brought to you by the

Video Development Initiative

with contributions from

University of Alabama at Birmingham
Australian National University
Canarie
Emory University
Georgia Institute of Technology
University of North Carolina at Chapel Hill
Ohio State University
Penn State University
Southeastern Universities Research Association
University of Tennessee

April, 2002

Preface

Introduction

The first public videoconference was held in April 1930, between AT&T headquarters and their Bell Laboratory in New York City. [Rosen] Microphones and loudspeakers transmitted the audio while, under a blue light, their images were captured and transmitted as they looked into photoelectric cells. An article in the April 10 edition of the New York Daily Mirror described the audio as clear and the image as inoffensive (a term commonly used for driver's license photos but not often heard today for videoconferencing!) It was at this time that the value of face-to-face conversation at a distance was expressed.

Shortly thereafter, in 1933, the FCC was formed when much radio and television traffic began to collide. In 1934, the standards wars between companies began with the FCC intervening to establish hearings and approve standards. In 1941 the first analog standard for television, with 4.2 MHz of bandwidth (525 scan lines and 30 frames per second), was adopted. By the 1950's we had 83 channels covering the frequencies 54 to 890 MHz.

But it was thirty years after that first AT&T videoconference, before the first videoconferencing product was introduced on the market. In 1964, AT&T introduced its Picturephone at the New York City World's Fair. This system, marketed as an exclusive executive tool, required 1 MHz processing power (considered daunting at the time) and provided the first data sharing feature. In 1971, the first transatlantic videoconference occurred between two Ericsson systems (a product named LME.) And some twenty years later, desktop videoconferencing clients became available.

Intel, PictureTel, and VTEL were some of the early desktop players. Others - Zydacron, VCON, Polycom - followed as the market grew and changed. That market is continuing to grow and change and it is unclear whether H.323 will continue as the predominant standard for videoconferencing over IP or if other standards (ad hoc or formal) will move ahead. More likely, all of the various technologies being used to enable virtual presence today will change quite radically as we learn more about what we can and might do with virtual presence in the future.

Welcome to Version 3.0 of the cookbook. We are happy to bring you another update. If you are new to videoconferencing and the cookbook, we hope you find it easy to use and that the cookbook helps ease your entry into one of the newest and most promising uses of the Internet. If you've already read Version 1.0 or 2.0 or if you are a veteran of videoconferencing, we hope that you find the new content here interesting and useful.

Preface
This section has been reviewed and updated as needed.

Uses of Videoconferencing
Uses, applications, and case studies have been pulled together into this section.

Popular Collaborative Technologies
H.323 is our focus and we highlight current technology here. We also take the opportunity to introduce several other technologies for videoconferencing.

Emerging Collaborative Technologies
This section describes a few things we might see down the road like satellite videoconferencing and even teleportation.

Basic Requirements for Successful Videoconferencing
Basic and add-on components are featured in this section. Tips on selecting vendors and negotiating contracts with them are included.

Best Practices and Etiquette
Here we give tips on how to look and sound your best through technical discussions on the audio and video environments and the basic "rules of the road" when in a videoconference.

Practical Steps
So you're wondering what these interfaces look like? Stop by here for a preview.

Network Matters

This section describes the typical network connection needed for good videoconferencing. It includes a list of typical problems seen when the network connection is under configured or experiencing problems. Tips on who to talk to at your site and tools to monitor your connection are suggested. This version discusses firewall use and home networks using cable modem and DSL.

Advanced Components and Management
This section describes the complex equipment whose role includes multipoint conferencing, gatekeeper security, and signaling translation. Good information for new site administrators is here like profiling your campus needs and other deployment issues.

Related Topics
This section covers various related and interesting things such as selecting and tuning your PC for videoconferencing, developing a room for videoconferencing, and multicast. We've added a new section on broadcasting and archiving videoconferences.

ViDe Favorite Recipes
Well, it *is* a cookbook...

Glossary
Videoconferencing related terms.

Appendices
Standards specifications, cookbook references, listservs, hotlists.

Contributors
This section has been updated to include authors, graphic artists and reviewers for all versions.

How to Use this Cookbook

The Interface

Since this cookbook is a web document, we should give you a few guidelines on how best to view it. First, this is a framed document. The left-hand frame contains a general table of contents. Clicking on any item in this list will cause the content in that section to come up into the main frame on the right-hand side (where you are probably reading this now.)

Secondly, this cookbook has been developed to work with Netscape Navigator 4.0+ and Internet Explorer 4.0+. Some aspects of the interface may not work with earlier versions of these browsers. It has been designed to work best with a 17-inch monitor. If you are using a smaller monitor, scroll bars will appear in several of the frames. Your window should be resized as large as possible before loading this document. Sizing afterwards can cause formatting problems.

You will notice that the bottom frame contains a number of buttons.

This button will load a printable copy of the cookbook in a separate browser window.

These buttons will allow you to move back and forth in the cookbook sections.

This is not the same back and forward as the window history back and forward. The window history can be accessed, as usual, with the browser commands within the main content window.

This button will load a more detailed table of contents in the main window.

This button will load the glossary of terms into a separate browser window. Using a separate window allows you to look up terms as you are reading cookbook content.

We intend to expand and improve this document in future versions. We would like to hear your comments and suggestions. This button will load a feedback form in a separate browser window.

Lastly, the bottom frame contains copyright and contact information. You can easily reach the cookbook editor by clicking on the link.

What is Videoconferencing?

Videoconferencing in its most basic form is the transmission of synchronized image (video) and speech (audio) back and forth between two or more physically separate locations, simulating an exchange as if the two (or more) participants were in the same physical conversation. This is accomplished through the use of cameras (to capture and send video from your local endpoint), video displays (to display video received from remote endpoints), microphones (to capture and send audio from your local endpoint), and speakers (to play audio received from remote endpoints). Although there are many factors that serve to modify or increase the complexity of this basic definition (several of which are discussed in this cookbook), it is useful to keep the concept simple in the beginning when deciding why or how you may be able to use videoconferencing for yourself or your organization.

In understanding the role that videoconferencing could play, consider two general situations: a) those where you are already able to communicate with someone who is not physically nearby, but you wish that communication could be richer, and b) those where you wish to access or communicate to a location that may or may not be nearby but is limited by situational or physical constraints. Distance education often comes to mind first when considering the former situation, but several other existing types of communications can also be enhanced or extended. These include organizational and cross-organizational meetings, counseling, foreign language and cultural exchanges, and telecommuting. Communication is already occurring in each of these applications, but could be made more compelling, more effective, or less expensive through the use of videoconferencing. (Imagine a telephone call where you can see the speaker, or a television through which you can talk.) For the latter situation, the introduction of videoconferencing has enabled communication to restricted areas such as clean rooms, nuclear facilities, operating rooms, and the space shuttle. It has been used to observe wildlife in their natural habitat, to establish interactive surveillance and security, and, combined with micro-instrumentation, to observe inside the human body. This side of videoconferencing may not come to mind as readily as the enhancement of simple communication but it can be quite powerful. Simply imagine situations where you might like to be a "fly on the wall", with the ability to interact if desired. To imagine even further, consider that videoconferencing can be point-to-point (between two endpoints), or multi-point (combining two or more endpoints into the same "conversation"). When you begin to combine diverse endpoints into one setting where audio and video from each can be shared in real-time, whole new levels of interaction are enabled and entirely new ideas for communication can result.

Once you determine that videoconferencing is for you, you need to be aware that it is not currently a "plug-and-play" technology. Videoconferencing actually began over a decade ago with the introduction of expensive group conferencing systems designed to send and receive compressed audio and video over network connections that could guarantee a dedicated rate of transmission and predictable service (i.e., point-to-point T1 or fractional T1 communication links, switched connections using ISDN, or ATM). Standards surrounding how the audio and video would be compressed, how the endpoints would communicate with each other (i.e., initiating/terminating calls, negotiating audio/video compatibility, indicating error conditions during a call), and how the video streams would travel over the network eventually evolved but systems were not fully interoperable at the start. Still, evolution persisted and useful videoconferencing did finally emerge. Arguably the most popular and extensible early compressed videoconferencing was enabled via the ITU (International Telecommunications Union) standard called H.320. However, even with H.320, videoconferencing remained largely restricted to a) those who could afford the technology and network connections to establish meeting rooms, and b) those who were able to travel to a videoconference enabled meeting location.

As time has gone on, the above restrictions have lessened. The technology itself for conducting videoconferencing has become less expensive, more flexible, and now includes options for desktop videoconferencing as well as group videoconferencing. More ubiquitous network types, particularly TCP/IP as used on the Internet, are commonly used to provide less expensive and more flexible connections. In conjunction with this, a new ITU standard has emerged for supporting audio/videoconferencing over IP. This new standard, H.323, was first approved by the ITU in 1996 And has evolved through several additional versions since then and been implemented in a wide variety of commercially available products. Those products are the primary focus of this cookbook, however, information regarding other standards and means for videoconferencing is also included. In addition, the cookbook will touch on many other factors required for a thorough understanding of videoconferencing. These include the importance of standards in general; videoconferencing needs assessment; application possibilities; basic equipment selection and use; and advanced components and services. It is both hoped and anticipated that this cookbook will help you to move from imagining what you might do with videoconferencing to a successful and effective videoconferencing deployment.

Who are the Intended Readers?

This videoconferencing cookbook has been prepared for academic and research users on advanced IP networks around the world. We feel that the span of that topography and experience will make the cookbook valuable to any academic institution desiring to implement videoconferencing for local, state, regional, national, and international communications.

The application examples here are targeted towards the academic and research community in particular and include meetings (one-on-one to many-on-many), classes, and collaboration. The audience levels will range from the beginning user of videoconferencing to the intermediate/advanced user of videoconferencing to the new organizational integrator. It is expected that the beginning user has operational skills on a Windows type of workstation, including general software installation skills. It is expected that the new integrator has knowledge and skills relating to their local and extended networks as well as general server-asset support.

In attempting to analyze the audience attitudes toward videoconferencing, it is acknowledged that attitudes will range from excitement and abandon to caution and order to even skepticism and apprehension. The expectations of videoconferencing are likely to range from top end audio and video ("Why, it's like you're right across the table!") down to good audio with passable video ("Is that a new hair-do or is your camera malfunctioning?")

For the beginning user the objective is to familiarize you with the concept and uses of videoconferencing. To that end we will lay out a potential strategy for selecting and purchasing a videoconferencing product, suggest steps to follow in learning the use of the product, and share ideas about how to introduce the product into your professional life.

For the intermediate user the objective is to bring out new ideas for your use of videoconferencing and to familiarize you with some advanced features and enhanced components for videoconferencing.

For the new integrator the objective is to familiarize you with potential uses of videoconferencing at your site, to introduce the different components and services that will be required to support those uses, and to share experiences and shortcuts for such support.

Why are Standards, Openness, and Interoperability Important?

H.323 is an International Telecommunications Union (ITU) standard for videoconferencing over IP. It is an umbrella standard that specifies mandatory and optional requirements in several areas to enable a complete "call" or communication sequence. The standard also defines four major components that may be part of the call - terminals, gateways, gatekeepers, and multi-point control units. The reason for the standard is to enable interoperability between different vendors' implementations of these components. As is the case with all standards, there is a danger of either over-specification or under-specification. If the standard is over-specified, it may become difficult to implement in the form of a cost-effective product. If the standard is under-specified, there may be room for different interpretations that lead to equally compliant yet non-interoperable implementations. Version 1.0 of the H.323 specification left significant latitude for vendor interpretation. This latitude enabled wide differentiation in the marketplace but led to poor interoperability among early products. Subsequent versions of the standard are addressing this issue by becoming more specific in key areas but interoperability between vendor implementations remains an issue, as does interaction of the various H.323 components across the Internet vs. an intranet.

Fortunately, market forces have resulted in several strategic partnerships among videoconferencing vendors, which will tend to increase interoperability in this arena. In some cases vendors have sought to acquire complementary products in order to offer complete "turnkey" solutions. In others, joint ventures have been formed to assure interoperability within a broader product line.

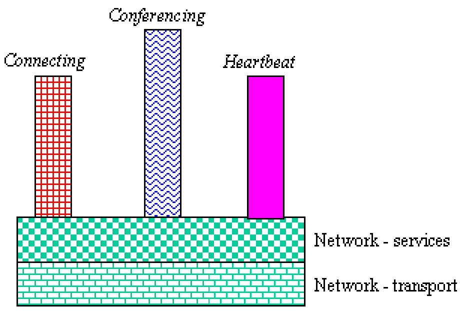
We are still at the early stages of the H.323 lifecycle. While the specifications paint a picture of seamless conferencing over the Internet, today's reality is still somewhat uneven interoperability among H.323 products most reliably suited for intranet deployment. As the standard evolves through future versions (version 3.0 was finalized September 1999) and as product cycles have time to reflect that evolution, this situation is improving.

What are the Basic Ingredients?

Like all good recipes, developing a videoconferencing environment requires a combination of good ingredients, careful preparation, and practice. Attention to detail will only enhance the experience.

The requirements to deliver audio, video and applications across a network can be thought of as a birthday cake. The network forms the multi-layer foundation, with all its attendant icing, and the ultimate goal is to support the candles so that everyone can see the light, i.e. to support the videoconference between participants.

This is best explained in a picture. Preferably drawn by somebody who has some artistic skills, but in their absence…



As this picture beautifully demonstrates, there are 5 major components to videoconferencing, and only one of those is the actual conference that is the audio, video and data that two or more sites want to share. It is possible to take any videoconferencing product and identify its components like this, and it provides a simple model for comparisons of different products. All of these components are covered in greater detail in later sections of the cookbook, but a brief summary is provided here.

The **network transport** is the actual network layer between your sites. This can be IP (Internet Protocol), ISDN, ATM, Frame Relay, DSL, carrier pigeons, dog sleds, whatever. It is a link typically provided across long distances by commercial providers. Each of these network types has its own peculiarities which lend themselves, or not, to videoconferencing. This cookbook will focus predominantly on IP-based products, but will also look at how *gateways* provide a mechanism to connect IP-products with ISDN-products and others.

The **network services** are special features provided by the transport layer. These include concepts such as *quality-of-service* where some kinds of traffic are given preferential treatment to other traffic, and *multipoint or multicast* services, where traffic from one site can be delivered to more than one destination. This last concept allows you to run videoconferences between multiple sites, potentially allowing everyone to see everyone else. It may also provide varying levels of *security* for your videoconference.

The **"Connecting"** candle provides the mechanisms by which two or more sites can call each other. This requires special features such as *directory services***,** which are like dialing instructions, *authentication and authorization services*, which allow you to control who can use your videoconference systems and how, and *accounting services*, so you can bill people afterwards. This candle is usually always on, since it needs to be available all the time. Some products though make do without it, which may or may not be important in your environment.

The **"Heartbeat"** candle is only lit during a videoconference. It provides information behind the scenes so that a videoconference can run as smoothly as possible. This includes *performance feedback*, where a receiver can tell a sender to slow down when it can't handle the incoming flow, or that it is losing information somewhere along the network path, and also *connectivity feedback*, where a receiver can tell the sender it has accepted a call, or conversely that it has hung up and the sender needs to stop sending.

The **"Conferencing"** candle is of course the whole point of this cake. It contains the *audio and video signal*, usually compressed and encapsulated in some fashion, and also supports *data or application sharing* which allows you for example to view presentation material at full quality or a remote computer desktop, or share files around amongst the participants.

Having built this cake, you have to have some way of eating it and hopefully enjoying it. This is provided by the equipment at each site, and is usually called the *terminal*. These can range from large-scale room-based systems say with a board table or a lecture theatre, multiple displays and cameras, right down to ultra-portable plug-in units for a laptop, and a whole range in between. Supporting these terminals requires different levels of audio and video quality input and output. Some terminals do the work for you, whereas others do not.

The cake above generally doesn't care how you eat it. You can usually use different kinds of terminals to attach to the same videoconference, as long as they talk the same baking (network and candle) standards underneath. At the same time, you can use the same room audio and video systems to drive a variety of different videoconferencing products.

Finally, note that *video-streaming*, or *video-on-demand* services, are functionally very similar to a one-way videoconference, regardless of being "live" or pre-recorded. You will find many videoconferencing standards re-used in streaming systems, and the overall birthday cake structure is the same, with a few changes to the candles. This re-use allows for some interesting arrangements, such as streaming a videoconference to receive-only participants, or introducing video-streams into a videoconference.

Uses of Videoconferencing

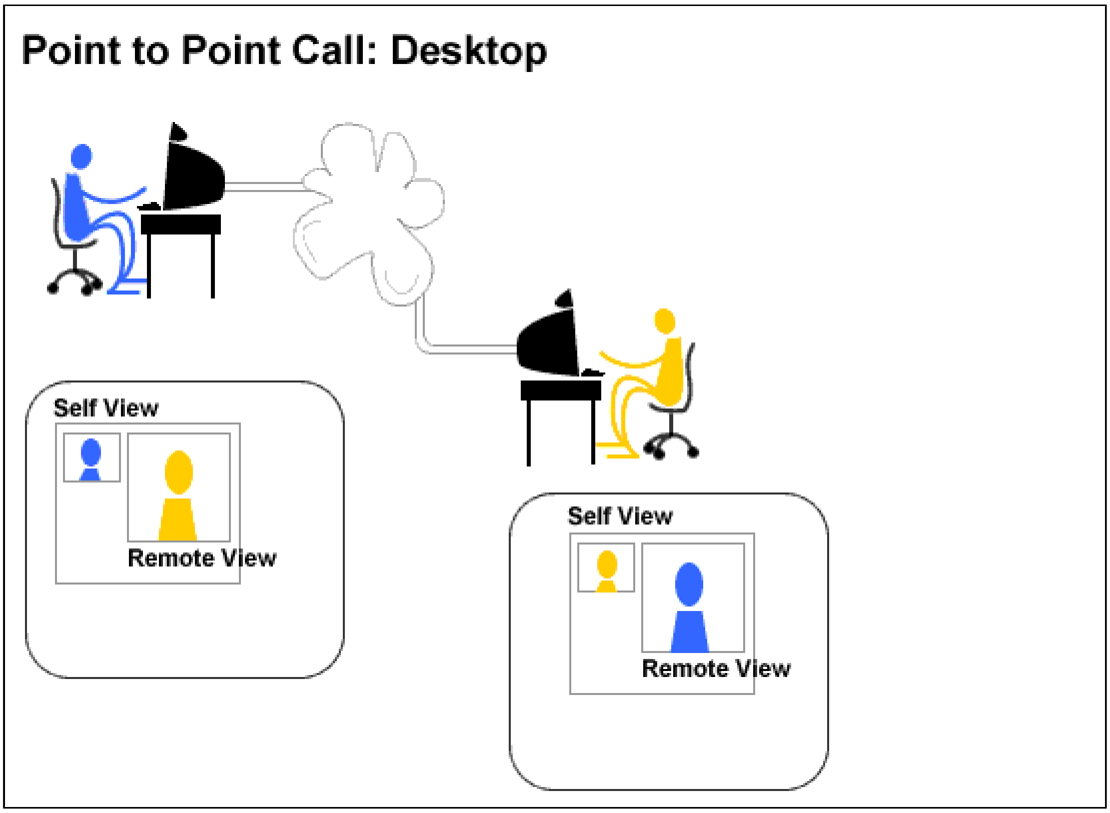
One of the most frequent questions we get is "How would I use videoconferencing?" In this section we have taken several approaches to an answer. First we'll give some general idea of how videoconferencing can contribute to higher education. Next we'll zero in on a few areas that are using videoconferencing to improve or broaden their reach and service. Finally we will describe explicit case studies including those with direct ViDe member involvement.

