while implementing software compression. We have found 333MHz Pentium II systems to be adequate, and achieve excellent performance with a dual 450MHz processor Dell Precision 610. The multiple processors are useful when running application sharing. Single processor machines often experience temporary frozen video, and dropped audio, when an application program uses significant CPU time, such as when an application starts. An alternative to using a faster machine is to invest in a video encoder card with hardware compression. The system CPU still implements the decompression of the incoming video, and given the cost of hardware based compression
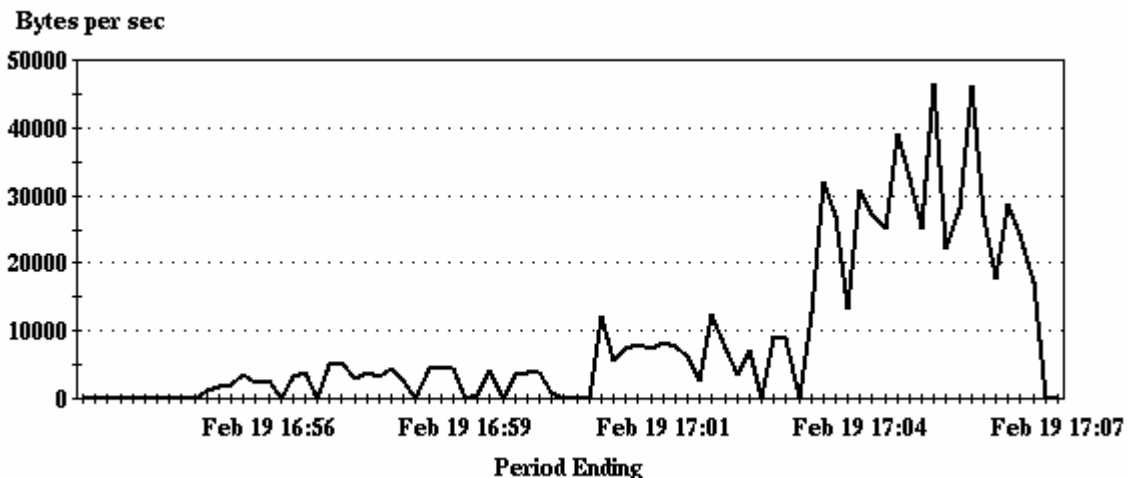

Figure 2 – Transmitter Bandwidth

cards, $800 and up, it may be worthwhile investing in a faster computer since a faster machine accelerates other non-video applications.

Dedicated H.323 systems often include a hardware compression card with multiple inputs, allowing you to switch between multiple cameras such as a document camera, etc. Turnkey systems are appearing which support both high quality H.323 and H.320 384kbps ISDN calls.

While measuring packet loss and latency are objective measurements, video and audio quality are subjective. Video quality is a function of frame rate, consistency of the frame rate, image resolution, image noise, and compression artifacts. Noise, echo, and clarity similarly effect audio. Additionally, when audio and video are compressed, transmitted, and decompressed on the remote end, the video and audio may get out of synchronization. This is most noticeable when a person's voice does not match their lip movements, and therefore is commonly called lip-sync.

H.323 bandwidth & Packet Loss

Many factors effect the bandwidth use, such as the image size, hardware capabilities, the software product used, compression algorithm used, what network speed is configured into the software, and the image complexity and dynamic nature of the imagery. RMON probes are being used in our test configuration to provide traffic plots of the traffic generated from H.323 sites. Baseline tests show that an active desktop H.323 session can use between 20kbps for a voice only conversation to over 256kbps for a high quality desktop videoconference. Dedicated H.323 systems with hardware compression may use 2 to 3 times this bandwidth.

Most of the H.323 clients allows the bandwidth to be adjusted by the user, subject to a maximum imposed by a gatekeeper if present. Depending on the product, this bandwidth may be set in KBPS, or in terms of your connection type, such as "28.8 Modem", "ISDN", "LAN". Using the right setting is important in getting the optimal connection quality. Higher bandwidth settings allow for faster image updates. This works as long as the bandwidth is really available on the network. While it is a subjective judgement, we did not find the image quality useful for 14.4 Modem or 28.8 Modem settings. These setting are effective for audio only, and for application sharing without video.

When the H.323 configuration is set for more bandwidth than is available on the network, the resulting video and audio quality suffers drastically as the data packets are lost. This setting is often configured when the software is initially installed, and therefore gets overlooked as a parameter that may need to change on a per call basis. When calling between two sites on an uncongested LAN or Campus network, the LAN setting will give the best quality video. Using the same "LAN" setting to call a second site that is reached through a remote link such as a loaded T1 link, or an ISDN link, will result in a broken video image and audio that suffers from frequent drop-outs.