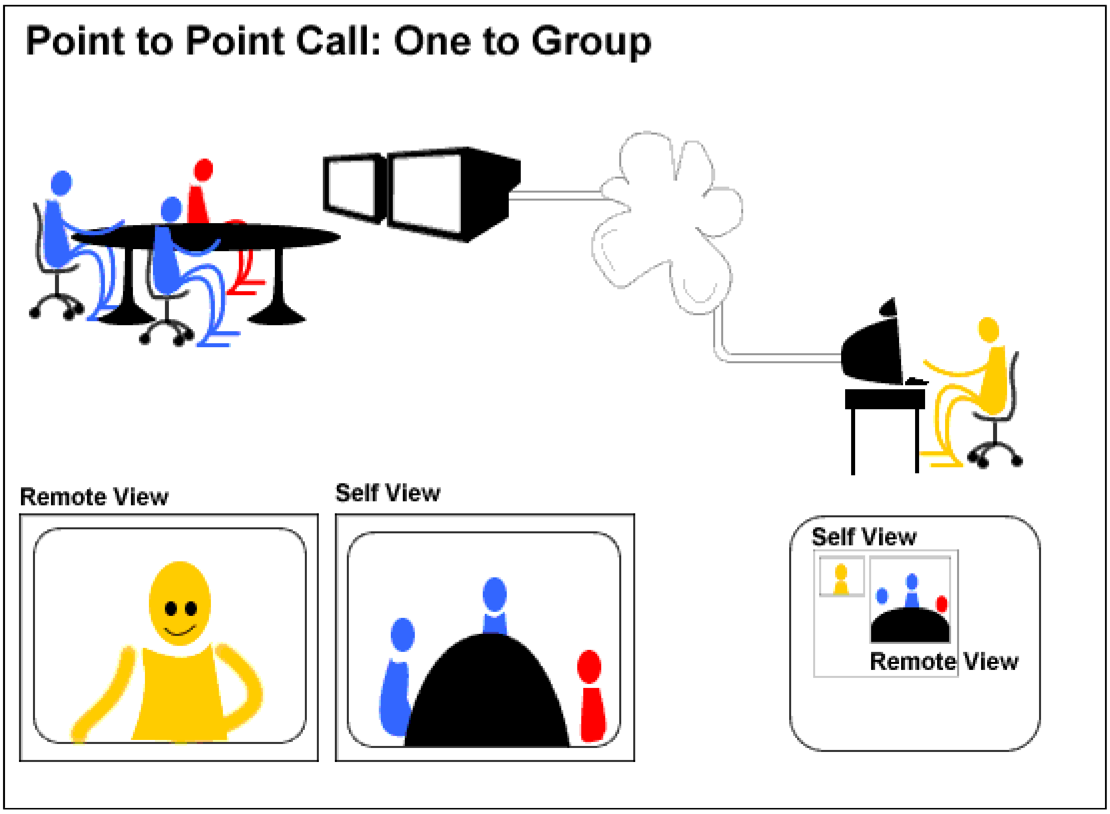
General Uses

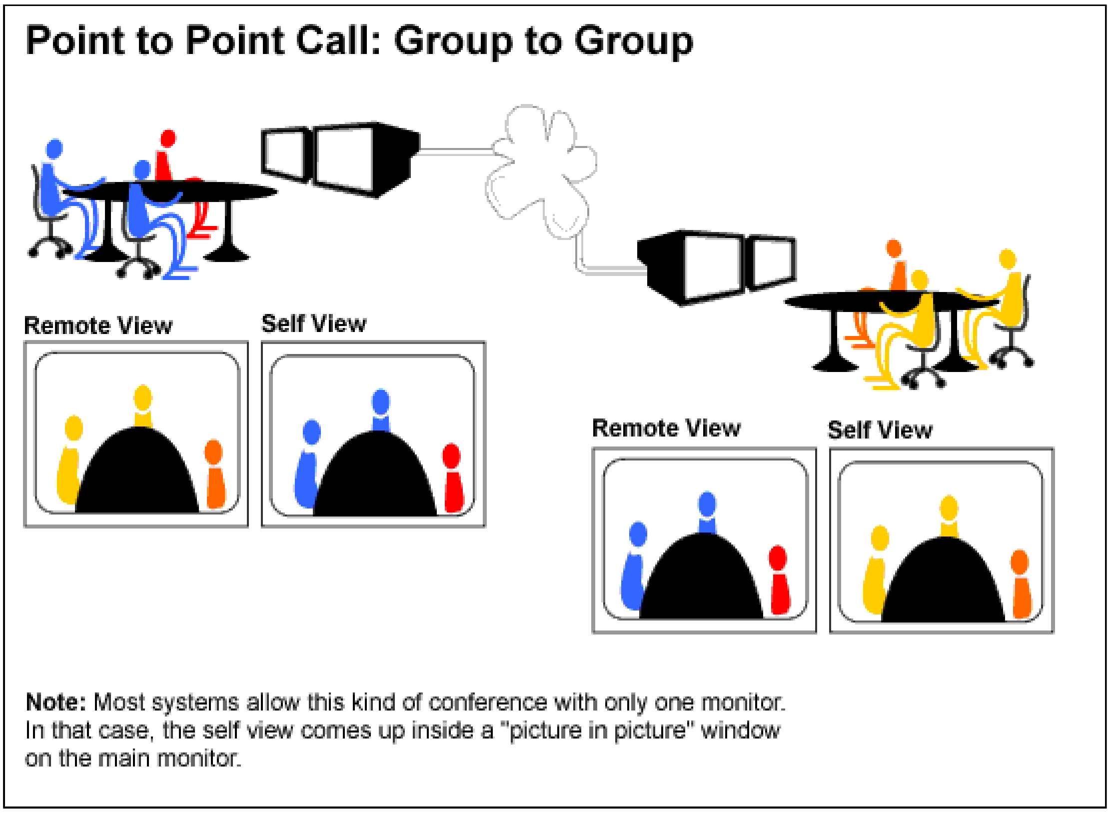
Facilitating attendance at meetings is one of the simplest yet most popular uses of videoconferencing. For meetings that already regularly take place and require face-to-face communication, videoconferencing can substitute for the actual physical presence of remote participants. This reduces travel costs as well as travel time and makes meeting attendance more convenient. It can also make meetings more likely to occur. Frequent and/or ad hoc meetings that might not have been scheduled due to travel costs and timing can be enabled via videoconferencing and enhance the sense of teamwork among people at different locations but working on the same project. Videoconferencing provides remote participants with much of the face-to-face familiarity that comes with physical presence, including elements of facial expression, body language, and eye contact. If videoconferencing is readily available on individual desktops, the cohesive effects of this enhanced communication can be even greater. Collaborative work can then be enhanced further through the integration of videoconferencing with collaborative electronic tools (data transfer, shared whiteboards, shared applications.) These will be discussed later in this section.

Continuing where the above section left off, we'll try to describe how meetings can be handled via videoconferencing. The two following sections will explain some further issues with specialized meetings - classes and collaborative meetings (meetings where work is getting done in the meeting).

In considering the use of videoconferencing for meetings, it helps to think broadly about what a "meeting" really is. In the following illustrations, meetings that include videoconferencing are shown as instances of one-to-one, one-to-many, and many-to-many communication.

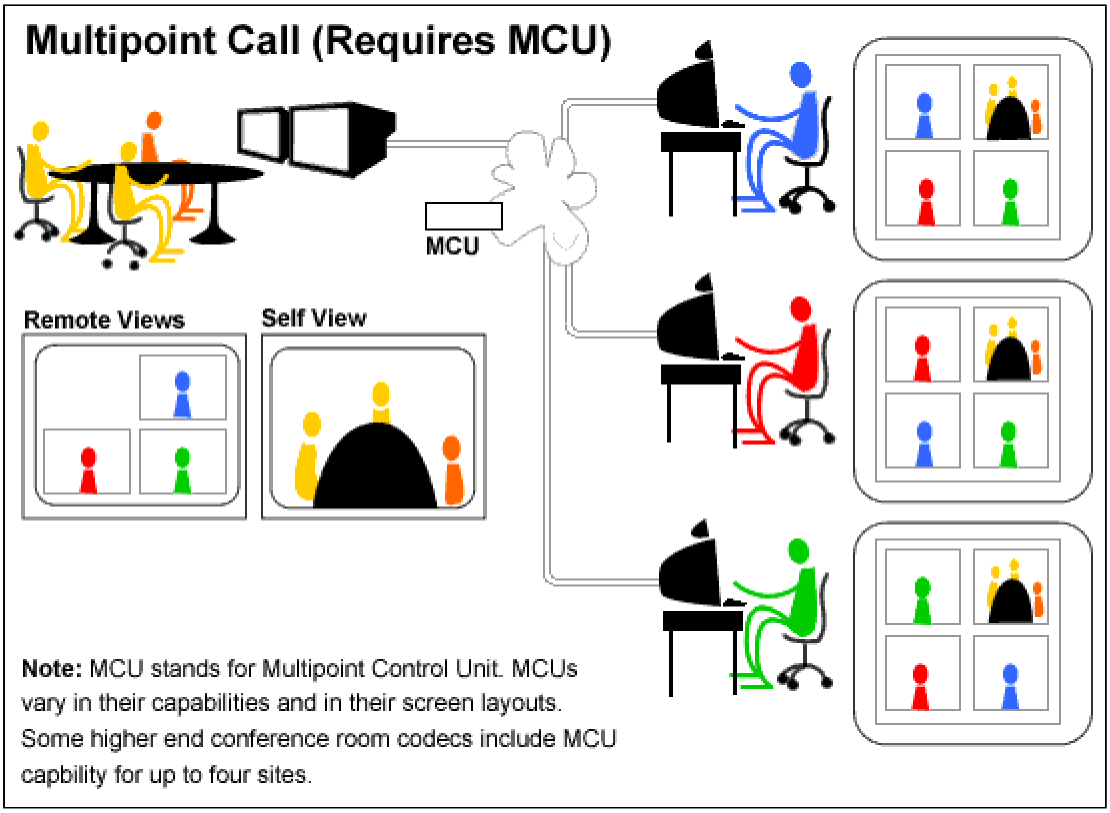






In each of the above cases, the quality of the audio and video are critical to the success of the remote participation. Both will affect whether or not the remote participant(s) feel like they are truly part of a meeting (not just an observer) and also whether or not the other participants treat them as part of the meeting. For meetings, though it may seem a bit counter-intuitive, audio is probably more of a "show stopper" than video. Minor hiccups in the video (pixelization, freezes, etc.) are often tolerated by users. Similar hiccups in the audio make a meeting almost worthless. If there is anything you can do to keep audio quality consistent and high, you should do so.

In the specific case of a multi-point meeting -- where more than one location is participating remotely, several factors affect the success of the remote participation. These include the view participants have of each other, how well participants can hear each other and be heard by each other, and how participants determine who is leading the meeting or "has the floor" at any given time. Features for controlling these factors are discussed in greater detail in the Advanced Components and Management: MCU section of this cookbook, but are previewed below.



What participants see may be

- **Voice Activated** - where the incoming video from the current speaker's location is displayed to all other sites.
- **Continuous Presence (sometimes called "Hollywood Squares")** - where each location can see all other locations (or a selected subset of all locations) at the same time.

What participants hear may be

- **Half duplex audio (sometimes likened to "walkie-talkies")** - where participants can only hear one speaker at a time (audio from the dominant speaker's site suppresses audio from all others) and must indicate somehow when speaker control should be passed.
- **Full duplex audio** - where audio is "natural" in the sense that everyone can hear everyone else at all times.

Meeting control may be

- **No Control**: where full-duplex audio is continuously available from all sites and people can "talk over each other". In this case, the lead speaker is determined by general consensus of those present, just as in a physically proximate meeting. (The view of participant sites would still be either continuous presence or voice-activated).

- **Chair Control**: where a feature is included in the videoconferencing technology (either at the terminal endpoints or in the MCU) to pass chair control via some designated mechanism ("electronic hand raising.") The site possessing chair control is seen and heard by others until chair control is passed.
- **Lecture-Style**: a variation on Chair Control. One site is designated as the lead site and can enable/disable Chair Control access by other sites as well as enable/disable other sites from being heard, or being viewed.

As with any new technology, successful integration of videoconferencing into existing activities requires attention to the needs of the people who will be using it. The determination of what is acceptable and useful must be based on the reaction and comfort level of the end users. In the case of simple point-to-point meetings, there is not a lot of new learning required for participants to successfully interact with each other as long as the video and audio quality do not interfere. Care should be taken to ensure that participants feel they can see and hear each other clearly. More information is available in later sections (see Practical Steps and also the "Developing a Productive Videoconferencing Room" under Related Topics), but typical rules of thumb include:

- Microphones should be of sufficient quality to pick up the speaker's voice naturally (in terms of volume and physical position) and without excessive background noise.
- Microphones and speakers should be positioned so that they do not cause feedback and interference with each other, such as when the microphone picks up the sound from the speakers. Using directional microphones will also help limit the interference.
- Camera quality should be good enough to capture an acceptable image (test with users at the remote site to see how you are coming through) and cameras should ideally be auto-focusing and should auto-adjust for lighting conditions so that participants do not need to adjust them while conferencing.
- Speaker volume and camera position should be user-adjustable, or have proven acceptable auto-adjusting ability with user override capability.
- Displays for incoming video should be positioned as naturally and comfortably as possible for inclusion in the meeting and to enable/encourage eye contact.
- Any conference controls that do not duplicate natural conditions (i.e., voice activation in multi-point conferences) will need to be introduced to users ahead of time since they may need time to practice to become comfortable and effective using them.

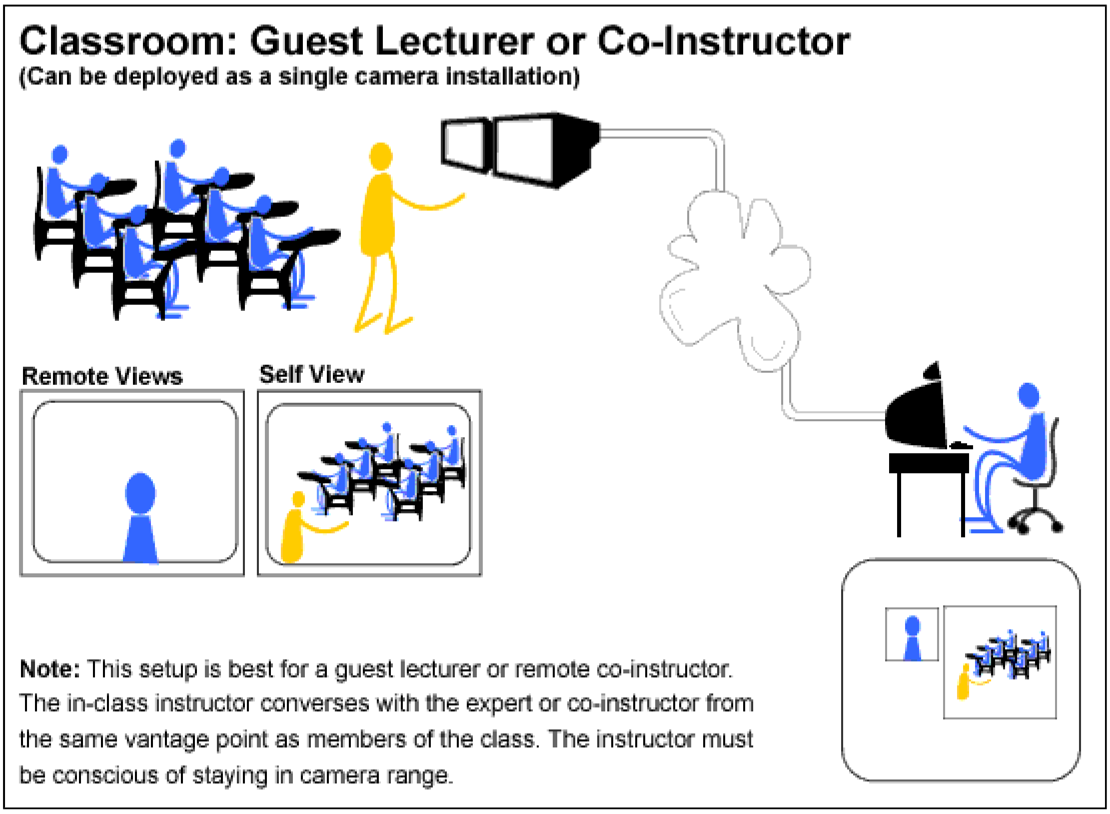
Attention paid to the total "look and feel" of the meeting scenario prior to conferencing helps to ensure that the technology will enhance rather than detract from the success of the meeting.

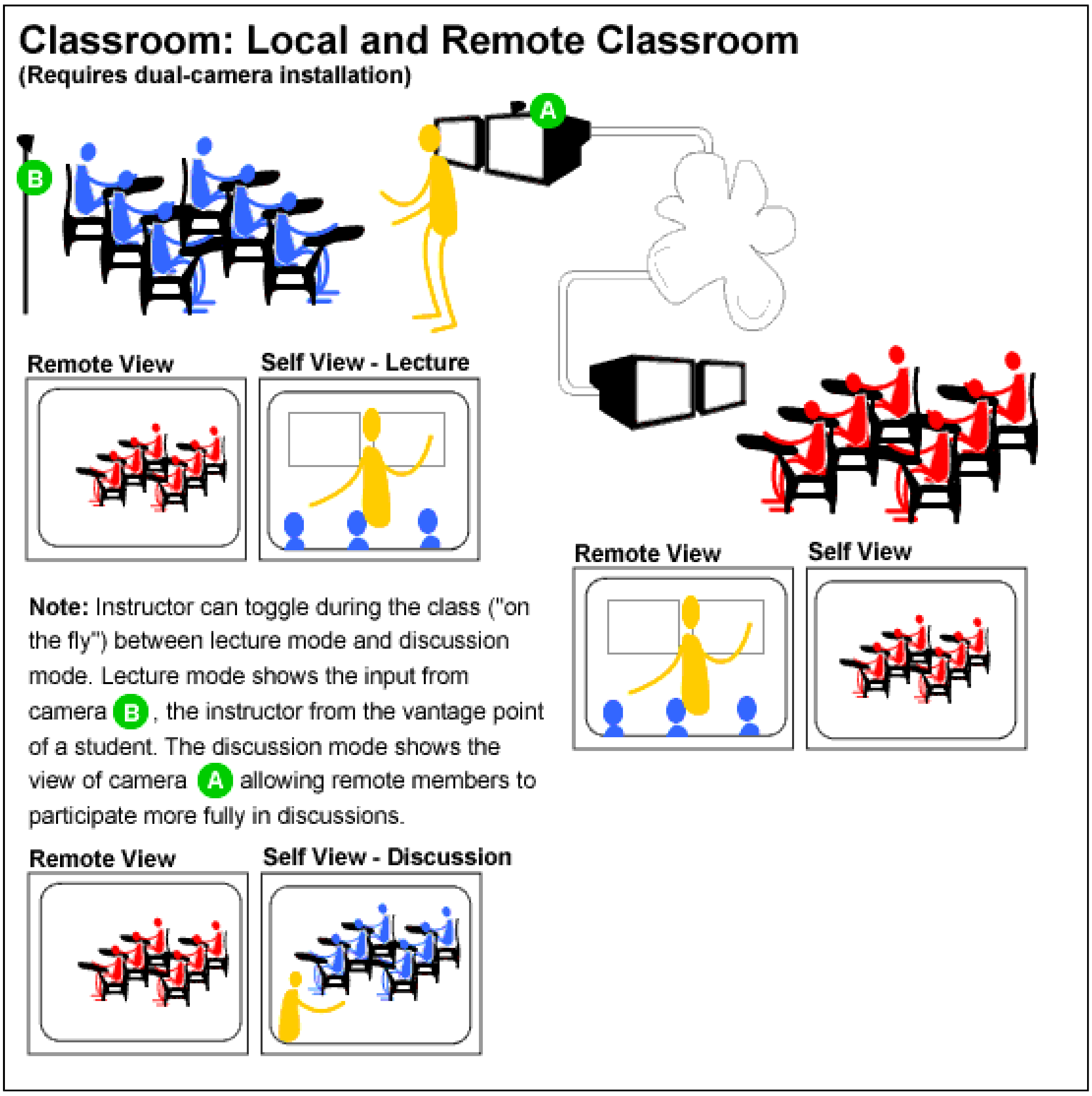
Classroom

A particularly exciting type of "meeting" that may be enhanced and expanded through the use of videoconferencing is classroom instruction. Certainly all of the factors listed above for furthering the success of videoconferencing within general meetings affect the classroom as well. In addition, the introduction of videoconferencing into the classroom means that at least some things about the nature of the instruction necessarily have to change. We'll try to outline some of those changes here.

In one case, remote participants may be additional students that the instructor must now accommodate in terms of instruction and integration with any physically present participants into one student group. Remote participants should not feel that they are getting less out of the class than their physically present counterparts and physically present students should not feel that the presence of remote students is detracting from their instruction. In another case, remote participant(s) may be additions to the instruction itself, such as expert speakers or co-instructors. As with any team-teaching, a cooperative balance of instructional duties is required but this can be made more complicated if video presence cannot compete with physical presence. For instance, instructor accessibility in the physical classroom can easily overtake the presence and command of the remote instructor, encouraging side conversations and inattention to remote instruction.

Two of the most typical classroom scenarios are illustrated below:





Yet another aspect of videoconferencing in the classroom is that the "participants" being shared via the videoconference connection might not always be human. An instructor may want to incorporate an alternative video source (e.g., a document camera, a VCR) for sending to remote locations, or may want to receive video from an alternative video source at the remote site. The potential for combining video inputs and outputs can seem endless and readers are encouraged to fully explore these options when evaluating videoconferencing equipment for use within a classroom.

Most importantly, use of videoconferencing in the classroom requires special attention to the comfort level, teaching style, and instructional techniques of the instructor. In the ideal world, preparation for the use of videoconferencing in the classroom would be minimal. However, today's reality dictates that there will have to be some adapting and learning on the part of instructors to use videoconferencing successfully for instruction. Practice time outside of actual class time must be available and utilized to effectively integrate the technology with their own instructional style and methods, thereby ensuring a natural flow of classroom activities by the time the technology is experienced by the students.

Collaboration

As the previous sections describe, videoconferencing can be used very effectively for meetings and classes. Travel costs and stress can be reduced while personal interaction can remain high. More people can be reached with knowledge and information when videoconferencing is used in the classroom. This section will describe going past the mere communication of presence or presentations. Collaboration is the process of working together. Videoconferencing systems can be designed to support rich multimodal interactions between sites.

A videoconferencing terminal will generally come with a number of software tools including electronic whiteboards, ftp, and chats. The whiteboard can be useful for dynamic lectures, collaborative diagramming, brainstorming, and sharing notes. ftp can be used to transfer files quickly without the need for a separate operating system window. Chat can be useful when audio quality is poor or unavailable for some participants or when a subset of participants needs to communicate privately.

An interface is often provided to enable sharing of third party applications that may be installed on participating workstations. This is particularly useful when group work is supported by project-specific software applications. Optimally, communications between the terminal end stations -- while they are sharing these tools and applications -- should be standardized to ensure the highest level of interoperability, access and accuracy. The most common implementations are supported by the ITU standard, T.120. As stated in the DataBeam Tutorial on the T.120 Series Standard, "Established by the International Telecommunications Union (ITU), T.120 is a family of open standards that was defined by leading data communication practitioners in the industry. Over 100 key international vendors, including Apple, AT&T, British Telecom, Cisco Systems, Intel, MCI, Microsoft, and PictureTel, have committed to implementing T.120-based products and services."

Two terms often heard in discussions of T.120 are application sharing and data collaboration. The distinction here primarily revolves around who has control of material or application. In application sharing, the owner of the material or application is allowing the other participants to view it only. In data collaboration, the owner of the material or application is sharing both the view and the ability to modify the material (or run the application.) We will illustrate the use of these through several examples.

Videoconferencing endstation clients that support application sharing and data collaboration often do so through buttons or pull down menus. In most cases, a button will be clicked or menu item selected while the relevant application window is active. The process is very simple. A mouse click will be assumed in these examples. (See Practical Steps for full graphical examples.)

Lecture: Large Class - You are an instructor who has the need to show material from a presentation, web page, or other application that you use to deliver course material. In this case (say it is a large group coming from several distributed sites), you simply want to present the material in one direction. So after activating the window, you simply click on the application sharing button. The material immediately shows up on the screens throughout the conference. As you navigate through the lecture, each screen changes to follow. (Note: it is not necessary for the application to be resident on the receiving machines.)

Lecture: Small Class - This case is similar to the one above except that you are working with a much smaller class. In this case, you might want to have more than just video and audio dialog between yourself and the students (and student to student.) Perhaps you'd like to include some problem solving aspect to the class. You might bring up an electronic whiteboard or other application and start up the shared data collaboration so that each student might present their ideas on a topic or solutions to particular problems.

Presentation Planning - You are an educator, scientist, engineer, technologist. You have been working on a project with others in your field who are separated by quite some distance. Several of you are doing a team presentation so you would like to prepare your slides together. After activating the call between the presenters, one of you will bring up the presentation software and click on the button for application sharing (if only one person will be typing) or data collaboration (if all of you will be entering material.) You are able to discuss the material, analyze the potential audience, schedule each section in your face-to-face dialog. As you agree on layout and topics, you can enter them directly into the presentation.

Proposal Preparation - You are an information technology director who is working with another information technology director at a different school. The two of you are proposing a joint project in educational technologies over advanced networks. You are preparing your material in your favorite publication software. After activation of the call, one of you will bring up the document and click on the data collaboration button. The document will appear on the other director's screen. Each of you can now type into the document. Control is transferred back and forth simply via mouse clicks. Changes will appear on each screen.

Student Projects - It is very common to assign group projects, particularly in higher-level classes and as term projects. This is a good team building strategy that allows the students to tackle larger problems and learn from each other. As long as the students have been located at the same campus, or reasonably close by, this works well. While application sharing and data collaboration could still be used locally (say for those night owls who don't want to drive late at night), a great deal of diversity can be added to the project if the students are in separate locations. Students in environmental studies might be teamed together from diverse locations such as a coastal environment, a mountain environment, a desert environment, etc. The students can use data collaboration to prepare their final reports, run data analysis for all to see, etc.

Scientific Research - You are an engineer and you are studying aircraft wing design with several colleagues who are distributed around the country. You have implemented a large-scale application on a parallel computing system at one of your sites (actually, it could be anywhere on the network!) The person at that site can begin the application and click on data collaboration so that each of you can interact with the model as it runs and see the results as they happen. You are also using CAD software (which runs in an X-Windowed environment) to analyze the output further. One of you will start up the CAD software and click on application sharing. All of you can then view the structures and discuss what happened, what to try next, etc.

These are but a few examples of the diverse uses of videoconferencing for collaboration. In thinking of your own scenarios, consider aspects of your project work or instructive activities where data is being passed back and forth in the form of file or document transfer but is currently being acted on or viewed individually. If manipulation of this data is really intended to support the development of a common product or understanding, these are aspects of your collaborative work that may be enhanced through application and/or data sharing.

Specific Applications

Telemedicine/TeleHealth

Telemedicine is a growing field made possible by improved and more widely available and affordable telecommunications services. Once narrowly defined as direct provision of medical services using telecommunictions technology, the term telemedicine is now being supplanted by "Telehealth", and covers "use of electronic information and communications technologies to provide and support health care when distance separates the participants" (US Veterans Administration) or "use of electronic information and telecommunications technologies to support long-distance clinical health care, patient and professional health-related education, public health and health administration" (Office for the

Advancement of Telehealth, US Health Resources and Services Administration) .

These definitions include activities providing direct and indirect clinical services such as teledermatology, but also include educational and administrative uses of these technologies to support health care, such as for continuing education/administrative videoconferencing. As convergence of video, voice and data networks occurs in the marketplace, it is not surprising that services previously offered on multiple media are now migrating to converged Internet, providing opportunities for convergence of a number of health care activities occurring at a distance.

- **Brief history of Telemedicine/TeleHealth**

  The earliest recorded use of telemedicine was a 1950's Nebraska demonstration project using closed circuit television to provide mental health services from a university medical center to a state hospital 100 miles away. Forty years ago the NASA space flight telemedicine program began so that medical personnel on the ground could monitor astronauts' biomedical responses to space flight and to provide any necessary medical care. NASA's "Telemedicine Space Bridge to Armenia" Project provided medical assistance in response to a severe earthquake in Armenia in 1988. Using a live, two-way satellite link medical personnel at hospitals in Salt Lake City, Houston Texas, and Maryland conducted many sessions with Armenia physicians for a variety of medical consultations. Due to the enormous expense of these pioneering efforts it is only within the last ten years that the practice of telemedicine has begun to move from pilots to public availability.

- **Current Practice of Telemedicine**

  Remote clinical diagnosis via videoconferencing is currently used in some rural areas. A rural doctor or nurse practitioner consults with a physician based at a metropolitan or university hospital. Using videoconferencing technology and specially adapted medical tools, the remote doctor can see the patient, talk with the local health care practitioner, hear a heartbeat through a remote stethoscope, see images from ear/nose/throat exams, or examine skin conditions. This application has typically required leased T-1 telephone or ISDN lines, which can be quite expensive. Due to a number of issues (cost effectiveness, patient/physician acceptance of the technology, licensing and payment issues), the most common use of such facilities has been to provide health care to prisoner populations. Prisoners have a legal right to receive needed medical treatment, but the cost of transporting a prisoner to a medical facility is extremely high since at least two guards and possibly and ambulance are required for transport during trips requiring an entire day. This high transport expense has provided cost-justification for telemedicine in states such as Virginia and Texas.

  A notable pioneer in broader acceptance of telemedicine can be found at the East Carolina University's Telemedicine Center (http://www.telemed.med.ecu.edu/). Their telemedicine program employs an array of interactive video and audio technologies to deliver clinical care and education to the rural population of eastern North Carolina. Since 1992, the Center has supported over 7,500 telemedicine consultations in over 35 different medical specialties, and over 10,000 distance learning and continuing medical education activities. ECU's Telemedicine Center includes an operational communications hub providing connections between points of need and global medical resources utilizing POTS, ISDN, T-1, Microwave, Satellite, and IP technologies.

  The US Veterans' Administration (VA) has had a number of clinical telemedicine programs, including teleradiology/filmless digital imaging, telepathology,

telenuclear medicine (MRI), home telephone monitoring of cardiac pacemakers, telephone liaison care programs, and more. (http://www.va.gov/telemed/about.htm)

Tandberg (http://www.tandbergvision.com) is a vendor offering a system called the HealthCare System III. This turnkey solution combines an H.323 videoconferencing station with medical examination tools that can be read at a distance, thus allowing the doctor to see both the patient and an auxiliary transmission, including sonograms, store and forwards, telescopic images, or a computer or graphic image. NASA is designing a highly portable Telemedicine Instrumentation Pack (TIP) to collect medical audio, video and data from the patient in space.

Another approach is to build your own telemedicine system. By combing a videoconferencing device, a PC, and medical devices that can interface into a computer, one can build a videoconferencing system that best meets the needs of the end users. Georgia Tech's Biomedical Interactive Technology Center (www.bitc.gatech.edu) has developed such a system for use in Sarov, Russia, a former Nuclear City. The system combines off the shelf PC's with a PictureTel codec, digital cameras, medical instruments interfaced to the computer, scanners, printers, and a video microscope. A system like this can be customized to meet the exact needs of the users, but, it lacks the polished, integrated interface one can expect in a commercial system. As the new field of telemedicine grows and evolves, homegrown systems such as these will help refine the needs of systems, and will influence the product design of future commercial telemedicine products.

In an effort to push the field of telemedicine forward significantly, the National Library of Medicine (www.nlm.nih.gov) in 1996 awarded 19 multi-year telemedicine projects intended to serve as models for evaluating the impact of telemedicine, assessing various approaches to confidentiality in telemedicine, and testing emerging health data standards. A symposium summarizing activities in and results from this program was held in March 2001, and proceedings are available on-line. (http://collab.nlm.nih.gov/tutorialspublicationsandmaterials/telesymposiumcd/Contents.html)

An emerging application in telemedicine is home health care delivery using videoconferencing. Patients being monitored for a number of conditions might be sent home with a PC-based telemedicine appliance that can be used to connect them back to a hospital or doctor's office via ISDN, DSL or cable modem connections. Such systems often combine medical diagnostic devices such as stethoscope, ECG, otoscope, etc. Companies pioneering in this field, such as CyberCare Technologies (www.cybercare.com) and American TeleCare (www.americantelecare.com), do exist, but at least one of them is facing severe financial difficulties.

- **Current Issues associated with Telemedicine/TeleHealth applications**

There are many unresolved questions in the field of telemedicine:

Acceptance: The Veterans Administration states that more specific evidence of efficacy/therapeutic and diagnostic impact/cost analysis are needed in many areas of telemedicine; in other words, telemedicine is not widely accepted as an approved medical practice by doctors or patients.

Security: Security of telemedicine sessions and related information is a crucial issue. The Health Information Portability and Accountability Act of 1996 (HIPAA) guarantees privacy of patient medical records by federal law, and with individual as well as institutional accountability for violators. HIPAA compliance becomes mandatory in Spring 2003. (http://www.hcfa.gov/hipaa/hipaahm.htm)

For videoconferencing systems used in telemedicine, an implication is that reliable session encryption will be mandatory. An excellent white paper describing the state of H.323 encryption is available from UKERNA (United Kingdom Research and Education Networking Association). (http://www.ja.net/development/video/vip/reports/south3.pdf) The report describes the H.235 security recommendations that are part of the H.323 standard, but finds little implementation of this standard within H.323 clients. Some vendors such as VCON have built proprietary encryption schemes into their products; these proprietary solutions may solve an immediate problem but prohibit interoperability. The authors conclude that encryption at the H.323 application level is in its infancy so encryption can presently only be offered widely via network layer encryption, for example by use of VPN or IPSec.

Regulation and Reimbursement: Until the late 1990's, private and public third-party payers generally did not have explicit policies to pay for telehealth services. The Office for the Advancement of Telehealth in the US Department of Health & Human Services' Health Resources and Services Administration (http://telehealth.hrsa.gov/) is an excellent source for information on regulatory and reimbursement issues. The Balanced Budget Act of 1997 was a significant step forward for telemedicine because it allowed Medicare payments for patients in rural communities having no health care services at all; this prohibited payment to medical specialists such as cardiologists or radiologists whose expertise might be valuable to a general practice physician. The types of services covered were required to be "interactive" and thus did not apply to "Store & Forward" technologies such as when X-rays or laboratory tests might be forwarded to a specialist for diagnosis. The rules were expanded by the Benefits Improvement Act of 2000 to allow coverage to include any type of medicine to any rural area and to any entity that had participated in a Federal telemedicine demonstration project (for example, Veterans Administration Hospitals, even in urban areas). The practice of medicine across state lines is severely restricted by current laws.

- **Standards and Guidelines for Telemedicine/Telehealth Videoconferencing Systems**

  The Office for the Advancement of Telehealth maintains guidelines for evaluating telemedicine/telehealth systems at http://telehealth.hrsa.gov/pubs/tech/technome.htm. These guidelines are based on goals of interoperability, compatibility, scalability, accessibility, and reliability. Guidelines are also provided by area of medical specialization, such as dermatology or psychiatry.

- **Additional Resources:**
  - **Telemedicine Information Exchange** was created and is maintained by the Telemedicine Research Center with major support from the National Library of Medicine. http://tie.telemed.org/
  - The **American Telemedicine Association** seeks to bring together diverse groups from traditional medicine, academic medical centers, technology and telecommunications companies, e-health, medical societies, government and others to overcome barriers to

the advancement of telemedicine through the professional, ethical and equitable improvement in health care delivery. ATA was established in 1993 as a non-profit organization and is headquartered in Washington. Links to telemedicine organizations in other countries can be found on the ATA web site. http://www.atmeda.org/

Telecommuting

Telecommuting is a fast growing trend in many geographical areas - both rural and urban - as companies and employees realize many benefits that might be gained by this. For both large and small companies, moving some of their workforce to home offices for all or even part of the work week can enable the company overall to lower costs by reducing the size of building space that is required for regular operations. (Note that network and telecommunication costs are more likely to be reallocated rather than reduced since work-at-home employees still need to be provided with these typical and expected services in their home office.) In addition, as home and work times are blending more and more within many fields, the ability to assign the legitimate title of "telecommuting" to time spent working at home helps this time to be recognized and supported as core work hours rather being potentially unaccounted for. While we may dispute whether or not such blending is a good trend, most people who already contribute significant at-home working hours would agree that the increased recognition for such work definitely is. A third benefit of telecommuting when coupled with the increase in global Internet connectivity is that people in many fields are able to now live where they most need or want to be while performing work that may have its basis or effects elsewhere.

In contemplating the richness of a home office environment, videoconferencing can provide a natural and productivity-enhancing improvement to telecommuting, enabling more effective communication between employees who are at home and those who are in the office. Telecommuters can also use videoconferencing to interact with similarly equipped colleagues at other organizations, ranging from meetings to remote presentations and even conference attendance. Obviously this range of profession communications requires more than a "postage stamp" image and poor quality that is often associated with Internet-based conferencing.

One way to achieve a home-based connection that can support the quality that is necessary for reliable telecommuting is through the use of ISDN lines at the home end that are used to dial into an ISDN (H.320)-to-IP-based (H.323) conferencing gateway that is positioned on a well-connected IP network. Of course, this also means that there would be monthly ISDN charges associated with all calls, as well as per-minute charges at whatever rate applies to use of the PSTN between the home office and the gateway. The increasing availability of broadband Internet service options, however, is beginning to offer a less expensive alternative to ISDN in many locations. Both cable modems and xDSL (Digital Subscriber Line) services can support at least low-bandwidth H.323 calls (256K - 384K) and "fiber-to-the-home" initiatives hold the promise of increasing conference quality and capability in the future.

Once more bandwidth is available to an average home and particularly if that bandwidth can be guaranteed, videoconferencing is likely to become widely used for telecommuting. Beyond just conversational interaction, improved conferencing capability will enable data or application sharing (when two employees must jointly work on one shared document, or interact within the same application) and high-impact remote presence for teaching, presenting, and collaborating. Notably, as telecommuting increases in capability and also popularity, we will need to understand and adjust to changes beyond just those related to the changing technology. The concept of "work space" and "home space" will blur even more than they have today, creating a need to re-examine how and when we interact with each other within these spaces. For instance, an increased capability for rich and real time communication on a global scale from your home will not alleviate the problems associated with interacting across different time zones. It may, in fact, compound them. If you may someday soon be able to present your latest research paper to a remote audience at your dinner hour or attend a meeting in the middle of your night, does that mean you will necessarily want to do that? Does it mean that others will come to expect you to? (And, if you did, what would you wear?)

Tele-Education

There are many opportunities for enhancing education with videoconferencing. These include remote-teacher/remote-specialist applications, as well as classroom-to-classroom interaction. It can also be scaled to more than 2 sites, allowing virtual education "teams" to be created. It can used in the K-12 education system, as well as in-house corporate training, government services coordination e.g. for emergency services, and many other education areas. In most cases the equipment can be scaled to suit the situation, allowing for a regularly scheduled education program in fixed locations, or conversely supporting high-impact once-off events such as a visiting dignitary or specialist. It can be used to link sites within a region, or scaled up to international sizes. In the K-12 sector for example this allows students to interact with students their own age in other languages and cultures.

In the remote-teacher approach, a simple H.323 system, attached to a TV or a projector allows a class to interact with a remote instructor. Depending on the situation, the remote instructor may use a fixed or a mobile system (e.g. a zookeeper, or museum staffer), or include a mixture of live and recorded video (e.g. for medical training). Most H.323 (and other video-conference) systems allow for flexible encoding arrangements, where a camera can be managed separately to the encoding device. Typically the receiving site is fixed, such as a classroom, but the interaction still works in both directions. By using a "multipoint-conference-unit"(MCU)/"bridge", or some form of network multicast, more than 2 sites can be supported at the same time, and all participants can interact with each other, if that suits the occasion. It should be noted though that most education events tend to be "lecture-mode", which is predominantly (but not entirely) one way. Managing audio and video in both directions can be a major technical challenge in some situations — e.g. in large lecture theatres with hundreds of participants. The human eye and ear can focus much better than their electronic counterparts.

There are many projects active around the world today. One excellent mailing list focused on the K12 sector is the Ed1VidConf list, at ED1VIDCONF@A05VM.RVR.IL.AMERITECH.COM. Subscribing is non-trivial; the easiest way is to contact the moderator Linda Woods Hyman, lwhyman@pacbell.net, for assistance.

Another large-scale project, which combines H.323 (Internet-based) and H.320 (ISDN-based) videoconferencing with satellite transmission, is the World Bank's Global Development Learning Network (http://www.gdln.org/). The GDLN aims to support developing countries with education and policy support from other countries in their own region and the rest of the world. Typically it is aimed at Universities, government departments and Non-Governmental-Organization (NGO's). Around 100 sites will be connected to the GDLN by around 2005.

Judicial Applications

The judicial system has found videoconferencing to be a cost-effective and productive technology for its needs. A number of counties have begun to install videoconference systems in jails and courthouses. These systems are used most commonly for "video arraignments", where a prisoner will go to a videoconferencing room in the jail. Another system in the courtroom has dedicated cameras to show the judge, prosecutor and defense attorney. The prisoner can see all of the members in the court, and the information is simultaneously recorded with a split screen 4 ways, so a complete record of the proceeding is made. This provides several benefits to the court system, as it reduces the number of defendants who need to be transported from jails to courthouses, reduces the overcrowding of courthouse holding facilities, reduces the security risks associated with transporting and handling defendants, saves time and saves money. Most of these types of systems currently run over private networks using proprietary technologies, but they could be designed using H.323 technology over a secure intranet as well.

Remote Laboratories

Two combined forces are helping to make Remote Laboratories a reality. First, there is the ever-expanding technology, both on the end-point side and in networking infrastructure. Second, there is the tightening of budgets along with the increasing cost of large laboratory instruments and equipment. Another area where videoconferencing is expanding is to provide a way for researchers to share, for example, a scanning tunneling electron microscope using existing (and ever cheaper) network infrastructure. The names for these extensions to videoconferencing are often referred to as "Remote Laboratories" or "Collaboratories".

Remote Laboratories allow scientists to actually *"do science together*" across great distances. Videoconferencing allows scientists to discuss their science, teach and learn from one another, and even give suggestions about things like further research ideas. But Remote Laboratories actually bring videoconferencing, data collaboration and remote instrument control into the lab settings, providing a way for scientists to plan together, work together, and experiment together, both (or all) parties having access to the data, the instruments, and the documentation that is generated by any scientific endeavor.

Remote Laboratories are bigger than "just" videoconferencing. They encompass the entire way a group of people will interact. And since most of that interaction will be mediated by technology, special considerations need to be paid to integrating the modes of interaction into a cohesive, productive, and natural whole. Some of the technological elements of a collaboratory are Intranets: "shared spaces" applications, (with real time, but also persistent, data collaboration environments); videoconferencing (either inside of shared space or augmenting it); and special software or hardware interfaces that allow control of instruments, (things like dials, meters, and response windows.) These sit on top of regular office interaction technologies such as telephony, messaging, group scheduling, and text chat software.

As an example, say that we'd like to create a remote laboratory for advanced optics, like those found in digital systems. Such a laboratory conducts research into the optimal integration of optics and microelectronics. Engineers distributed around the country (and world) would investigate how optics can best be used to increase the capabilities of high-performance digital systems and would work closely with industry to bring the resulting knowledge into the mainstream. A small core team at a main location would interact with a geographically distributed network of partners drawn from the national laboratories and universities. The partner laboratories would act as sources of specific technologies and reservoirs of expertise and specialized capability, whereas the core component would undertake activities generic to the application.