To illustrate this an H.323 call was set up between two LANs connected by an ISDN line. On the "transmitter" LAN, a high-speed desktop was connected to a VCR video source, and an RMON probe attached to that LAN. The probe output is shown in Figure 2. On the "receiver" LAN, another RMON probe was set up to measure the traffic after passing through the ISDN line. This is shown in Figure 3. While video was also sent in the reverse direction, the RMON probes were not configured to capture the reverse direction traffic. Both plots cover the same time interval, covering three
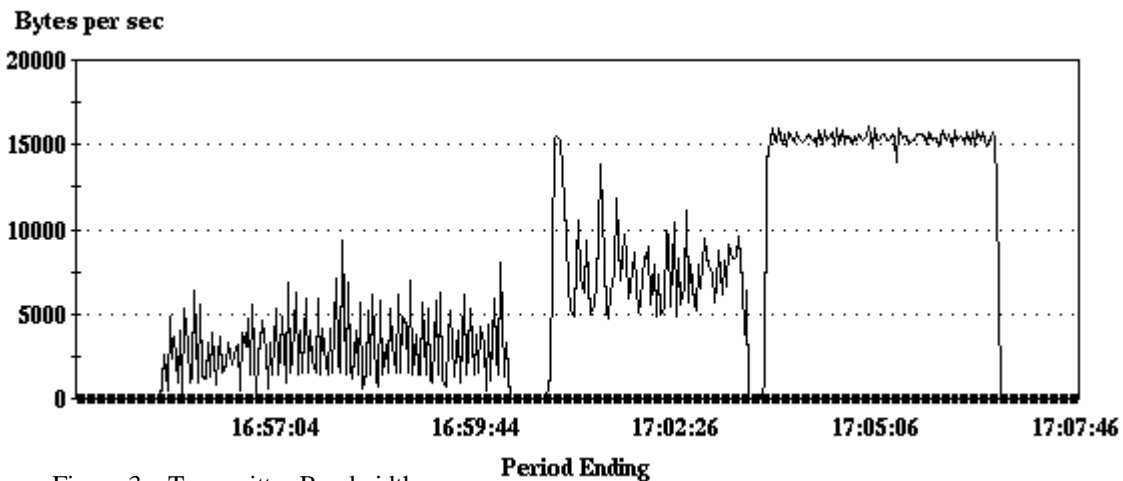

Figure 3 – Transmitter Bandwidth

separate calls. In the first call, shown on the left, the H.323 client was configured to use "28.8 modem" bandwidth. This resulted in approximately 3000 bytes per second, or 24kbps, of traffic. Audio and video quality is good, but the frame rate of approximately 1 frame every 2 seconds did not provide a sense of motion. The second call was configured for "ISDN" bandwidth. This resulted in an average of about 7500 bytes per second, or 60kbps. Video and audio quality are good, and the frame rate of approximately 3-5 frames per second provides a sense of motion and continuity. The third call was configured for LAN connectivity. The transmitter sent approximately 30,000 per second, or 240kbps. This was about twice the capacity of the ISDN line, which can be seen on the receiver plot which shows 15000 bytes per second (120kpbs). The video image was unrecognizable due to partial updates, and the audio was very garbled.

WAN Network Design Issues

One of the major issues in the WAN is that H.323 uses UDP instead of TCP. TCP is not practical to use for a video transmission protocol as TCP uses packet acknowledgements and retransmits to guarantee end to end connectivity. Due to the real-time nature of video applications, if the packet does not make it through the first time, it is more important to process the next packet then to try to delay the session while waiting for a re-transmit of the previous packets. TCP also has a congestion control mechanism that reduces the amount of bandwidth that it uses (slows down the transmission speed) when packet loss occurs. UDP does not provide this feature. Therefore if an H.323 session using UDP, and a FTP session which uses TCP are competing for bandwidth on the same link, the FTP session will slow down, but the H.323 session will continue to flood the link. This type of behavior can make H.323 (or any UDP service) a poor application to run over congested links.

Different queuing paradigms can help or hurt a networked application's performance. Modern queuing algorithms allow traffic to be manually prioritized. Of course, one of the first issues in prioritizing traffic is deciding what traffic should get priority. This can quickly get tangled up in a corporate decision making process. The final result will probably not be reasonable to implement and manage on a large network. The fair queuing techniques implemented in some routers attempts to share the bandwidth amongst IP flows. This allows an improvement in overall performance. Fair queuing is relatively easy to configure as it is generally either enabled or disabled for a given router interface. It may also provide some

relief from the questions of setting policy on network bandwidth since all applications are treated equally.