The inherently geographically distributed nature of this structure poses significant challenges to efficient operation, especially in light of the close collaboration that will be required among core and partner laboratory researchers in order to incorporate new technologies into working systems. The operational functions of such an organization can be categorized and described as:

- **Administration, including general accounting, travel, and purchasing**: The web has been used very successfully as the basis for many administrative operations, including general accounting, purchasing, and instructional aspects. Intranets often handle this function, but sometimes they are part of group messaging software.
- **Presentations such as lectures, seminars, and courses:** With higher video bandwidth networks, lectures and seminars would be delivered to participating researchers, lab users, etc.

through videoconferencing. This is "just" videoconferencing.
- **Meetings, conferences, and informal interactions among researchers**: The mechanisms of travel and conventional telephony would be augmented with an advanced videoconferencing functionality that borders on "virtual presence." Here's where it starts to get really interesting.
- **Laboratory interactions involving experimental facilities as well as one or more people**: A "virtual laboratory" environment would be created by configuring all laboratory instrumentation to be compatible with automated control, and by augmenting network videoconferencing capabilities with currently available, network-windows-based laboratory instrument control software.
- **Publications: reports, journal papers, conference presentations:** Standards-based representations for electronic imaging, textual information and graphical information, and a methodology for efficient co-authorship would further the electronic exchange of data and facilitate the preparation and dissemination of reports and publications. Videoconferencing could be used for joint preparation.
- **Public relations and recruitment**: increase the ease with which interviews, both for press and for potential hires can occur.

It must be said that Collaborative Environments are more complex than videoconferencing in a vacuum. Since the applications of the end-points are many, and since they will need to interface with other applications and systems, it's important to be flexible in the design and implementation phases of such a project. Experimenting (with certain pieces of the system) and then integrating (with more trying than buying of *off-the-shelf* software / hardware packages) keeps you flexible and lets you understand each piece of the puzzle separately, instead of some tangle of technology.

Remote Laboratories require each participant to know what's going on, which is one of the hardest parts of technology-mediated communication. "What's happening when I'm not there?" "What is the general feeling around the office?" "What's going on?" The highest aim of such a system is "virtual presence" so that, for example, as remote lab participants arrive to work, they can visit with and be caught up on developments of their remote colleagues. They would do this as they check their email, get coffee, and plan their day. The ability to see who's online, who's in the lab, and get further information about what they're doing (and how it's going) goes far to give life to a remote laboratory. The problems with doing this are substantial. It's a big task to have people check in and check out, but this can be handled by some groupware or shared space software packages. It's an even more considerable task to get people to update meta-information on their projects and activities more than once per week. There are also *big brother is watching you* overtones to a system where others can remotely turn on a collaborator's personal videoconferencing system.

In short, today there is no established, easy way to provide remote laboratory facilities, either off the shelf or using a customized approach. There are many issues to deal with, both technically and socially. But the payoffs to such a system can be tremendous – remote laboratories can enable scientists in ways that were unthinkable just a few years ago.

Surveillance & Security

We usually think of videoconferencing as being two-way and interactive, with both parties equally participating in the exchange. However, videoconferencing can also be used as a one-way monitoring technology simply by selectively muting audio and/or video on one side. This aspect of H.323 conferencing is not often, if ever, on the scope of vendors' current products, but some of today's H.323 products could be adapted for particular uses, and future development along these lines is certainly plausible.

To imagine scenarios where H.323 conferencing could be adapted as a surveillance/security tool, you really only need to look at situations where video is already being deployed for these purposes. Consider in particular those applications where connecting the video device to a LAN could enhance functionality and flexibility, or reduce costs. For instance, many campus and corporate environments have extensive LAN/WAN infrastructures that include both Internet and intranet connectivity, and in parallel provide a closed circuit TV/security system. An H.323 surveillance solution would allow monitoring devices to be positioned anywhere there is a LAN connection and allow video from these sources to be viewed, recorded and easily integrated with other organizational data systems. Many H.323 systems provide remote-control cameras with pan-tilt-zoom capabilities, which can be controlled through some of the H.323 protocol extensions. Some cameras are themselves enhanced with intelligent features such as auto-tracking, auto-signaling when they detect movement, and some are even programmable through serial ports and connections to PCs. One example is the Sony EVI-D30/31 series and their VISCA ports, and other vendors have similar products.

This type of "surveillance" can extend beyond support for security applications. Remote undisturbed monitoring of situations that cannot readily be observed (such as wildlife in their natural habitat, or astronauts working in the space shuttle) can be used for education, training, even entertainment. Ready proximity to other data applications available via the LAN/WAN would allow for easy integration of the captured video into documentation, presentations, educational materials, and even broadcast events. H.323 videoconferencing terminals can also be placed in selected areas to provide "remote attendants", e.g., a "video receptionist", a "video test proctor", or "video lab support ". The other end of the terminal would then reside on the desktop of the actual person performing the function, allowing them to monitor what is going on in the room and also communicate readily when assistance or mediation is necessary.

In any of the above scenarios, an MCU (multipoint conferencing unit), or network multicast, could be used to enable monitoring of several areas at once from a single vantage point.

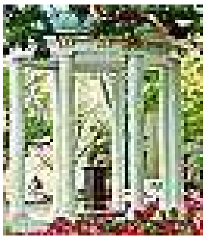
Granted, H.323 terminals truly capable of replacing today's specialty security and surveillance systems may need to evolve in several areas such as robustness, weatherproofing (for outdoor applications), size/shape, etc. However, most high-end terminals today support external cameras with a composite or S-Video/S-VHS input, decoupling the camera from the encoder. With the use of a cheap wireless video transmission system, the camera can be completely independent of the terminal, or the camera-terminal combination can be connected to your network in places where the network cables don’t reach.

Case Studies

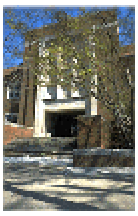
The following examples of H.323 applications have been selected to illustrate the range and variety of emerging H.323 use, and the potential H.323 has for supporting collaboration and resource sharing on our campuses. Whereas H.323 videoconferencing is used in many of the same application domains as its precursor, H.320, (specifically, Distance Education, Telemedicine, and for communication purposes), these case studies should demonstrate that the convenience and cost-effectiveness of H.323 have been recognized by the academic community, and there is no shortage of creative application of the technology. The applications described herein are at various stages of maturity, but all are beyond the mere conceptual stage.

Place your mouse over the image for more information on the project.

University of North Carolina School of Social Work: Teaching and Training



over the Internet.



PEPNET Videoconferencing Testbed

Evaluation of H.323 Videoconferencing for Medical School Planning on a High Performance Statewide Network.

Virtual Rounds: Sharing of live animal clinical cases via H.323

ViDe "Large Scale Video Network Prototype" Project

Popular Collaborative Technologies

- H.323
  - Hardware Assist
  - OpenH.323
  - Software Clients
- VRVS
- MPEG
- Motion JPEG (MJPEG)

- Access Grid
- Web Clients

H.323

In order to communicate effectively, a common language between the two (or more) participants must be chosen. Without a common language, there will be little or no understanding, resulting in a passing of merely "noise" between the parties. This doesn't apply only to people; computers particularly must have common ground when communicating to one another, notably because of the context-free world in which they operate.

Networks of computers, therefore, are built on standards and protocols, selected so applications that are dependent upon the network can exist and operate at their fullest capabilities. There are several notable organizations that work to develop standards, both formal and ad hoc, across a variety of fields. The ITU (International Telecommunications Union) is one such organization, a prominent group that develops internationally recognized recommendations and standards to enable computers, radios, and other machines to interact with one another.

The ITU H.32x family of standards handles multimedia communications. This family includes H.320 (communication over ISDN [integrated services digital networks]) and H.324 (communication over SCN [switched circuit network], better known as traditional phone services).

H.323 is a communication standard produced by the ITU, initiated in late 1996, and aimed at the emerging area of multimedia communication over LANs (local area networks). It is an outgrowth of the traditional H.320 technology but optimized instead for the Internet. H.323 has since been revised to include voice-over IP and IP telephony, as well as gatekeeper-to-gatekeeper communications and other data communications that involve packet-based networks. These networks include IP-based networks like the Internet, Internet Packet Exchange (IPX) LANs, and WANs. H.323 is widely supported by many commercial vendors and used throughout the world in commercial and educational markets.

The H.323 standard specifies a great deal of information about the properties and components that interact within an H.323 environment. It specifies the pieces that combine to provide a complete communication service:

- terminals, either PC or stand alone devices, these are the endpoints of the communication lines
- gatekeepers, the brains of the network; providing services like addressing/identification, authorization, and bandwidth management
- gateways, which serve as translators when connecting to a dissimilar network (such as an H.324, for example)
- MCUs (multipoint control units) which allow multipoint conferencing, or communication between more than two parties at once (much like a traditional conference call on a telephone)

In addition to component types, H.323 also describes protocol standards, permissible audio and video codecs, RAS (registration, admission, and status), call signaling, and control signaling. H.323 specifies a mandatory level of compliance and support for the above specifications for all terminals on the network. More detailed information about H.323 is available through several links shown in the appendices.

Hardware Assist

In most videoconferencing over IP, including H.323, the endpoints compress the audio and video data to be exchanged so that transmission of that data over the network will be faster, less likely to degrade and will consume less network bandwidth. The trade-off for this added network efficiency is a need to compress the media stream as it leaves the sending endpoint and then decompress the media (audio/video) stream for interpretation once it arrives at the receiving endpoint. This compression and decompression require processing power and additional time. The less processing power there is, the more delay is introduced into the communication.

H.323 videoconferencing is often considered a desktop technology but PCs in general are not powerful enough yet to do high-quality full-screen, full-motion video compression and decompression, especially in addition to all the other applications that a PC might be running at the same timea videoconference is desired. This has forced leading H.323 vendors to integrate hardware-based codecs ("compressor/decompressor") into their videoconferencing products. These codecs are specifically designed to offload the compression and decompression task from the PC, allowing the endpoint overall to achieve good performance. In the past and often still today, the codec is included as an additional PCI bus card in the PC. Examples of this include the Zydacron OnWAN and the VCON Escort series. These and similar PC products can support a high frame rate (15 - 30 frames per second) and extended range of call quality/bandwidth settings (128K - 1.5Mb) so that videoconference quality seen at one's desktop can equal that of larger and more expensive room-based conferencing systems.

More recently, with the advent of USB (Universal Serial Bus) technology, the trend is more toward external devices that plug into a USB port. This allows for the simplicity of "plug and play" installation for seemingly simple videoconferencing devices that are also capable of providing relatively high call quality. The extra processing power required is included in the USB camera/device with the USB port providing the "bandwidth" necessary for the compressed video to pass from the camera/device to the PC. Examples of this approach are Polycom's ViaVideo and the VCON ViGo.

Another addition to the "hardware line-up" of videoconferencing endpoints is standalone non-PC-based"appliances",such as the Polycom Viewstation. These appliances are specialized hardware devices (system/camera combinations) that provide high quality medium to large-scale videoconferencing capability. They do not run other programs such as a PC-based endpoint might and they are larger and more expensive, yet often as simple to use, as USB devices on a desktop PC.

Overall, though hardware-based endpoints cost more than their software-only counter-parts (which range from inexpensive to free and are discussed in the following section), the extra cost is often justified in order to achieve videoconferencing call quality that is acceptable beyond just novelty or casual use.

Open Source H.323

Open Source Software

The basic idea behind the open source movement is that when the source code for software is freely available, programmers will read and evaluate that code. They will then make improvements and fix bugs more quickly than any company will be able to with similar proprietary code. Other major advantages of open source code include cost and portability. Free open source code can make software, such as H.323, available where money is an issue. Open source code is portable in that it can be compiled to run on different types of computers and different operating systems.

The Open Source Initiative (OSI), a non-profit corporation that promotes open source software says "Open source promotes software reliability and quality by supporting independent peer review and rapid evolution of source code. To be OSI certified, the software must be distributed under a license that guarantees the right to read, redistribute, modify and use the software freely." http://www.opensource.org

Open H.323

"The OpenH323 project aims to create a full featured, interoperable, Open Source implementation of the ITU H.323 teleconferencing protocol that can be used by personal developers and commercial users without charge." http://www.openh323.org. The OpenH.323 project is "coordinated" by QuickNet, an Australian manufacturer of hardware-based voice codecs for H.323.

The OpenH323 project includes:

OhPhone	A text only H.323 client
Open Phone	A GUI based H.323 client (Currently available forWindows).
OpenMCU	A conferencing server for H.323
OpenAM	An answering machine for H.323
OpenGK	A H.323 gatekeeper
PSTN Gateway	Allows an H.323 client to receive and make calls on the Public Switched Telephone Network (PSTN). PSTN Gateway requires appropriate hardware.
Dump323	A stand alone program that can take hex data and decode the Q.931/H.225 and H.245 packets and present them in a human readable form.
G.711 and GSM	Audio codecs which are supported in software. Most other codecs are covered by patent or other legal restriction and cannot be included in open source without license fees.
H.261	A Video codec which is supported.
Support for:	QuickNet hardware DSP
Clients for:	Windows, Linux & various BSD's SIP implementation

It is difficult, of course, to know just how much OpenH.323 is being used - or any open source software, for that matter. One possible indication is that on SourceForge, the largest online depository of open source software, components of openH.323 are in the 96th percentile for "activity" (downloads, requests for support, postings, etc.). The Open h.323 GK is listed in the 90th percentile, for example.

GnomeMeeting

GnomeMeeting is a GUI-based client for Linux. The client will be in the Gnome package for the next version of Linux. There is evidence of an active and stable GnomeMeeting development community. http://www.gnomemeeting.org/.

GnomeMeeting was written by Damien Sandras as his final year project for his degree in Computer Science Engineering at Universite Catholique de Louvain Department of Computing Science and Engineering in Louvain-La-Neuve.

This program allows Linux and FreeBSD users to videoconference with industry standard H.323 applications such as the Microsoft NetMeeting program for Windows. The program has proved extremely popular with users, and is now included in many Linux distributions as well as with FreeBSD.

GnomeMeeting Features:

Support for audio codecs	LPC10, GSM-06.10, MS-GSM, G.711-Alaw, G711-uLaw

Support for video codecs	H.261-QCIF, H.261-CIF
Support for	H.245 Tunneling, Fast Start, auto answering, do not disturb mode
Register and search	ILS directories
Modify	soundcard and camera setting
Docklet	for Gnome or KDE
GUI	
International support	
Easily configurable	

Software Clients

NetMeeting and CUseeMe are two examples of software based H.323 clients. Software based clients are often cheap to implement due to the low cost of simple USB cameras and cheap microphones. As well, the software is often free which makes a software-based solution quite appealing to organizations with little or no funding for videoconferencing. The caveat is that software clients require very powerful systems to function well, and some clients don't work properly in conjunction with other H.323 systems. Software clients use the main system CPU to encode and decode video. This causes a great burden on the system, often causing choppy video or other problems.

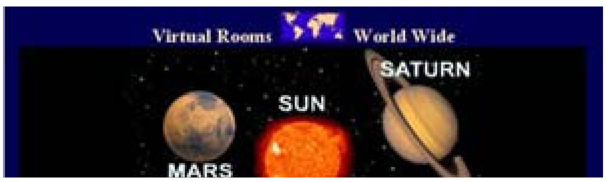
VRVS

The Virtual Room Videoconferencing System (VRVS) provides a low cost, bandwidth-efficient, extensible means of videoconferencing and remote collaboration over networks within the High Energy and Nuclear Physics (HENP) communities.

Since it went into production service in early 1997, deployment of the Web-based system has expanded to include thousands of registered hosts running the VRVS software in at least 59 countries. A set of "reflectors", interconnected using unicast tunnels and multicast. manage the traffic flow at HENP labs and universities in the US, Europe, Asia, and South America. VRVS provides versatile collaboration tools: MBone (vic/rat), H.323 (Polycom, NetMeeting), QuickTime, Desktop/Application sharing and Chat on various platforms. Scheduling of rooms is also included. Recent and ongoing developments include support for MPEG2 videoconferencing, shared collaborative environments, QoS over networks, etc. The goal is to support a set of new and essential requirements for rapid data exchange, and a high level of interactivity in large-scale scientific collaborations.

Virtual Rooms

When two or more people have a meeting, they usually join in a room or office inside a building and talk together for a while. When the participants of the meeting are far away, they join and talk inside a



virtual place that we call Virtual Room. They can be miles away from each other, but using their browsers they can all join the same virtual space (the same Virtual Room) and see and talk among them, like if they would be together in the same physical room.

Reflectors

A reflector is a host that interconnects each user to a Virtual Room, by a permanent IP tunnel. The reflectors and their links form a set of virtual sub-networks through which audio, video or data, flow. The use of the reflector technology allows the system to be highly extensible, and assures the quality needed for videoconference transmission. Participants at any location join videoconferences (in one or several virtual rooms) by contacting their "closest" reflector. In order to make an efficient use of bandwidth. Packets (video, audio and data stream) are just sent through the tunnel that links two reflectors, if there are participants in the same virtual room on both sides. In addition, the network reflector topology is chosen taking into account both geography and the bandwidth available on each network link, in order to optimize the network-connectivity paths. At the registration process, the end-user is automatically attached to his/her nearest reflector. When the user starts a connection with VRVS, the system sends a connection request to its reflector. If the reflector doesn't answer, the user is automatically connected to his/her nearest backup reflector. For more information on reflectors, please see VRVS Reflectors.

Web guides on making videoconferences, scheduling rooms, joining a conference and more are available at the VRVS site.

MPEG

MPEG-2 Videoconferencing is one of the higher quality forms of videoconferencing. Some of the bigger players in this arena are GDC, Amnis Systems Inc. (formerly known as Optivision, Inc.) VBrick, and the former Litton Network Access Systems (technology now provided by StarValley Systems http://www.starvalleysolutions.com/). Some systems are PC based with dedicated MPEG-2 video cards for encoding and decoding, while other systems are purely hardware-based designs. The conferencing systems are told what IP and port to send their video to, and they listen on another port for incoming information. After a quick handshake and some call signaling, the conference is established and both ends begin sending and receiving. Streams can be sent at different sizes, ranging from 1.5Mbps up to 15Mbps. One must keep in mind though, that more data compression means more latency. This is fine for meetings and conferences, but can be unacceptable for applications such as a band playing together over a distance. There is not yet an MPEG-2 conferencing standard, so there are still problems with vendor interoperability especially for call setup.

MPEG stands for the Motion Picture Engineering Group. It is an industry-based standards setting organization that specializes in audio and video compression and transmissions. They currently have three published standards that relate to video compression: MPEG1, MPEG2 and MPEG4. The newest, MPEG4 is on version 2, but its development is continuing. The two other standards that MPEG is working on are MPEG7 (for describing multimedia content)and MPEG21 (which is to define a multimedia framework) don't currently have a significant role in the videoconferencing world.

The following table summarizes (and vastly simplifies) the various formats and their capabilities:

	Typical Image Size	Typical Bandwidth	Max Bandwidth
MPEG1	352x240 (std profile)	1.5 Mb/s	2.5 Mb/s
MPEG2	720x480 (main profile@main level)	5 Mb/s	15 Mb/s
MPEG4	720x480 (main profile, L2)	2 Mb/s	4 Mb/s

MPEG1 is an older standard that was originally designed to compress about 30m of audio and video onto a CD. It is fairly fast to compress and decompress, but for the number of bits, it doesn't make for impressive video quality. Typically the bit rate is around 1 to 1.5Mb/s. Since H.263 compression as used in most H.323-based systems produces better picture quality with about the same processing and fewer bits transmitted, MPEG1 isn't usually a contender for videoconferencing systems.

The MPEG2 standard includes is a commonly used compression scheme for video (and confusingly it also includes a transport mechanism). There are a number of products that use this format to produce reasonable quality (close to broadcast TV) images. While it was developed for broadcast applications, vendors have found that by using standard definition video, fast encoding and decoding cards that the latencies introduced allow it to be used in an interactive environment. MPEG2 is a very complex standard that includes many variations on resolution and format (18 in all) that extend from standard definition TV to high-definition TV specifications. It is the encoding currently used in many consumer TV products like DVD players, satellite TV receivers and digital cable TV receivers.

MPEG4 is a newer standard (1999) that includes a video encoder that is more modern than the one used in MPEG2. Like MPEG2, it has a wide range of profiles that range from very low bit rates meant for wireless transmission to very high-rates for video editing and exchange. While this standard also includes many other non-video multimedia components, the compression technology is just now starting to appear in video-based products. For lower bit rates and resolutions it can probably easily replace MPEG2 and for the same number of bits provide a better image (or for fewer bits the same quality of image). This may not happen too quickly in the TV/broadcast arena since there are already so many set top boxes and DVD players in people's homes. MPEG4 requires more processing power than either MPEG2 or MPEG1 to encode and decode which means that in the short term it may not be feasible to use it for real-time interactive work. As with the other MPEG compression schemes, it was developed for broadcast or streaming applications where the latency isn't as much of an issue as it is with videoconferencing. Its application to videoconferencing may only be a matter of time as machines get faster and hardware is specialized to use this standard. The comprehensiveness of the standard, the inclusion of non-video components and its flexibility for graceful degradation as network conditions deteriorate make it a promising candidate for future videoconferencing applications.

[JVC has recently announced a prosumer grade DV camcorder that has an MPEG4 streaming feature built in. They don't specify the delay characteristics of the stream, but they do mention that a possible application might be videoconferencing. The system currently only provides for a maximum bit rate of 384Kb/s at a slow 15 frames per second and a small size of only 352x288. This is equivalent to a H.323/H.263 videoconference, but typically they have a 30 frames per second rate.]

Motion JPEG (MJPEG)

MJPEG, for 'motion-JPEG', is an encoding technique that simply jpeg-compresses each video frame before transmission. There are several advantages to this approach, and some disadvantages. The biggest drawbacks are that there are very few products that support it, and it has been seen as a failed technology of the past. Both of these are changing again, and MJPEG is making a strong comeback.

The main advantages of MJPEG are that JPEG-compression is very cheap to do in hardware, is extremely fast, quite robust, and it supports almost any size video frame you want to transmit (subject only to an 8x8 tiling restriction, so images may need to be padded in software). This means you can use MJPEG for video frames from sub-QCIF size up to full-size 1080i-HDTV (1920x1080) and beyond.

Because the codecs are so fast, and there is no inter-frame compression used, MJPEG has the lowest latency of any codec currently on the market. This low latency is great for more natural interaction than say H.261, H.263 (the most common H.323 codecs) or MPEG. In general, MJPEG systems also perform much better for lip-synchronization than most H.323 or other codecs. Why this should be so is unclear, but the lower codec latencies do allow for greater flexibility in the decoder and playback setup.

Because MJPEG uses the normal JPEG tiles, and no inter-frame information, errors or packet loss on the network will only impact a tile, or row of tiles, in an image, and the error does not propagate for several frames. This makes it a very sensitive network-testing tool!

A common approach with MJPEG codecs is to give users two choices of compression. One sets a fixed compression level, or "Q-value" in JPEG terminology, for the DCT codec and just compresses each video frame and transmits it. While this provides a very stable and high-quality video stream, it can use up bandwidth very quickly, especially if there is a lot of fine detail ("high frequency data") in the images. An alternative approach allows one to set a bandwidth cap, and the Q-value is allowed to vary frame-by-frame to meet the bandwidth cap. This means the image quality will vary somewhat from frame-to-frame. For normal videoconferencing situations where there is little change though this is not perceptible.

Since MJPEG compresses on a frame-by-frame basis it automatically picks up the frame rate of the raw video, so there is no issue with handling NTSC, PAL, SECAM or any other codec, all in the same system. Conversely, some products allow you to specify a particular frame rate (as a fraction of the source stream) and again use the Q-factor versus bandwidth arrangement above to fine-tune the performance of a video-link.

Another advantage for MJPEG-encoded material is that it is always frame-accurate, which makes editing a lot easier — unlike inter-frame compressed material where you need to work from keyframes, or decompress and recompress intermediate frames with its attendant loss of quality. A variety of software video-editors use MJPEG for on-disk storage for this reason, only converting to MPEG or other format when writing to tape or VCD/DVD.

DCT-based MJPEG is not for the narrow-bandwidth-minded. A PAL/NTSC signal at 25/29.97fps looks very good at 2Mb/s. Much below that bandwidth the image rapidly becomes blocky, as the Q-factor has to be driven to lower values. Typically you can improve the picture dramatically then by going to lower frame rates, but that will only work so far. For bandwidths below 2Mb/s at PAL/NTSC resolution you would be better advised to go to a codec such as H.263 or MPEG, or go to a lower resolution. Above that however MJPEG truly shines.

A more recent change is the support for the newer JPEG-2000 standard. This has moved away from the DCT compression used in "old" JPEG, and also H.261/H.263, MPEG and others, towards "wavelet" compression of image data. The compression achieved is significantly better, it has more flexibility (e.g. you can identify regions of frames that are more important than other regions, and so get a better treatment from the codec), it has better error-resilience (one of its applications is in wireless video transmission to low-bandwidth devices) and you can embed metadata more easily with the image data. At this stage, support for MJPEG2000 is still in its infancy, and all done in software, but it has some significant potential. Sample material at 64-256kb/s looks very impressive, and even 32kb/s CIF is quite comfortable.

Products that support MJPEG include ATM-based codecs such as Marconi (ex-Fore's) StreamRunner products, PC cards such as LinuxMediaLabs' LML33, or software such as Pegasus Imaging Corp's (www.jpg.com) PICVideo products and Morgan Multimedia's mjpeg (and mjpeg2000) software. Try entering 'MJPEG' into your favorite web search engine.

The Access Grid

The Access Grid (or AG) is a project that has grown out of its roots at Argonne National Labs. You can get a lot of details at its main web site: www.accessgrid.org, but a brief summary will give you the flavor.

The fundamental ideas behind the AG are that it is "always-on", i.e. you can walk in at any time, that it uses IP-multicast as its transport, so there is no MCU involved and so it can support a large number of simultaneous sites, and it is based around openly (typically free) available standards and open-source as far as possible. It does not, of its own, support gatewaying into other systems such as H.323, H.320, VRVS and similar, although these functions can be achieved through related projects. It also provides mechanisms to share data, such as slides and web pages, and some high-end AG sites support virtual reality display systems.

The rooms used for AG meetings are much the same as a high-end videoconference room, and so rooms can be used to support AG activities as well as H.323 activities. A "standard" AG room (or "node") provides a large display system, capable of ~3000x1000 pixels. Usually this is achieved by using 3 data projectors. An echo-cancellation system is used to connect multiple microphones in the room and provide a clean audio signal to the encoder. Multiple cameras at one site are supported simultaneously — so that any site can transmit as many streams as suit their needs. This is a benefit over many H.323 systems which really only support one video stream. However, multiple outgoing streams, and the use of multicast, means that the bandwidth usage scales up with the number of sites involved. For example, 25 sites each sending 4 video streams result in each site receiving 100 streams, at the full transmitted bandwidth. Even using the same video-codecs as in H.323 (e.g. H.261) at only 256kb/s the bandwidth quickly surpasses 30Mb/s. Higher-bandwidth codecs are under investigation. This is not for the faint-of-heart, or narrow-of-bandwidth!

An AG node typically uses 3 or 4 dedicated desktop-PC computers, mostly running Linux. In general most sites have settled on a set of hardware for video capture and display and audio capture and emission, which have respectable driver support in the Linux community. However, given the ever-increasing range of audio/video hardware, there are always groups tinkering with the latest hardware.

The experience of the AG though is very distinctive and worth the effort. While much of it is still at the research stage, and doesn't have the commercial polish of most H.323 products, it provides a very impressive videoconferencing environment, especially when the number of sites involved increases beyond a handful. Events with 100 sites have been held — and one can see all other sites at the same time. This is totally different to large-scale MCU-cascades in H.323 where one is aware there are many sites connected, but you can only see a few at a time. If you can afford the bandwidth, the equipment, and the bleeding-edge feel, an AG node is certainly worth experiencing.

Web Clients

Web clients offer the hope of simplicity to end users of videoconferencing. The idea of simply pointing the web browser to a site and almost instantly being able to connect to another user via the browser is a compelling and simple idea. However, the results leave a lot to be desired.

The clients seem to be fairly easy to install. Assuming you have a working webcam installed on your computer, the ActiveX controls or plugins detect the camera and configure things fairly automatically. However, these products are generally not standards based and they can only talk to other instances of the same software.

The best use of these products may be to create online groups. Online support groups, user groups, group chats, etc can be supported by the products (although most currently seem to be point to point). Online technical support is another promising application, useful in helpdesks and other support functions. Adult uses are also one main use of these technologies at the current time. In the R&E environment, the only appropriate use of these technologies might be to support student faculty meetings or student group interactions in distance learning environments.

Leading products in the consumer space are FocusFocus.com, Lotus's SameTime, and FVC's Click to Meet are aimed at the corporate market. A good listing of products for web-based conferencing is available at http://www.strom.com/places/wc.html. Note that some of the listed products only support voice or data collaboration, not video.

This is a market that may change rapidly over the next two years, so check the web for new products and solutions.

Emerging Collaborative Technologies

[Application and Data Sharing](#)
[Teleportation Technologies](#)

Simple Internet Protocol (SIP)

SIP stands for Session Initiation Protocol, and is a part of the IETF standardization process. SIP is a signaling protocol for establishing calls and conferences over IP networks. The session setup, change or tear down is independent of the type of media that will be used in the call; a session may include different types of data, including audio, video, and many other formats. SIP originated in the mid 90's (about the time H.323 was becoming finalized as a standard) so that it would be easy to invite people to view an IP multicast session like a shuttle launch on the M-Bone. Those who are most excited about SIP believe this development is as significant as the HTTP protocol, the technology behind web pages, that allows a single page with clickable links to connect you to text, audio, video, and other web pages. This possibility is behind SIP's rapid adoption as a Voice Over IP standard.

SIP is modeled after other Internet text-based protocols such as SMTP and HTTP, and was designed to establish, change, and tear down calls between one or more users in an IP network in a manner totally independent of the media content of the call. Like HTTP, SIP moves control over the application to the endpoint, eliminating the need for a central switching function.

An excellent resource site for learning about SIP has been developed at the Computer Science Department at Columbia University http://www.cs.columbia.edu/sip/ . Information about the SIP IETF working group can be found at http://www.ietf.org/html.charters/sip-charter.html

SIP Architecture

Main components of the SIP architecture are:

1. SIP User Agent

   The User Agent is the SIP endpoint or end-station software. The User agent functions as a client when initiating session requests, and also acts as a server when responding to a session request. Thus, the basic architecture is client/server in nature. The User Agent is "intelligent", in that it stores and manages call state. The User Agent places calls using an email-like address, or a telephone number (E.164). As an example: SIP:user@university.edu . This makes SIP URL's easy to associate with a user's email address.

2. SIP Server
   a. SIP Proxy Server

      One type of SIP intermediate server is the SIP Proxy Server. Proxy Servers forward requests from the User Agent to the next SIP server, and also retain information for accounting/billing purposes. In addition, the SIP proxy server can operate with stateful (i.e., circuit-like) or stateless (i.e., TCP-like) communication. The stateful SIP server can "fork" incoming calls so that several extensions are rung at once and the first to answer takes the call. SIP proxy servers can use multiple methods to try to resolve the requested host address, included DNS lookups, database lookups, or relaying the request to a "next" proxy server.

   b. SIP Redirect Server

A second type of SIP intermediate server is the SIP Redirect Server. The Redirect server responds to the User Agents request by providing information about the requested server's address so that the client can contact that address directly.

The role of these SIP servers is to provide name resolution and user location. The combination of Proxy and Redirect servers gives SIP great architectural flexibility; the user can employ several schemes, simultaneously, to locate users. The SIP architecture is particularly well suited to support mobility.

3. SIP Registrar

    The SIP Registrar provides a location information service; it receives information from the User Agent and stores that registration information.

The SIP architecture makes use of **SDP** (Session Description Protocol). SDP was and early IP Multicast conference tool developed to describe audio, video and multimedia sessions. In fact, any MIME (Multipurpose Internet Mail Extension) type is supported, similar to e-mail's ability to support all types of message attachments. The session description can be used for negotiating an agreed upon set of compatible media types.

As a result of this architecture, an end-user can request to establish a session to another user by an email-like address; using SIP Proxy and Redirect Servers and Location Services that exist within the IP network cloud, information about that remote users location can be used to establish a connection. Even when the remote user is mobile, proxy and redirect can be used to forward the connection request to the user's current location. Sessions can involve multiple participants, similar to a multi-point H.323 call. Communications within a group session can be via multicast or a mesh of unicast calls, or even a combination of both.

SIP Architecture Supports New Types of Services

- A type of "call forwarding" allowing users to specify where they are so that incoming calls can be forwarded there, or to chose to forward calls to "voice mail" or any other answering service.
- Call participants can manage the call; this allows one or more callers to decide to bring in a new call participant or to cancel a connection in the call.
- Ability to return different media types allows an incoming call to be answered by a web page with information that can be used to complete the call.
- "Presence" information — the User Agent can be used to indicate whether the user is present (available to take the call) or absent (not able to take the call).

SIP In the Marketplace Today

There are a number of commercial SIP implementations available in the marketplace today, mostly software telephones or other voice over IP products, including a line of SIP architecture products from Cisco. However, the most significant market impact will come from Microsoft, who has announced plans to stop all H.323 development and move exclusively into SIP product development. The Windows XP Operating system comes with a SIP User Agent built in. It is called "Windows Messenger" and it turns your PC into a software telephone (a voice over IP device), with the added features of video, chat, data sharing and more.

Network World Fusion conducted an interoperability test of Windows Messenger in January 2002, registering the Microsoft client with a Synamicsoft SIP Proxy Server and placing calls to and from a

Pingtel xpress IP phone. The calls were not only successful, but voice quality was reported to be business-quality.

Microsoft has designed its Windows Messenger so that users are required to register with Microsoft .NET.

Relationship of SIP and H.323

Both SIP and H.323 are standards for call routing, call signaling, capabilities exchange, media control, and additional services. H.323's strength has been its interoperability with the Packet Switched Telephone Network (PSTN) and availability of affordable and dependable desktop and room videoconferencing systems/appliances. SIP is a protocol developed specifically for the Internet and promises greater scalability and flexibility. H.323 is likely to remain the predominant conferencing technology for the next 2-3 years, with SIP coming into greater use at that point.

Interoperability with H.323

Standards organizations are already working on SIP-H.323 interoperability, promising the possibility of a reasonable transition period between H.323 and SIP technologies. Two organizations especially interested in this topic are the IMTC (International Multimedia Telecommunications Consortium, a non-profit corporation of 100+ organizations around the world http://www.imtc.org/) and also the ETSI (European Telecommunications Standards Institute http://www.etsi.org/). The Open H.323 Organization (http://www.openh323.org/) has already released a working H.323 to SIP gateway.

Wireless and Satellite Video

Wireless local area networks are now becoming common, and so the question arises as to whether videoconferencing can be done over such networks. The answer is a qualified yes. Video and audio are more demanding of network quality than are simpler things like email and web browsing. Dropped packets and network traffic competition cause video and audio to fail long before email and web browsing. The net result is that the wireless distance that can be achieved for video is significantly less than that for other applications.

Another factor that works against wireless video is that wireless networks are completely shared by all users, as if dumb hubs were being used on a wired network. Most wired networks today use switches, which isolate the traffic from one station to another. Thus if there are many users on wireless network, then video will degrade faster than other applications.

Videoconferencing is now being done via earth satellites, with significant success. But it suffers from the same effects as wireless video, and a few more. The latency is inherently greater due to the speed of light delay to and from the satellite. This is at least 1/2 second. However most users quickly adjust to that effect and then ignore it. Other issues are that the satellite data rates may be different for uplink than for downlink, making symmetrical videoconferencing more difficult. The satellite may also operate in some form of store-and-forward transmission mode, making the jitter larger.

Application and Data Sharing

Developers have put a great deal of emphasis on the audio and video aspects of collaboration. With the benefit of standards, most of these products can interoperate with each other and the use of these videoconferencing endstations is becoming more widespread. This means that people are trying to merge videoconferencing more and more into their daily routine. In doing so, the demand for sharing data and applications (as in a normal face to face meeting) is on the rise. This might be the usual Word document or PowerPoint presentation. Stretching the limits in our research and education environments, it could also be a complicated MRI scan or even the output of an animated simulation.

Such sharing can be viewed in a number of ways. We can speak of sharing the application (such as running a shared version of Excel or Matlab) or of sharing the data under distributed versions of the application. We can look at what functionality the tool may provide such as whiteboard, chat, file transfer, imaging, polling/voting, animation, or remote control. We can look further at features such as standards compliance, open source, scalability, specification of lead person, scheduling, authentication, and security. Is the tool fully distributed or does it require a central server through which all data must pass? And then there is perspective – what perspective is given to the end user? Are they sharing a document or their office? Are they sitting at their desk or are they joining others in a room or large conference hall? Is the sharing one-way or two-way?

Early tools (1990's tools) had the benefit of the T.120 standard. T.120 defines how to establish and manage interactive communications between two or more participants' desktops. It also defines these facilities over a number of different network protocols in a manner that enables data communication services independent of the underlying network. It provides support for applications, defining the startup mechanisms, exchange of capabilities, control or conference lead, etc. It also defines general, high level interoperability of commonly required functions such as file transfer and still image exchange (including whiteboards) through standard protocols. But several issues challenge the future of T.120:

- In combination with H.320 and H.323, while the data stream is prioritized it receives in the range of 6-20 KBPS of the available call bandwidth making the sharing of complex or dynamic documents very difficult and sometimes impossible.
- Only a view is shared. The application runs on the originating workstation (the workstation of the person who "owns" the data) and only a view of that window is shared. Changes in software licensing may prohibit this form of application use in the near future.
- The rich functionality of the Web offers a growing number of competing ways to implement sharing of resources, including your personal ones.
- Maintaining standards is difficult and development under them is costly.

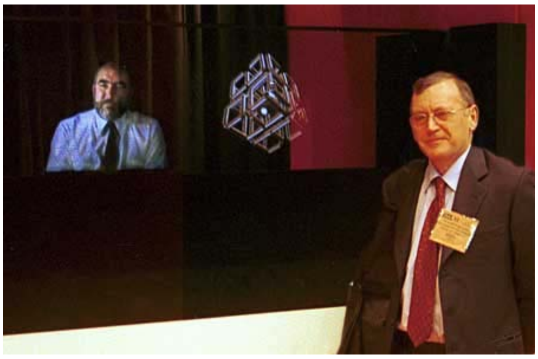
Standards experts speculate that T.120 is loosing attention as developers see more opportunities with web enabled tools. These web tools provide new perspectives with more attention given to how people actually work together in person. In fact, this very thing seems to be the competitive edge that is driving development. But this raises some concerns for the higher education environment. Some of these tools are expensive. A corporation is able to establish its own "corporate standard" by selecting a specific tool for deployment across its entire organization. Institutions of higher education are more prone to independence, limited funding opportunities, and diverse purchasing guidelines and requirements. Determining how to equip our faculty, staff, and students to collaborate across our institutions is going to be challenging.

ViDe's Data Collaboration Working Group has undertaken a project to follow and analyze the abundance of tools on the market today. We encourage you to track our progress and to participate in these studies as your time permits.

Teleportation Technologies

Several companies have developed systems which "teleport" a 3-D, full-sized holographic image of a person into a room. The virtual image of a real person can be placed behind a lectern or desk, in front of a wall, curtain, or other background. This virtual person can see the person or audience in front of him or her and can make eye contact with them. They can hold a normal dialogue. This form of virtual presence makes it seem as if the whole person is right there in the room.

Teleportec Teleportation Technology.



Reality Interface's Teleportation Conferencing System

High quality audio and video components eliminate audio latency problems. While the systems operate over ISDN, ATM, or IP, the unique features require specialized lectern and desk equipment.

Basic Requirements for Successful Videoconferencing

Basic Components

As discussed in our introduction What is Videoconferencing?, any videoconferencing terminal must have a few basic components to "get the job done": a camera (to capture local video), a video display (to display remote video), a microphone (to capture local audio), and speakers (to play remote audio). In addition to these more obvious components, a videoconferencing terminal also includes a codec ("COmpressor/DECompressor"), a user interface, a computer system to run on, and a network connection. Each of these components plays a key role in determining the quality, reliability, and user-friendliness of the videoconferencing experience as well as any given videoconferencing terminal's suitability to particular purposes. A basic understanding of each of these component's roles will help you map videoconferencing technology capabilities to your specific application needs.

The Main Camera

By nature of the general definition of videoconferencing, at least one video source is typically present at each endpoint. The most common video source is a single main camera that captures live movement occurring at one end so that it may be sent to the other end in near real-time. ("Near real-time" is an important concept in the success of a videoconference and is covered more in the sections below on the codec and in our later full section on the network connection.) The detailed characteristics that distinguish one camera from another are a topic in and of themselves and cannot be fully covered here. However, when selecting a camera for videoconferencing, it is important to understand that the quality of your camera heavily determines how your video will appear to the receiving end. It is often our first reaction to attribute video quality to the receiving system - (i.e.," Why does their picture looks so bad when we spent $XXXX on this system?!" Yet, if you cannot see the other site clearly, their camera is quite often the culprit.) It follows that, when evaluating camera quality, you need to be sure you are shown how your image will appear to others. In addition to image quality, cameras vary in terms of other features that will affect both their usefulness and their cost. Among these are: the ability to pan, tilt, and zoom, wide angle versus narrow angle lens, manual focus versus auto-focus, manual iris versus auto-iris, auto-tracking, remote control, and/or RS-232 control. Naturally, as features are added, cost goes up. Considering the impact of the main camera on the success of a videoconference, it is extremely important to imagine ahead of time how the camera will actually be used (i.e., room setup, number of participants, user temperament, etc.) and then ensure that the selected camera can support those uses.

The Video Display

In addition to capturing local video, a videoconferencing solution must include the ability to display the remote video that is being received. This incoming video is displayed on a monitor, most often a computer monitor, which influences how clearly the remote site can be seen and also how many people at the receiving site can easily see it. "Typical" display monitor quality considerations such as screen size and resolution affect the size and clarity of the incoming video window and also the integration of the incoming video window with the application interface that surrounds it. The quality of the image within the video window itself is, however, more directly related to the performance and capabilities of the codec and to the quality and bandwidth of the network connection. In the case of a desktop videoconferencing terminal, most offer a scaleable video window that shares space on a PC desktop with other program/application windows. In such cases, the conference aspects most heavily influenced by the capabilities of the computer monitor are the appearance of the video window itself (not what is inside it) and the ability to manipulate that window within the larger display. In some cases, an entire display monitor can be dedicated to displaying incoming video (a "full screen" conference) while a second monitor is added for call control and data sharing. A final note: Video resolutions supported by H.323 are CIF (352 X 288 pixels) and QCIF (176 by 144 pixels). Since these resolutions are fixed, increasing the network bandwidth of a call beyond a certain point will not show an appreciable difference in video quality within any given video frame. However, additional bandwidth enables higher frame rates (i.e., the sending of additional video frames per second), which can have dramatic improvements on the smoothness and video quality of motion.