Table 1and Table 2 provides a brief look at the effects of FIFO queuing vs. fair queuing router configurations on an FTP and H.323 session competing for the same bandwidth. This test was done on the same configuration as the bandwidth tests discussed earlier. When the H.323 session is configured for ISDN bandwidth, it uses about half of the ISDN links capacity. The TCP flow control mechanisms used by FTP result in the FTP session using the remaining bandwidth. In this configuration the router's queuing configuration does not matter, as shown in Table 1.

When the H.323 session is configured for LAN bandwidth, the H.323 session attempts to consume all of the available bandwidth. With FIFO queuing, the congestion control in TCP slows the FTP session down to very low bandwidth, while the incorrectly configured H.323 session uses the majority of the available link capacity. This is shown in Table 2. When the router is configured for Fair Queuing, the H.323 session and the FTP session achieve nearly the same throughput, and the FTP session is still effective.

Table 1 Queuing at H.323 ISDN setting

	H.323 @ ISDN speed	
	FIFO Queuing	Fair Queuing
H.323	7000 bytes/sec	7000 bytes/sec
Ftp	8000 bytes/sec	8000 bytes/sec

Table 2 Queuing at H.323 LAN setting

	H.323 @ LAN setting	
	FIFO Queuing	Fair Queuing
H.323	15000 bytes/sec	8000 bytes/sec
Ftp	< 100 bytes/sec	7000 bytes/sec

The fair queuing technique provides an effective mechansim to limit the impact of H.323 on other applications. It does not guarantee that sufficient bandwidth will be available to support H.323. Since fair queuing attempts to evenly divide up the available bandwidth, 19 FTP sessions and one H.323 session would each receive about 750 bytes/sec or about 6 kbps. This would result in unuseable video quality for any H.323 bandwidth setting.

LAN Network Design Issues

Given the bandwidth, packet loss, and latency issues, H.323 can be implemented in the LAN if the current network is not saturated. If the LAN is currently

saturated, then H.323 will probably perform reasonably well but may seriously impact the performance of TCP based applications such as ftp, http, etc. The fair queuing techniques discussed earlier are generally not useable in a LAN environment and therefore will not help if the LAN network is running at full capacity. For saturated LAN networks, the solution to supporting H.323 will likely require adding higher speed backbones and switching architectures. This growth curve should be monitored for impact on network design issues. A network with 100Mbps trunks, and 10Mbps switched Ethernet, or 10Mbps shared amongst a small user pool, should suffice to provide good H.323 operations. In a campus network environment, with multiple switch hops and router hops, H.323 should behave well for most uses if existing performance is reasonable. As with any application, as H.323 usage grows the network will have to be scaled accordingly or gatekeepers added to the network to preserve the network performance.

Future: Gatekeepers, RSVP, and Differentiated Services

In supporting sufficient bandwidth for H.323 on a wide area network, it may be necessary to give H.323 packets priority over other IP packets, as long as the H.323 behaves properly. Determining and assuring correct behaviour is the realm of the gatekeeper. One idea is to couple the gatekeeper to a bandwidth reservation mechanism such as RSVP or the IETF's differentiated services efforts. Since the gatekeeper is controlled by the network administrators, it is possible to police the H.323 sessions. While there are mention of RSVP enabled gatekeepers, this still requires a network which supports RSVP or differentiated services. Differentiated services is not yet a finalized standard, and RSVP has not caught on for wide-scale deployment.

Conclusions

Supporting limited use H.323 in an Enterprise network is possible. Scaling it to medium use levels can be achieved with some attention to the wide-area network design. A solid network design will probably serve better than sophisticated and time consuming management techniques. Ignoring H.323 can be catastrophic since it can cause very poor performance for other network applications. Replacing the office telephone with H.323 desktops and central telephony gateways may be commonplace some day Supporting this level of reliability for H.323 over the WAN will require some form of IP quality of service, which is currently the focus of much attention in both the research world and the marketplace.

References

[H.323] Recommendation H.323, "Packet-based multimedia communications systems", 2/98, International Telcommunications Union

[DiffServ] "A Framework for Differentiated Services", <draft-ietf-diffserv-framework-01.txt>, Yoram Bernet, et al, Oct 1998, IETF Internet Draft.

Filename: H323Usenix3.doc
Directory: P:
Template: D:\Program Files\Microsoft Office\Templates\Normal.dot
Title: Supporting H
Subject:
Author: Randal Abler
Keywords:
Comments:
Creation Date: 02/23/99 10:45 AM
Change Number: 3
Last Saved On: 02/23/99 2:43 PM
Last Saved By: Randal Abler
Total Editing Time: 1,482 Minutes
Last Printed On: 02/25/99 1:09 PM
As of Last Complete Printing
    Number of Pages: 6
    Number of Words: 3,211 (approx.)
    Number of Characters: 18,306 (approx.)