Audio Components

Within a videoconference audio is as important, and often considered more important, than video. If we lose video or experience poor video quality in a conference but audio remains intact, we can still accomplish many of our communication objectives. The conference would simply become a teleconference rather than a videoconference. In contrast, poor or disrupted audio quality effectively shuts down a videoconference, often sending participants scrambling to find a "native audio" telephone to complete the meeting. In light of this, the devices that capture local audio (microphones) and those that reproduce remote audio (speakers) are critical conference components. Coupled with this are characteristics associated with comprehensible full duplex (simultaneous two-way) transmission of audio, such as echo cancellation, noise suppression, and audio mixing. These features are influenced by a combination of the microphones, speakers, and codecs. Similar to the camera discussion, it would be impossible to cover all features of audio performance here. However, one key to ensuring audio that supports conference requirement and expectations is to examine the location, quantity, and quality of your microphones and speakers. If cost is an issue, using a quality handset or headset can offer better results than a comparably priced microphone-speaker arrangement. Again, as features are added, cost goes up, though the cost differences may not be as pronounced as they are in camera selection. Since hearing is often the best test, you may want to speak and listen before you buy!

The Codec

The codec has been mentioned above as affecting both the video and audio within a videoconference. Indeed, the codec actually forms the heart of any videoconferencing terminal and is the main enabler of wide-scale videoconferencing. The word "codec" is a shortened version of "compressor/de-compressor" and is specifically applied to the wide variety of algorithms used for actually compressing or decompressing audio and/or video information. This compression has historically been necessary to make the audio/video data "small enough" to be practical for sending over expensive network connections. In this sense, there are many audio and video "codecs" (particular compression/decompression methodologies) that are supported as part of the H.323 videoconferencing standard (see the appendices for The H.323 Standard). For the purposes of this section, we are considering a broader meaning for codec: the codec as the portion of the videoconferencing terminal that is responsible for whatever compression/decompression of the audio/video signals is taking place.

This latter and broader definition allows for the codec to be either a software or hardware component, and confers great responsibility upon the codec for the success of the videoconference. The amount of data required to "describe" audio and video in a digital format is very large by today's data networking standards. Without some form of codec, the transmission of a videoconference requires extremely high amounts of network bandwidth. It is the codec that takes the sights and sounds captured by the local camera and microphone, and then compresses that information such that it may be transmitted across a network fast enough to enable near real-time communication. When the compressed information is received at the remote site, the codec within the remote site's videoconferencing terminal decompresses it and enables "play back" through the speakers and display. Though we think of the conference as a real-time conversation, the real-time feeling is a function of how fast each of the codecs are compressing/decompressing the data, and how fast and reliably the compressed data is traveling back and forth across the network. In light of this, some factors to consider when evaluating codecs are:

- **Is the codec a software or hardware component?**
  Hardware codecs are generally faster in completing their compression/decompression task, making near real-time communication more likely. Hardware codecs also often carry their own processing power "on-board" such that they do not rely on the resources of the underlying system. For instance, in the case of a desktop system, using a hardware codec may mean that you don't need a "souped-up" PC, or that you will be able to run other applications on your PC while simultaneously participating in a videoconference. On the other hand, software codecs are generally less expensive and easier to install (no special hardware required), but they tend to produce lower quality ("casual") conferencing with very low frame rates. In H.323 desktop videoconferencing systems, the codec typically resides on an interface board or in a software application. In H.323 group conferencing systems, the codec is most likely an interface board itself (you buy the PC) or is part of a turn-key system that is possibly proprietary but most likely PC- based.
- **What actual audio and video codecs (compression/decompression methodologies) does the more broadly defined "codec" support?**
  In order for a successful videoconference to take place, endpoints must be able to negotiate a common methodology for both audio and video exchange. Any given video terminal/codec (using the broader definition) may support a number of audio/video codecs (the narrower definition), some of which must be supported for a videoconferencing terminal to be considered "H.323 compliant". A video terminal/codec may also support proprietary audio or video codecs of the system developer's own design. When two of these video terminals are in the same videoconference, they may have access to improved functionality, quality, or reliability between them because they can each understand and use the proprietary features. When selecting a videoconferencing terminal, you should be aware of its range of support for various types of audio/video compression. You then need to consider whether or not this range covers the range you are most likely to encounter in your videoconferences.

The User Interface

All systems that are meant for use have a user interface. The friendliness of the user interface largely determines whether the system is embraced by end users, or left to be grudgingly approached on an "only-if-I-have-to" basis. The implications and importance of the user interface may easily be overlooked or taken for granted if the main functionality of the system is complex or interesting to the point of distraction. That may be the case with videoconferencing. Often we consider and compare videoconferencing terminals based solely on video and audio quality -- what it looks and feels like when we are actually in a conference -- and we don't necessarily stop to consider other features of the system. These other features may determine how we get into and out of conferences, what we can do in conjunction with a videoconference, and even what we know about how the call is going or what we have documented about the call once it's over. A sampling of specific features and considerations are listed below, some of which have already been touched upon and others that are addressed in greater detail in sections that follow:

- **How the video terminal application "works and plays" with others.**
  Is the system easy to install, de-install, etc. How much system capacity does the videoconferencing application use? Can other applications run comfortably and reliably when the videoconferencing application is running and in use? Is a wide range of system performance acceptable, or are system requirements stringent? Has the videoconferencing application been tested for interoperability with other H.323 terminals?
- **The "Dial" menu, or placing and receiving calls.**
  Is there any easy to access Phonebook for keeping track of frequently called numbers in a user-friendly way? Is there an automatic call log available for call history and/or error tracking? Can the data rate (call bandwidth) be selected for particular calls in a way that is easily understood?

- **Application sharing and data collaboration.**
  Are these features fully integrated into the videoconferencing application or are they provided using a "helper" application (e.g., NetMeeting) or perhaps not available at all?
- **Interaction with audio/videoconferencing devices.**
  Can a wide variety of audio and video devices be used with the terminal application or are only certain devices supported? Are inputs and outputs other than cameras and monitors supported (e.g., VCR in or out?) To what degree can audio/video features (e.g., volume, echo, color, brightness) be controlled from within the application? Is there support for the use of alternate or enhanced devices (e.g., Far End Camera Control, dual monitors, telephone handsets for privacy?)
- **Support for the H.323 standard.**
  How compliant is the video terminal with the current H.323 standard? How prepared is the terminal/developer/vendor to support future H.323 versions and directions? Does the video terminal make any concessions now to cover potential functionality gaps in the current H.323 standard? (e.g. user authentication, secure gatekeeper registration?)

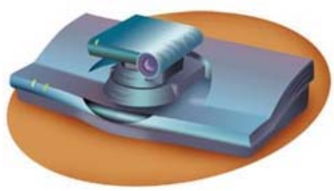
Though this checklist only provides a partial glimpse into the very volatile area of H.323 videoconferencing terminal development, it should prove useful as a starting point for the very important task of evaluating the user interface.

The Supporting System and the Network Connection
Though the supporting system and the network connection are not technically part of the basic components of a videoconferencing terminal, they have a definite effect on the terminal's perceived performance. To understand more about the influence of each of these, please see the sections Network Requirements and Selecting and Tuning Your PC

Examples

The VCON Vigo and Polycom ViaVideo are examples of hardware solutions for PCs. They connect to the PC via USB and have their own onboard video codecs that help reduce CPU usage on the PC. Such systems are often deployed as a desktop solution.

The Polycom ViewStation and VCON Falcon IP are examples of purely hardware- based solutions. The only additional hardware required to use one of these systems is a television. Completely self contained, they are excellent choices for conference rooms and classrooms.



Vendor Relationships

Selecting a vendor

The first step in vendor selection is a ***survey of the market,*** looking at existing technology, new standards and emerging technologies, customer deployment of current technology, customer satisfaction surveys, the experiences of your colleagues, and of course the web pages of all the videoconferencing vendors identified in the market survey. Even in cases where a vendor has already been pre-selected, such as in a statewide videoconferencing initiative, knowing the available technology and offerings of competing vendors is invaluable for working with the pre-selected vendor. In your market survey, concentrate on articles and web sites that survey the market, evaluate existing vendors and technologies and that predict future enhancements for videoconferencing technology. Sign up for electronic discussion lists that include users of videoconferencing technology. Ask questions about different vendors and their offerings. Sources for information include:

- *Websites*
  Look at general videoconferencing information, such as this website, and also vendor-specific web sites. Look at vendor press releases, generally posted at vendor web sites, for the past year to six months, to get a feel for development patterns in videoconferencing and for each vendor.
- *Discussion Lists*
  Subscribe to any electronic discussion lists where videoconferencing users discuss their experiences. Some vendors may host discussion lists for their customers.
- *Conferences*
  At conference booths, talk not just to vendors but also to other users spending significant time at the booths. Discover their level of expertise and deployment plans, and be sure to exchange cards. Current customers often take advantage of conference booths to discuss issues and concerns with vendors. This is a great way to develop a reference list independent of the list supplied by the vendor.
- *Journal Articles*
  Survey articles that compare functionality and possibly customer satisfaction are critical, but don't neglect predictive articles looking at future enhancements. A clear migration path for future technologies is critical for more expensive purchases.
- *Colleagues*
  Contact any colleagues that have purchased videoconferencing systems. Ask for copies of any purchase information that your colleague is able to share, such as RFI (Request for Information); RFP (Request for Proposals); purchase order outlining specifications and service requirements; and contract. For government agencies, after bids are awarded, this is generally freely available or, at worst, available through the Freedom of Information Act. Ask for your colleague's experiences with the selection process. Which vendors were eliminated, and why? Ask for experiences with installation and deployment. What level of assistance was provided with set-up and initial troubleshooting? Did the product perform as described? What ongoing issues and concerns are your colleagues experiencing?
- *Manufacturers*
  Contact references provided by the manufacturers of videoconferencing products of interest to you. Ask for customers with a large number and variety of conferencing needs. Manufacturers will, of course, give names of customers that are happy with their products, but this will also let you know why they are happy with that particular manufacturer.

After a market survey to familiarize yourself with the state-of-the-art for videoconferencing, it is necessary to select the functionalities that are both critical and desirable for your project and codify those functionalities into a ***purchasing document,*** whether an RFI, RFP or a purchase order. If at all possible, design an instrument that can be sent to a large number of vendors.

Be sure that the bid section will result in competitive pricing that can be compared uniformly across vendors. A good practice is to provide a bid sheet with individual line items for each meaningful system component. Meaningful system components vary by project and are best determined by the individual institution, after an extensive market survey. These line items can include the entire system (hardware & software); individual line items for component pieces (i.e., MCU, terminal software, etc.); and line items for services, such as installation, training, and ongoing maintenance. Be sure the bid component includes price breaks for item multiples, such as terminal software, cameras, etc. It is critical to request information about warranty and maintenance costs. One often overlooked pricing differential is warranty period, with some vendors offering three months and others a year or longer. It is common to require multi-year bids on ongoing maintenance costs for large-scale purchases, to insure that an organization is able to financially maintain a selected system over time and to insure that vendors do not offset low purchase costs with high maintenance pricing.

Although a purchase is implied, be sure to include language that states that the organization you represent is not required to issue a purchase order in response to bids received.

Distribute your purchasing instrument to the widest possible vendor pool. You will probably work closely with your institution's purchasing department, but do not rely solely on their list of identified vendors. Supplement that list with the vendors you discovered in your market survey.

Your purchasing instrument should require the names and contact information of all customers similar in size and mission to your institution. Do not ask for selected customers, but the complete list of customers meeting your description. A critical component of the selection process is the checking of references. Be sure to ask standard questions of each reference, for comparison purposes, as well as open-ended questions about their experiences. Many vendors may not provide a complete list, even though it is requested. If necessary, ask the contacts provided what institutions or companies they contacted for references, and expand your reference pool in this manner.

Contractual Issues

Depending on your institution and the nature of your videoconferencing project, no contract may be required, or there may be a purchase contract and a maintenance contract. For expensive projects, where expense includes not just the purchase itself but the staffing and training required for deployment, a purchase contract is a good idea. A purchase contract can provide the *financial and risk protection.* If the contract includes innovations not yet available, the purchase contract can outline staggered payments for scheduled deliverables. The contract can also define financial performance incentives for functionalities that are very new or that do not perform as specified, particularly if you select a vendor for very good reasons in spite of concerns expressed in reference checks about the performance of certain functionalities or problems with ongoing troubleshooting and support. Most vendors have an honest desire to serve customers well, but they are frequently understaffed and focused more on generating new business than on support for existing customers. Financial incentives (also known as financial penalties, when the vendor steps out of the room!) are an effective way to insure service and minimize risk, particularly for very new technologies. Vendors are more likely to agree to financial incentives for performance for large, expensive projects and for projects that will be heavily promoted by the purchasing institution.

For government entities, which of course include state universities, financial penalties can be tricky, but not impossible. Steep reductions in ongoing maintenance costs, free extension of the warranty period, payment in free enhancements, free additional terminal software; etc. can usually be worked out with your contracts department as well as with the vendor. The goal is to avoid enriching the coffers of your institution's "general fund" (which might go toward the purchase of uniforms for the football team), and instead to impose performance penalties that directly compensate your videoconferencing project.

Financial protection can and should include price caps for ongoing maintenance and should, if at all possible, lock in prices for enhancements that are part of the purchasing instrument response (and thus the contract), but not yet available for purchase. In terms of *innovations,* a contract is a good place to negotiate for functionalities requested in the purchasing instrument which the vendor is willing to develop but unable to currently supply.

Vendor Services & Support

Vendor services include installation support, technical documentation, ongoing troubleshooting and maintenance, and upgrades and enhancements to current service.

Installation and troubleshooting support can include any or all of the following:

- technical documentation.
- onsite installation (generally for an additional fee.)
- telephone and web-based support. Be sure to determine turnaround time for both telephone and web-based support. Ask about escalation routines and the vendor's help desk responsiveness standards before purchase.

The best installation and troubleshooting support includes preventive support. At a minimum, vendors should test all non-bundled and third-party hardware and software and provide a list of compatible products. Using tested, compatible products is the surest way to avoid installation and ongoing performance problems.

This is a particular concern where PC's are used to host videoconferencing products. Be sure that the vendor points out information on compatible equipment configurations, including:

- number of processors allowed
- operating systems that are compatible
- video cards that are compatible
- speed of processor that is required
- additional equipment requirements (cables for PTZ cameras, NT-1 for ISDN lines,
- IMUX's for multiple ISDN lines,...)

Technical documentation can include:

- paper or CD-ROM manuals.
- web-based documentation.
- web-based FAQ.

Manufacturers usually leave support up to the vendor (seller) and use the vendors as front lines of support on the first tier of troubleshooting. Ask the vendor or manufacturer for the support number at the manufacturer so that higher-level support calls can be initiated by you, the customer. The manufacturer's support line will almost always have better information on their products than the vendor's will. Maintenance services generally include several levels of service. Be sure the vendor specifies what levels of service or what enhancements can be provided. These can include:

- telephone support
  Be sure to ask if they have a toll-free number. They may not point this out up front.
- web-based support
  Is online chatting provided for support?

- on-site service
  How many on-site service calls do you get?
- free upgrades and enhancements
  Are the upgrades and enhancements software only, or is hardware included?
- software upgrades and patches.

Training can include:

- on-site training, at point of installation.
- fee-based training on demand.
- a training manual.
- interactive tutorial, frequently on CD-ROM or on the web.

Many vendors are not prepared to provide extensive training, but for large-scale projects where, for example, you will install videoconferencing services for 50 faculty members, training will be a critical component for the success of your project. If you are a project manager or an engineer, but not a practiced trainer, consider contracting training to a computer training firm specializing in technology transfer to non-computer professionals. The videoconferencing vendor and the computer training firm can be contractually required to work together to develop a training package. Be sure to specify who owns the training materials developed, which includes training manuals, course outlines and lesson plans. Be sure that either your institution owns the training materials or that the training firm and/or the videoconferencing vendor recompense your institution for the training materials developed at your expense if they wish to reuse them. We strongly recommend that large projects, particularly involving large numbers of end users, should include a significant training component.

Ongoing Maintenance and Upgrades/Enhancements

Before purchasing a videoconferencing system, be sure to identify, possibly through a non-disclosure agreement, any anticipated enhancements scheduled for release in the next six to fourteen months. In particular, pay attention to operating system changes in the market. Is the operating system you are using now changing to a new version that is not compatible with the current or future upgrade of the videoconferencing product? Will a hardware upgrade be required to move to a newer version of an operating system or a new operating system altogether? Will you be required to pay to replace your hardware if this is the case? More than likely, in such instances, you will have to bear this cost. If you include any planned enhancements in the purchase contract, be sure to minimize the risk to you contractually.

Be careful in a contract to negotiate only for enhancements to current functionalities and not replacement functionality that would result in the purchase of a non-standard current product. You do not want to risk ongoing problems with new releases and upgrades that will not interoperate with your nonstandard product. If current functionality requires re-working to customize service for your institution, you are probably buying the wrong product.

If you identify a significant enhancement to service that you contract with the vendor to develop, be sure to use the purchase contract or another contract instrument to spell out the specifications and the financial incentives for completion. If your institution's involvement in designing and testing the enhancement will be significant, consider a joint marketing venture, or at least a substantial innovator's discount for the purchase and ongoing maintenance of the enhancement. Make sure all joint venture or pricing arrangements are clearly established in the contract.

Add-on Components - Enhancement Software and Other Peripherals

Understanding the basic components of videoconferencing is a necessary first step in planning your use of the technology. Understanding how these basics may be supplemented or enhanced is a critical next step to ensure a successful application match. The following "add-ons" are typical of the changes that can be made to a basic configuration.

Input Devices or the Video Source

The video source, typically a main camera, can be varied according to the type and number of sources that are used. In varying the type of source, countless possibilities exist, really only limited by our imaginations and the readiness of the supporting technology. Some commonly used video sources include:

- A document camera, for keying in on documents, objects, participant movement, etc.
- A VCR, CD- or DVD-player or other playback device for sending pre-recorded video or multimedia information.
- TV receivers/decoders to support the sending of satellite, cable, and UHF/VHF programming.
- Specialty cameras, such as those integrated with microscopes, telescopes, internal surgical instruments, or other enhanced viewing devices. Interface to an electronic whiteboard for collaboration
- T.120 (an ITU standard for electronic data collaboration) peripherals such as digital drawing pads, scanners, or even additional computers for application sharing.

In varying the number of video sources, two or more inputs can be co-cabled to the same conferencing system and "swapped" in as the selected video source when changes in view are desired. An example of this is when the main camera is positioned to send a face-to-face view of the class to a remote site and an auxiliary camera is positioned in the back of the classroom to send a "head of the class" view of the instructor. Viewing input from more than one video source can be accomplished through the use of a video mixer, which takes input from two or more devices, "mixes" it together, and outputs the mix as a single video stream.

It is worth noting that having to choose one input device at a time for viewing is actually a limitation of the H.323 standard, as derived from the earlier H.320 standard. H.320 assumes bandwidth limitations that are not often considered within a TCP/IP environment (e.g., the Internet) and the H.323 standard carries this assumption forward. This limitation is not typical of videoconferencing approaches that "got their start" on the Internet. (See for instance [reference appropriate section here].) It is reasonable to think that both the H.323 standard and the products that incorporate it could evolve to support simultaneous viewing of multiple video sources from a single endpoint.

Other scenarios where multiple video sources are desirable include:

- remote surgery participation, where the internal view of the surgical camera is exchangeable with a view of the surgeon and the operating room;
- remote informational presentations, where the view of the presentation subject is complemented by the view of an expert who is speaking about it;
- a discussion surrounding a movie or TV broadcast, where a view of the program is alternated with a view of those discussing it;
- alternate views of the same room, e.g., where preset views of a main speaker, the close-up of a white board, and a view of all meeting participants is switched to match the focus of the meeting.

Output Devices, or the Video Display

Video display options are similar tovideo source options if you think of the term "display" in the broader context of "output". Video displays can vary according to the type and number of outputs (displays) that are used. Different types of displays include:

- A VCR, for recording incoming video.
- Room projection systems, to extend audio/video reception to accommodate a large space and/or audience.
- An encoder, to support preparation of satellite, cable, UHF/VHF, and streaming IP based broadcasts.

Also, when considering the number/type of displays, extra video displays are often necessary when extra cameras are introduced into a videoconferencing setup. Because it is natural for us to look at each other during conversation, it is natural for us to look towards the remote video display of the current speaker when we are in a videoconference. If the camera that is sending our video to the remote site is not positioned very nearly at or behind the primary video display, those at the remote site will view us as looking towards that display. To resolve this problem, extra displays that mirror the remote video should be positioned near each of the cameras that we will be "talking to".

One of the most common uses of additional video displays is that of "stretching" a smaller conferencing system (desktop or small-group) for use with larger groups. This is often seen in large classrooms, presentation facilities, or auditoriums where the remote display is mirrored to one or more large screens through a LCD panel or rear-projection system. Secondary displays can also be useful when a videoconference includes data collaboration or application sharing. Use of a dual monitor configuration on a desktop system, for example, provides the largest viewing of the remote site on the first monitor, while application and data manipulation take place without obstruction on the second monitor. Though surprisingly few H.323 terminals support this option today, it is a highly recommended configuration for videoconferences that heavily integrate data collaboration. Care should be taken, however, to avoid creating an environment in which multiple displays are required to participate in the conference, thus limiting the potential conferees to those with such capability.

User Interface Enhancements

Another area that is ripe for add-on and/or enhancement in a videoconference is the user interface to the videoconferencing endpoint itself. Most typically, you interact with a desktop conferencing system using a mouse and keyboard to interact with the videoconferencing software application the same way that you would with any other PC application. On group systems, this interface may be replaced with an alternate means of entering commands, such as a touchpad or remote control. These standard input devices can be supplemented through the use of some optional user interfaces. For instance, group systems that are placed in a classroom or presentation area equipped with additional specialized features (e.g., lighting and/or sound controls, projection equipment, computers configured for data access) may include a "teachers podium". You can think of this as a control panel (often fixed within an actual podium) that extends the controls for all the devices in the room into one aggregated and easily accessible location.

It is almost always advisable to involve professional help in planning and implementing a large group and/or multi-use system. What appears at first to be a simple change in the project scope or goals can make the installation of a videoconferencing system overwhelmingly complicated. Keep in mind each of the individuals and organizations that will be impacted by the system integration, at both the sending and receiving ends. Electricians, architects, Audio/Visual teams, and even carpenters will likely have a role to play in the project now. Professional videoconferencing design teams can help you understand how the pieces of the puzzle fit together and manage the project successfully.

Desktop systems are designed primarily for individual use and are more apt to support interface extensions that either improve personal productivity during a conference (e.g., data collaboration features) or improve integration with the user's normal routine (e.g., a telephone handset for privacy in a multi-person office). When considering data collaboration features, it is important to note that today's videoconferencing terminals vary in how they provide this functionality. Often, instead of being completely integrated into the videoconferencing application, data collaboration is accomplished by the videoconferencing application "calling" a helper application (i.e., NetMeeting) to assist. The features and smoothness of the data collaboration are then dependent on the helper application. This is a useful understanding to have when comparing/contrasting the user interface of different video terminals. In general, user interfaces vary widely across both group and desktop video terminals today and are a prime area for vendors to introduce product differentiation. The more time you spend actually using a system or walking through different usage scenarios before you buy, the less surprised you will be by how easy or difficultthe system is to use and how well it supports your purposes.

It is also wise to take into account the intended users' technical skill and comfort level when choosing a user interface for any given conferencing application. Videoconferencing systems are not a "one size fits all" technology. Many individuals can become quite intimidated when put in front of a PC interface. Conversely, when handed a remote control and a menu on a "TV" screen, they may become quite at ease. Always take the time to research the end user population and strive for a balance between functionality and user friendliness.

Network Connection

The network connection may not be an intuitively obvious area to "add-on" in a videoconference but considering the network connection in the broader sense of "the transmission services available to the videoconference" can be illuminating. This aspect of successful videoconferencing is covered in detail in [reference appropriate section here]. H.323 began as a standard for doing videoconferencing over a local area network (LAN) without Quality of Service (QoS) or other means of ensuring reliable performance. Given this, itmight seem that any LAN should support successful H.323 videoconferencing. However, the reality is that even H.323 videoconferencing will fare much better on network links that are not congested or are optimized in some way to provide a predictable level of service to the video application. True QoS standards that will be integrated into existing LAN/WAN protocol suites are currently under development and not ready for wide-scale deployment. However, some level of service guarantees can still be achieved through technologies that support "packet shaping" (giving particular data packets priority over others when sending them across the network) or simply through good network management (optimizing the network to support typical traffic patterns).

In addition in the near future, we might look to the network or the intermediate services that are evolving to be offered over the network (often called "middleware") to provide or support some level of security for a videoconference as its packets move from point to point. Services such as authentication (ensuring that packets are flowing between known and authorized endpoints) and encryption ("scrambling" video data between endpoints such that it cannot be viewed during transmission) may soon be services that are "added on" or made available to a videoconference as part of the connection. [Mary - Hopefully this sets us up to add a full section on middleware for digital video in a later update - once we know what that middleware will be ;-).]

Best Practices and Etiquette

Best Practices for the Video and Audio Environment

In this section we attempt to give a very brief (possibly oversimplified) look at how audio and video are captured and transmitted in a videoconference, what problems you might see during the videoconference, and how you can address those problems. Network issues can also affect the videoconference, but that discussion and problem treatment are addressed under Network Matters in this cookbook.

The Audio Environment

Audio is the most important part of a conversation. The audio system for videoconferencing consists of some combination of headset, handset, microphone, speakers, and digitizing device (hardware and software.) An ideal audio system is one that offers the widest frequency response (widest range of sounds or pitch) while using only a small amount of bandwidth and incurring minimal delay. For those who are interested, humans hearing is between 20HZ and 20KHZ, with intelligible speech being around 2KHZ. And studies show that 100ms delays are detectable but tolerable, 250ms delays are annoying, and 450+ms delays are unacceptable. [Network Week]

Click here for more detailed information on audio capture and transmission.

Questions to ask yourself about the audio quality:

Is the audio delivered at an appropriate volume with a minimum of background noise and hiss?

Your input device is likely to be a handset, a headset, or a microphone. First, if others are having difficulty hearing you, check your input device. The standard handset is known to deteriorate quickly. Try replacing it with your telephone handset. If the sound of your voice improves at the other end, you have a bad handset.

If you are using a headset, check the positioning of your microphone. Some headsets use microphone level output (meaning the sound of your voice generates the current required to carry the signal), therefore the volume will drop quickly as the distance between your headset microphone and mouth increases. For instance, you can double the output by decreasing this distance by 1/2.

The stock microphones are typically very basic units that can damage easily. Extension cables or damaged cables can add extraneous noise and hiss. A headset is often the best solution for basic equipment and good sound. If you plan to use videoconferencing often and desire semi-privacy during your conversations, invest in a good headset. Before doing so, verify that your videoconferencing client has microphone input or that you have access to a line-microphone input adapter.

Speakers and microphones play an important part. Does the system handle echo cancellation?

If you prefer to use speakers and a microphone instead of a headset or handset for your videoconferences, care must be taken in their selection. The standard speaker and microphone setups do not generally contain echo cancellation features. You can sometimes get by with the basic setup in a point-to-point call, but you will torture your colleagues in a multipoint call.

As your colleague's voice flows out of your speakers, your basic microphone will pick it up and feed it back through to their speakers or headset. Thus they will hear their voice echo back to them a fraction of a second later. The reverse case is also possible with the echo ing voice being yours back to you. This quickly becomes very distracting and annoying.

In a multipoint call, through an MCU, the echo begins to take the form of bells (or an even worse screech), with ever increasing volume and speed. The only way to survive such a call is to ask those on standard speaker and microphone setups to constantly mute their audio output when they are not speaking.

Companies like PolyCom make echo cancellation speaker/microphone combinations, called speakerphones, that work well in a variety of settings.

Most of this discussion applies to 1-3 people positioned at a desktop system. What if this is a large room?

Good room audio solutions are sometimes expensive solutions. Clients with standard echo cancellation features, used with basic speaker and microphone systems, work adequately but a "fish bowl" effect is sometimes noticeable. Frequency response and switching response become more important. A desktop caliber microphone may make the camera or MCU switch inappropriately as someone near the microphone shuffles paper or coughs whereas someone further from the microphone needs to shout in order to accomplish the same switching. Professional audio services should probably be consulted if the highest quality audio is expected for videoconferencing in a large room.

Does the videoconferencing client have automatic gain control to optimize volume on inputs and reduce background noise?

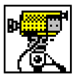
Most desktop videoconferencing clients require the end user to manually set the volume on the incoming call. In a point-to-point call, this isn't usually too cumbersome since you are dealing with one person at one volume level. In a multipoint videoconference, it would be desirable for the MCU to do automatic gain control or volume leveling across the callers. Such features do not exist in current MCUs and therefore each end user must potentially adjust their incoming volume according to multiple input (voice levels, equipment mix, etc.)

The Video Environment

Reading facial expressions and body language are the next most important parts of a conversation. As stated by Trowt-Bayard in "Videoconferencing, the Whole Picture", most of us are children of the television. We were born around or after the time that TV was "invented". Being such, our expectations on video quality are very high.

For those who remember, early television required much adjustment or fiddling with vertical and horizontal holds, adjusting the rabbit ears for better reception and sound, adjusting the contrast. Thanks to things like cable TV, digital video, and much higher bandwidths, there is no need to fuss with reception in this manner.

We've come so far. What problems could be left? Video techniques have been designed to accommodate those things to which our eyes are sensitive (like foreground and focus) and to devote less time and bandwidth to those things which our eyes might overlook (backgrounds, motion.)

Click here  for more detailed information on how video (both analog and digital) works. This may make it easier to understand how to achieve the best video quality possible.

Questions to ask yourself about video quality:

How is the video resolution? Do the colors flow smoothly? Is there any banding or dithering? Is there bleeding between colors? Do you see video artifacts such as blocks, splotches, and distortions? Other subjective measures include sharpness, contrast, brightness, color saturation, and stability (lack of snow or shimmering.)

First of all, test your focus. You can often test this through a local window, though sometimes a remote opinion helps. The location of focus buttons will vary so see your manufacturer's instructions for this detail.

Video formats are also defined for a particular pixel width and height (e.g. VGA is 640 by 480 pixels.) Has the encoding provided enough resolution for your purposes? Shrinking the picture size can help. (This is called scaling.) Enlisting an encoder format with higher resolution might be necessary, though associated bandwidth requirements will increase as well. Common Intermediate Format, or CIF, is a higher resolution format. QCIF, or Quarter CIF, produces compatible video at lower resolution (and bandwidth requirements.) Both CIF and QCIF can encode at 7.5, 10, 15, and 30 fps. If your client's vendor has chosen 7.5 fps, it will not handle motion as well. Some products also offer something called 16-CIF. Check your video format setting. If bandwidth is precious during your videoconference, consider dropping back to QCIF. If bandwidth is plentiful and resolution is important, try CIF (or a client that offers a CIF option.)

Could it be the network?

For an overall treatment of connectivity requirements, see "Network Requirements" in the Network Matters section of the cookbook. But remember, should you seem to be having an unusual amount of trouble during a videoconference (especially when previous videoconferences have gone well), check the paths to and from the other party. "Problems out on the net" do occur and you may be able to save yourself some unnecessary anguish if you postpone your meeting or accept issues over which you have no control. Tools like PingPlotter and traceroute can help you determine network outages or difficulties.

Does the client end station provide a consistently high frame rate (15 - 30 fps) during motion without sacrificing clarity?

With this question in mind, realize that frame rates will vary based on motion, dropped frames, network load, etc. "Jitter" and "stalled video" are symptoms of frame rate variation. Some client end stations have statistics on frame rates. If yours doesn't, check the bandwidth on your incoming and outgoing paths. They are likely to be different paths, which means that the video can look great in one direction and terrible in the other.

What happens when you move quickly or wave your hands? Linearity is a good measure of a client end station's sensitivity to motion and how consistently it maintains the frame rate. Video frames may be dropped. If this happens in bursts, the motion will appear jerky. If it happens in a predictable or uniform way, the motion will be smoother. Dropping the frame rate setting can sometimes help smooth the motion, though this option is typically not offered on most clients. Dropping the bandwidth manually may help and, in certain instances (like DSL), will significantly improve the video component of the conference. Better results are seen when using a garden hose on the roses than are seen when using a firehose.

Some client end stations perform video encoding on separate hardware. This encoding (typically H.261 or H.263) is a very computationally intensive process. Systems that do not provide extra hardware assist may experience some loss in clarity if encoding becomes a major burden on the workstation itself. Software codecs, especially on lower powered processors, may have more difficulty in supplying sustained and higher frame rates. Stopping all other applications may help.

The Joint Audio/Video Environment

Latency, as used here, is the delay between a video movement and the sound that goes with it -- the synchronization of sound with picture. As has been described above, the sound and picture in a videoconference are two distinct components that are produced simultaneously but captured and transmitted separately. It is up to the videoconferencing system to split them apart at one end, send them down the line, and put them back together at the other end. The videoconferencing system is also responsible for keeping them synchronized. Several things can impact this synchronization.

Questions to ask yourself about synchronization

Does the lip sync seem reasonable? Does a handclap synchronize?

It should be taken as a given that frames will be dropped. The codec should assume this and see to it that frames are dropped uniformly in order to maintain a sense of smooth motion. If yours doesn't and the latency is serious, consider a different client/codec.

Data sharing may have an affect. Data packets are given priority over video packets. If you have an active data collaboration going on, you may begin to see some latency in things like lip sync. Perhaps your participation in the application will distract you from the synchronization problems.

If you have other applications running, beyond those being used in the videoconference, they may be siphoning cycles away the codec, causing loss of frames and/or synchronization. This may be especially noticeable as you launch and stop applications. If you are experiencing serious synchronization problems, turn off other applications to see if that helps. A busy LAN and Microsoft Windows (the OS) buffering can also throw the synchronization off.

Videoconferencing Etiquette

Videoconferencing, by its nature, is a social activity. As with any social activity, there are acceptable as well as expected behaviors that accompany it. Some of these behaviors are the result of culture or the environment whereas some may be said to reflect "common sense". Of course, there is also a range of definition as to what is "acceptable", " unacceptable", "desirable" and/or "expected" based on individual interpretation and temperament. Finally, when compared to other well-established social activities that combine people with technology (e.g., talking on the telephone, watching a movie in a public theatre, driving a car), videoconferencing has not been around as long or had as much exposure. This combination of conditions results in the fact that videoconferencing "etiquette" is certainly not "carved in stone". However, there are some basic behaviors that will improve your own videoconferencing experience as well as that of the people you are conferencing with:

Testing, testing, 1, 2, 3…

Perhaps the most overlooked experience-enhancing behavior in a videoconference is simply to pay some attention to how others will be seeing and hearing you. In videoconferencing, much of the experience at one end is affected by conditions at the other. Most videoconferencing clients include a "self view" window. This lets you see how you appear to the remote end — whether or not you are completely viewable on camera, if there are distractions in the background, whether you are looking straight forward at the remote caller and not "gazing down from above" or "peering up from below". Even if the self-view window is not going to be kept up during the call, it's a good idea to preview your image in the window and adjust accordingly prior to the call. Unfortunately, this doesn't work for adjusting audio since your local audio is almost always suppressed from "feeding back" to you in local mode or even most test modes. In this case, testing and adjusting with a live call before a meeting begins or taking a few minutes to test and adjust at the start of a call is strongly recommended. Once a call is in progress, many people seem to tolerate poor audio or video conditions, not wanting to interrupt the conversational flow or simply because they figure it must be something "at their end". A short audio/video "rehearsal" is well worth the time spent as it contributes to making the technology as transparent as possible and enables comfortable, effective and rich communication.

Leaving well enough alone…

Once adjustments have been made at each end to produce optimal call conditions, perhaps the most important advice is to converse naturally and make as few additional adjustments as possible. True, some adjustments may be necessary in response to environmental changes (lights are turned on/off, background noise increases). However, unnecessary "twiddling" of audio or video can have very distracting results. For example, leaning forward and adjusting a desktop camera at your local end will produce the dreaded "giant palm monster" effect at the remote end, who see friendly faces of acceptable proportions replaced by a far-too-intimate view of all or parts of a hand. Also, if limited range or uni-directional microphones are being used, excessive movement or position shifting at the local end will produce audio break-up, swells and fades at the receiving end.

Are you still with me?

Once your camera and incoming view window have been correctly positioned so that "eye contact" has been established between you and the remote site, you should remain focused in that direction. Shifts in attention such as looking out a window, looking at other applications on the computer screen, "multi-tasking" with other work in your office, have the same effect as not looking someone in the eye when talking to them in person. It's important to realize that videoconferencing is much more like an in person exchange than a telephone call — body language and facial expression count!

Talking out of turn…

As with any in-person meeting, stray noises and side conversations within a videoconference distract from the primary conversation. This can complicate point-to-point meetings and becomes even more noticeable in multi-point meetings. It seems to be easier for participants to forget that they are truly part of a group conversation since the meeting room is virtual rather than physical. Side conversations at remote sites seem to spring up more readily than they would if everyone were in the same actual room. The microphones and speakers necessary for sending/receiving audio complicate matters further in that they do not differentiate between relevant and irrelevant sounds. They will readily pick up any conversation that is taking place near them and send it along. They will also just as happily pick up and transmit a sneeze with as much sound quality as a well-intentioned remark. Given these "imperfections" with technology (and with people!), it is good practice to mute your own audio when you are not speaking. In a point-to-point conversation, this isn't as necessary and may actually result in unnatural pauses in the conversation as muting at either end is turned on or off. However, it is useful in situations where audio may be poor at either end and can be used to minimize the effects of the audio problem on the overall conversation. In a multi-point conference, muting your local audio by default and unmuting only when you want to speak is almost always a good idea. This is especially relevant in the case of a voice-activated MCU since capturing the conference audio will also result in capturing the conference video. Any "side action" at your site will then be displayed along with any "side noise". Think about it - you may not want everyone seeing your facial expression as you dissolve in a fit of coughing, or watching you tumble oh-so-gracefully over the chair that you just knocked down!

Wow! Where'd you get that shirt?

Once you minimize audio distractions, it's time to think about minimizing video distractions. How and how far to go about doing this is a topic of some debate. "Traditional" videoconferencing has paid significant attention to proper lighting, room aesthetics, and attire, particularly in "board room" or group settings. This is a sensible approach to a technology that relies on cameras and monitors to create the conversational environment. Such "production" aspects are similar to those that are considered when producing high quality television and video presentations. These are especially applicable in preparing a conference room or a classroom for group use (see "Developing a Productive Videoconferencing Room " in the Related Topics section.) However, if conferencing is going to take place on a regular (maybe daily) and unplanned basis from desktops located in individuals' offices or homes, the acceptable degree of "sensible preparation" becomes less clear. If communication via videoconferencing becomes as commonplace as using the telephone, what will our norms for video etiquette be? Will we have to stop wearing favorite clothes if they have complicated patterns or loud colors, in case we get a call that day? Will we have to re-engineer lighting in our homes and offices, or setup "video friendly" areas to take all of our calls? What if we're mobile — get out of the sun? Step into a "video phone booth"? The answers to these questions are likely to change as the "human protocol" for videoconferencing evolves and as the technology becomes more capable of simulating "reality". During this evolution, it's important to consider what does and doesn't work well at any given time and in any given situation to ensure that you are making informed choices.

We're all in this together!

A final subtle but very important point of video etiquette is that, when you are in a videoconference enabled meeting, though participants are located in physically different places, it is truly a "real" meeting! At first pass, this means things like "you should be on time", "you should pay attention", "you should make sure everyone has the same information going into the meeting", "you should bring enough materials for everyone". In the case of a multipoint meeting, these considerations are more complicated in delivery but compounded in importance. For example, if hard copy materials will be used in the meeting, they should be sent to all locations ahead of time (not unlike preparing for a teleconference). If printouts will be made from electronic material presented during the meeting, you should be sure that all sites have the capability to print the materials. If particular local objects or room locations will be shown during a meeting, care should be taken ahead of time to ensure that camera views of these are available for remote participants.

Can I have some of that too?

A couple of other considerations are a little less obvious but really do make a difference, particularly in multi-point meetings when groups of people have been brought together at each of the participating sites. The first is that information which is specific to each local site (e.g., where the restrooms are, where to find a phone) may need to be distributed to those that are at the local site but isn't relevant to remote sites. Distribution of this information should be handled locally via pre-meeting communication, local handouts, or prior to the start of the meeting with local audio muted. In addition to this, if amenities differ from site to site, care should be taken to minimize group exposure to the differences in amenities. (In other words, if bagels and coffee are available at one site but not at another, it would be most polite to eat off camera!) Better yet, care should be taken to ensure that amenities are equal. Remember, it really is one meeting!

Behind the Scenes on Audio

We often hear about video encoding, but the audio must also be digitized. To do so audio is sampled every n seconds, where n is twice the highest frequency of the analog signal. The sampling rate turns out to be 8,000 samples per second. The sample is then quantized (in the range of 0 to 4095) and then compressed (to a value between 0 and 255, which will conveniently fit into one byte.) Compression is done through things like pause elimination and logarithmic schemes that take advantage of the fact that our ears are more sensitive to small, low volume changes. The signal is then transmitted at 64 Kbps (8,000 samples per second x 8 bits), though we will see below that in some cases the signal can actually be sent under 16 Kbps with good results.

See Nyquist's Theorem or discussions on "lossless digital representations" for more information on sample size determination.

G.711 is an audio standard that uses something called Pulse Code Modulation (PCM) encoding. PCM transmits the sampled audio as pulses that are coded to represent the amplitude of the original signal. PCM generally requires 64 Kbps bandwidth in videoconferencing clients. The H.323 standard requires support for G.711. G.723 is another standard, using Multipulse-Maximum Likelihood Quantization (MP-MLQ) for high quality at 6.3 Kbps and Algebraic-Code-Excited Linear-Prediction (ACELP) for good quality at 5.3 Kbps. Each of these has the added advantage of limited complexity. G.723 is an optional encoding format within the H.323 standard. Since a client can make no assumptions about optional features within clients from other vendors, you will generally only see G.723 used between clients from the same vendor.

The ITU G.728 standard is a 16 Kbps algorithm using low-delay code excited linear prediction (CELP). It is based on a standard analysis-by-synthesis CELP coding technique. However, several modifications are incorporated to meet the needs of low-delay high-quality speech coding. The G.728 standard was designed to provide speech quality equivalent to or better than that of G.721 (32 Kbps Adaptive Differential Pulse Code Modulation.) The G.728 coder was also designed to behave well in the presence of multiple speakers and background noise, and to be capable of handling non-speech signals. (Texas Instruments)

Behind the Scenes on Video

Cameras work much like our eyes. The camera utilizes lenses and solid-state chips to store the electrical charges representing the picture much as the parts of the eye do. But, whereas the eye transmits the picture to the brain in parallel, video networks must split the picture into scan lines and transmit the scan lines serially. (The name pixel comes from this process.) A complete picture is called a frame. Providing that the frames can be transmitted fast enough (the frame rate), the sense of smooth motion will be preserved. Today 30 frames per second (fps) is standard in a videoconference though 15 fps is often adequate.

Telephone companies have transmitted audio digitally for years. But while most video is captured in analog form, network transmission of that video is best done digitally. Analog systems are more susceptible to noise (interference, amplification across distances.) Screens are susceptible to artifacts (unwanted elements) and anti-aliasing (distortions.) Digital systems tend to be immune to these things. Repeaters (which take the place of amplifiers) are more accurate. Computers and networks are based in a digital world meaning streams of bits can more easily be pieced together when in digital form. So moving sound and video to digital form allows us to use the high processing power of computers for encoding, compression and storage and the economies of networks for transmission.

Video Standards

The common standards for digital video include the ITU-T's H.261 and H.263 Recommendations as well as the Moving Picture Experts Group (MPEG) standard. H.261 and H.263 are typically seen in today's H.32x videoconferencing clients whereas MPEG is more often seen in streaming or broadcast video.

The device that performs the encoding/decoding and compression is called the codec. Some systems provide hardware assist units for the codec duties. These units contain processors that have been specially developed for video applications, handling such things as encoding, decoding, compression, scaling, pan and zoom, local and far-end camera control. Other systems rely on software codecs. The latter will siphon cycles from the more expensive PC processor, possibly affecting other applications or, should the processor be low powered, the videoconference itself.

H.261 is the most widely used videoconferencing codec in the world. H.261

works with the two picture formats that support the three television formats known around the world (NTSC, PAL and SECAM). The higher resolution Common Intermediate Format (CIF) compliance in videoconferencing clients is optional under H.261, but Quarter Common Intermediate Format (QCIF) is mandatory. QCIF is lower resolution thus more suitable to small screens and talking heads. H.263 is another common, optional video standard. H.263 provides better quality at a lower data rate, but at the price of a fixed frame size of QCIF or 176 by 144 pixels. Simply stated, in H.323 videoconferencing, H.261 is mandatory, H.263 is optional, MPEG is not included in the standard at all at this time.

Video Encoding and Compression

Whereas H.323 transmits video digitally, video is typically transmitted via an analog signal. While a digital signal is comprised of a stream of discrete 0Õs and 1Õs, an analog signal is a continuous wave of varying frequencies and amplitudes. Frequencies describe how rapidly colors change (for example along edges) and amplitudes tell the magnitude of that change. The image must be encoded into digital form.

The cones in our eyes perceive color as changes in brightness (luminance), hue (color), and saturation (depth of color.) Our eyes pick up changes in brightness much more quickly than the other two. Therefore when transmitting digital video, more bandwidth is devoted to luminance. H.261 and H.263 use the Discrete Cosine Transform (DCT) to separate the image into these different parts. The coefficients describing these parts are then placed into a matrix according to a predefined order. For example, through color differencing, the chrominance (or color) information is subtracted from the luminance of the analog signal. Red, blue, and green difference signals result (the R, G, and B waveforms.) Eight bits are needed to represent each of the R, G, and B signals and each color component requires a frame of information. The digital image is then formed as a matrix or grid of these values representing intensity or color. The inverse DCT is used at the receiving end to reform the original video signal.

Uncompressed video typically requires bandwidths of 25-90 Mbps. Compression of the digital signal must therefore be used to further speed up transmission by lowering these bandwidth requirements by a significant amount. *Lossless* compression typically has a compression ratio of only 2:1, but it is very useful for transmission of documents when accuracy is critical. But this ratio is nowhere near adequate for video transmission. *Lossy* techniques (meaning that part of the signal is discarded) produce good compression ratios (100:1 to 200:1) though sometimes at a high cost in video and audio quality. Compression can also be *symmetric* or *asymmetric*. Asymmetric compression relies on more compute intensive compression techniques and is therefore more useful for video servers and broadcast applications. Symmetric compression is used for videoconferencing.

Compression also takes different forms: *intra-frame* (within frame) and *inter-frame* (between frames.) Intra-frame and inter-frame encoding/compression are useful for material that doesn't change often (redundant data) in the scene or between scenes. On average, only 9% of the pixels in a moving sequence change between frames.

Each form of compression has techniques for eliminating redundancy, thereby reducing the amount of information that must be transferred. Intra-frame compression works on the premise that redundancy exists in parts of a frame that are close to each other. Inter-frame compression works under the assumption that there is significant redundancy across frames over time (for example, walls and pictures don't change.) Some compression algorithms even have techniques to predict the next frame or describe frames in a fractal (equation) form, further reducing bandwidth needs. The waving hand is often used to test video quality per the encoding and compression functions of the client's codec.

See discussions on difference pulse code modulation, conditional frame replenishment, motion compensation, and motion vector prediction for information on interesting inter/intra compression techniques. These items are beyond the scope of this cookbook.

Several other features are often found in H.261 codecs. They include Forward Error Encoding (FEC), Video Multiplexing (timing, buffer sizes, etc.), and Layered Block Encoding (which appears to be a layered plan or protocol for the structure of the transmitted block to ensure the orderly transmission.) These items are beyond the scope of this cookbook.

H.261 requires intra-frame encoding via Discrete Cosine Transform (DCT). Motion compensation (or inter-frame encoding) is optional and, when included, is done through a combination of motion prediction, prediction error correction, and fresh image transmission (one out of every 32 frames.) H.261 codecs are required to only decode motion compensation and, when used with codecs that encode motion prediction, the quality is better and the transmission is more efficient.

Network Matters

Network Requirements

Videoconferencing was originally deployed over networks that could provide some guarantees about the level of service that would be delivered to the application. The ISDN and/or dedicated T1 circuits of the H.320 standards-based world provided predictable delays over dedicated paths. This allowed videoconferencing vendors to create products to work within these parameters. However, dedicated circuits are also expensive circuits.

H.323 standards-based videoconferencing was engineered for videoconferencing that takes place on a data network without any quality-of-service standard, such as the Internet. Such networks were not originally intended for delivery of sensitive near real-time applications. The data network is used for multiple purposes: e-mail, web browsing, and other activities are inter-mixed with H.323 videoconferencing.

The audio/video information within a videoconference is segmented into chunks by the application, encoded and compressed, put into a series of data packets and sent over the network to the remote end at basically constant intervals. The data packets may arrive at their destination at slightly varying times, if at all, and possibly out of order. To keep the "real time" impression of an interactive videoconference, the packets must arrive, on time and in time to be re-ordered for delivery through the videoconferencing terminal.

There are five fundamental network problems for videoconferencing over networks such as the Internet. They are *bandwidth, packet loss, latency, jitter and policies.*

Bandwidth is the fundamental requirement that there be enough space in a network path for all of your packets to get through unimpeded. For a rough idea of scale, a typical ISDN videoconference uses around 128-384kb/s. IP-based H.323 video systems can use the same bandwidth, although in general they tend to go higher since the network is cheaper, so bandwidth of around 384-768kb/s is very common. Higher-quality videoconferences can go to 1.5-2.0Mb/s, and if you want to go to broadcast quality the sky is the limit — 6Mb/s for NTSC/PAL transmission, 20Mb/s for prerecorded HDTV and higher for ‘live’ content.

This bandwidth need is symmetric — each end will transmit and receive this amount of traffic. If you are in a multi-point videoconference then you need to keep in mind that the MCU/bridge is seeing all of the streams at the same time, even if it is not forwarding them on. So if you have an 8-site videoconference running at 384kb/s, every site sends and receives 384kb/s to the MCU, and the MCU receives and forwards 8*384kb/s = 3Mb/s roughly.

Packet loss is when packets fail to arrive correctly. This can be due to insufficient bandwidth along the path (when congestion occurs, routers will drop packets), or perhaps errors in transmission. Errors occur most commonly on wireless links such as microwave, satellite or local wireless Ethernet. They can however also occur on copper and even fiber links. Packet loss results in effects such as "tiling" within the video window, missing pieces or blank areas within the video window, and/or disruptions in audio.

Latency is the time delay between an event occurring and the remote end seeing it. Latency is introduced both by the encoding/decoding process, and hence depends on the equipment used, and also by the time it takes packets to traverse the network. There is little you can usually do to change the network latency, on any large scale, beyond getting directly involved with a carrier or a research network. The speed-of-light is a limiting factor especially on satellite networks or international cable links.

Excessive latency increases the chances of people "talking over one another" because they don't realize that the person at the other end has started speaking too. Another problem is that the latency for the audio and video may be different, and hence lip movements don’t appear synchronized with the audio. This is a function of both the terminal and the network, and can vary dramatically — some products try to compensate for it. You should experiment to see if it is an issue for your applications.

Jitter is the random variation in latency, due to things like other processes running on the terminal (for example on your desktop PC), other traffic temporarily blocking the path through the routers along the way, or even the network path changing during a videoconference. In extreme cases this results in packets arriving out of order from their transmitted order. Jitter results in uneven and unpredictable quality within a videoconference. Terminals will try to compensate for this by buffering the traffic up to some finite time, before playing it out to you. This increases the latency even further.

Policies are introduced by things like firewalls and network address translation (NAT) devices that are generally used to try to hide or protect network elements from the wider Internet. H.323 uses dynamically allocated ports, usually, and is thus not very firewall-friendly. Unfortunately there are very few technical solutions to these, and require you to discuss policy issues with your, or their, network managers.

Given all of these potential problems, what can you about them? It is best to make no assumptions at all about your network's readiness for videoconferencing. You should discuss your videoconferencing plans with your network staff early on in your decision process, and ask them if your network will support videoconferencing in the locations you have in mind, including off-campus locations.

Nowadays most campus or corporate desktops have 10Mb/s or 100Mb/s Ethernet interfaces to the LAN, so a video-conference is not much of an impact there. However, your LAN needs to support "switched" Ethernet, rather than "shared", to be sure the videoconference does not impact others. A "shared" LAN divides the bandwidth between everybody, so on average you get much less, plus you increase the risk of packet loss and jitter.

Category-5 (or better) horizontal network wiring combined with fiber optic vertical wiring is recommended. Many university campuses have older wiring installed in some locations; if your wiring does not meet these specifications you should upgrade it.

There are additional features offered by Ethernet switches that may prove of value. One of them is support for multicast (see the multicast section elsewhere), and another is 802.1Q traffic prioritization. This allows you to give some devices a higher priority for network traffic than other devices — this is sometimes described as "layer 2 quality of service".

Across your campus, and out into the wide area is where most bottlenecks occur. Wide-Area links especially tend to be expensive and often narrower than you would like. It is possible for your network staff to provide a variety of traffic prioritization schemes at the IP layer, in the routers (but you need to be careful that other users don’t take advantage of it as well). This prioritization will only work as far as the edge of your network. Beyond that it depends on who you connect to. A commercial carrier will generally treat all traffic as equal, but they tend to over-provision bandwidth anyway. A research network (such as Abilene, GrangeNet, CA*Net, GEANT and similar) will often have a quality-of-service research project, which can handle priority traffic across the wide area. In all cases your network staff should be able to guide you.

The exact path along the network between video terminals, or from terminals to the MCU, will also affect the performance of your conference. Network packets do not necessarily take the shortest path from one location to another; routers determine which path is taken. A router must examine the destination address of the packet and then calculate where to send it. Every pass through a router is called a "hop". Because a calculation is involved, every "hop" adds a bit of delay to the total time required to transit the entire path, increasing latency and jitter, and also provides bottlenecks (routers have only so much memory and CPU capacity), increasing the risk of packet-loss. The fewer hops you have along the path the better off you usually will be.

You can learn the network path involved using a tool called "traceroute" (on Unixen, "tracert" on Windows). Traceroute will reveal all the IP hops involved and also provides some information about the amount of delay, in milliseconds, at each hop. The traceroute utility only checks the path FROM your computer TO the destination you specify, but not the reverse direction. Routing is not symmetrical: the path from A to B will not necessarily be identical to the path from B to A. Therefore, it is important to initiate a traceroute from each location. As a result, a videoconference may run wonderfully in one direction and poorly in the opposite direction (in fact, those symptoms indicate asymmetric routing). In most cases you want the path to be symmetric, but in some rare situations you may have no alternative (e.g. using one-way satellite transmission with a terrestrial return). Also be aware that traceroute does not identify the Ethernet switches along the way, which may be where a bandwidth problem occurs.

In general, it is strongly advisable to talk to the network staff at both ends of a planned videoconference. They can advise you of potential issues, monitor the network for problems, and possibly engineer a better setup for you. Just keep in mind they are usually harassed, over-worked and often quite conservative. There are many good mailing lists and websites that provide contacts with other network engineers who may be able to help. Some of these lists and sites are provided in the appendices.

Network Address Translation (NAT) and Firewalls

Videoconferencing is often recognized as a difficult service to negotiate with Network Address Translation (NAT) and Firewalls. The H.323 standard defines a service that utilizes bi-directional communication for call setup with connections characterized by dynamic port utilization and a high data rate. There are both UDP and TCP parallel connections made on the dynamically negotiated ports. The situation is further complicated by the transmittal of this port information in the payload section of the IP data stream that makes up this negotiation. (RFC 2663) The difficulty in the deployment of videoconferencing applications is due to the complexity of the H.323 protocol and its dependence on some of the utilized components (i.e. gatekeepers, MCUs, etc.) and the variety of vendor implementations.

NAT and Firewall technologies were deployed to solve important security issues by limiting access to an internal network's ports by filtering inbound Internet traffic. In addition NAT provides IP address space to the internal network by brokering a single port to multiple internal ports. Because of the widespread use of NAT and Firewalls, and given the characteristics of videoconferencing systems, these particular network components and their configurations have become an obstacle to deploying H.323 application systems. NAT is actually a form of firewall but is so widely deployed in home DSL/Cable routers and 802.11 Wireless Access Points that it deserves individual consideration.

How the VC requirements are incompatible with NAT

All TCP/IP applications depend on network routers to control the direction of IP packets to insure that each packet reaches its intended destination. Routers that function in their purest form do not pose a problem to H.323 traffic in particular. What frequently causes problems with H.323 traffic is a mechanism to conserve IP addressing space called Network Address Translation, or NAT. NAT accomplishes this by transparently sharing a single IP address with multiple hosts. By translating the IP address used on a private network to an IP address suitable for the public Internet the NAT-enabled router can map private IP addresses and private ports to external IP addresses and external ports and therefore support multiple private IP addresses.

The important aspect to remember of the relationship between the H.323 application and the network router is that the application is unaware that traffic is running through a NAT facility. Another problem exists because NATs typically set up its port mappings by examining the applications packet header. While this works well for most applications, the H.323 standard requires IP address and port information to be stored in the data portion of the IP packet.

Solutions to VC / NAT problem

Most NAT implementations will allow for the configuration of a DMZ host or in other words to pass all inbound traffic to a particular internal host. This is not always without some packet modification that would preclude H.323 communication and certainly introduces significant security vulnerabilities into an internal system. A more acceptable alternative is an application layer gateway or proxy that is able to interpret the call setup traffic and create the required port configuration. Most clients support proxies but do require address configuration.

PhonePatch is an H.323 proxy server for Windows and Linux.

Open H.323 Proxy is an open-source solution utilizing the Open H.323 libraries.

The GnomeMeeting community is promoting an alternative to NAT that allows H.323 signaling packets to pass unmolested (RFC 3103).

How the VC requirements are incompatible with firewalls

As stated earlier, NAT is a general type of firewall technology. The other main types are packet filtering, circuit-gateway, and application proxy. Any discussion of firewalls and their various implementations and deployment strategies is essentially eternal so this material should be considered superficial to the extreme. Consideration must also be given that firewall configurations commonly require asynchronous profiles for internal and external clients and frequently videoconferencing will transverse multiple firewalls. The following is very brief description of these firewall technologies and their implications for H.323 Videoconferencing.

Packet filtering firewalls block or allow connections based entirely on addressing information in the IP header. As we discussed earlier, H.323 embeds routing information in the actual data or payload of the packet that responds to a signaling request, so packet filtering provides no method of associating the UDP request with the routing information.

Consequently, the only way for a packet filtering router to support H.323 is to open up all UDP and TCP ports above 1024 in each direction. Obviously both source and destination firewalls would have to be configured similarly and therefore significantly reducing the protection the firewall was implemented to provide.

A circuit gateway firewall allows a UDP request to initiate the opening of dynamic ports for a limited time to allow for streams associated with the application. Most firewalls utilizing this method have some understanding of common application protocols such as Telnet and FTP. Their behavior is predictable because of static port requirements and the result is an expected state. The circuit gateway firewall can provide adequate support for videoconferencing if it can disassemble the payload and respond by opening the requested ports dynamically. The disassembling of the packets and associating the contents with the UDP signaling request is complex, requires more than standard application savvy, and can introduce latency.

An application proxy firewall is evident to the application because it implements a partial H.323 stack. The proxy performs address translation but can mange the internal and external ports by actually participating in the signaling process and thereby opening the requested data ports. The proxy firewall behaves transparently to both internal and external clients after the connection is made but making that connection is not without cost. Since the clients are aware of the proxy they must be configured to both place and receive calls through the proxy IP address. The clients must be configured with this address (like any proxy) and the H.323 identity fields must be completed to enable access.

Solutions to Videoconferencing problems with Firewalls

Of course one solution to supporting videoconferencing across the firewall would be merely open the required ports to all network traffic. While this may seem impetuous it is a solution recommended by at least one product vendor. Fortunately there are better alternatives available. Many vendors have now implemented firewall technologies into router products that acknowledge the limitations of older implementations and properly support H.323 and SIP network traffic. These solutions can detect the videoconferencing signaling requests and take appropriate action to allow the traffic to transverse the router or firewall.

Aravox is a RadVision partner with firewall products that support videoconferencing protocols with very low latency.

Cisco supports videoconferencing with its own PIX firewall products and through integration with other vendors' products.

Check Point Software's Firewall products are H.323-enabled.

Port Forwarding

What's all this about ports?

One piece of information contained in the packet header is something called the port number. Source and destination port numbers, combined with source and destination addresses, uniquely identify a connection. Port numbers often refer to applications or services. Some port numbers commonly seen, particularly with respect to digital video, include:

Application/Service	**Port Number (Range)**
ftp	20-21
Telnet	23
HTTP	80. 8080
Kerberos	88

Pop Mail	110
Lightweight Directory Access Protocol (LDAP)	389
T.120 (Data sharing)	1503
Gatekeeper Discovery	1718
Gatekeeper RAS	1719
H.323 Call Setup	1720
Audio Call Control	1731
H.263 Video Streaming	2979
H.245 (Call Parameters)	1024-65535
RTP (Video Data Streams)	1024-65535
RTP (Audio Data Streams)	1024-65535
RTCP (Control Information)	1024-65535

(A long list of protocol and port numbers can be found at the Internet Assigned Number Authority (IANA) Protocol Numbers and Assignment Services, IANA port number, and the RFC 1700 web pages.)

This connection identity method must, obviously, change somewhat where NATs are used. Much as how the gateway or router remaps unregistered, private ip addresses to registered, public IP addresses, it also remaps the port numbers. This is done using a port -mapping table to remember how the ports were renumbered for outgoing packets. The router can also reverse this process so that returning packets reach the correct incoming destination.

Modern firewalls and routers can be configured to pass contacts through to certain sites and certain ports. For example, if we want to provide a service to the outside world, say an anonymous ftp service, we can instruct our firewall to allow packets through on port 20/21 if they contain our server's destination address. In the case of videoconferencing through a private network, we simply tell the router which port numbers to forward (basically, which application's packets to allow through.) PolyCom's ViaVideo is an example of a videoconferencing application that handles port forwarding as well as NAT.

DSL and Cable Modem

Today, due to improved pricing, many people are connecting to the Internet from home via DSL and Cable Modem solutions. Both solutions are much better than dialup modem and are even usable for some video applications. It is important to understand a few things about these solutions before you get started.

How fast are they?

The uplink speed and the downlink speed are not generally equal. These solutions have been developed to deliver content to you fast. Downlink speeds are higher, designed to bring [web] content to you quickly. Uplink speeds are designed more for typing and moderate file transfers.

DSL technology converts your phone company's copper wire into use for high-speed network access (even while you're talking on the telephone.) DSL connections are dedicated from your home to your provider's central office and speeds are a function of the distance to your provider's central office, the gauge of the phone lines, and the DSL technology. Some recent information on DSL speeds show:

DSL Technology	Uplink	Downlink
SDSL	90-680 Kbps - 1.5 Mbps	640 Kbps - 1.6 Mbps
Residential ADSL	90-680 Kbps	1.5 Mbps
ADSL	1.1 Mbps	7.1 Mbps

(This author sees 211 Kbps on uplink and 1.2 Mbps on downlink.)

Cable modem provides shared broadband services in private network segments that ultimately meet on the commodity Internet. (This author saw peak performance similar to her DSL performance, though the cable uplink speed may have been somewhat faster.) As a shared service, performance will vary particularly during peak times such as afternoons or evenings or as your neighborhood cable modem subscriptions increase.

Finally, while both technologies provide high-speed access, performance to the destination site may vary due to routing variations and possibly commodity Internet congestion.

What does this mean for video?

Most of us have grown accustomed to a minimum bandwidth of 384 Kbps and 30 fps for H.323 videoconferences. While the downlink for cable or DSL can handle this, the uplink cannot. Additionally, many end station clients are constantly sampling and readjusting bandwidth and frame rates based on the congestion feedback they collect. This can add up to very poor quality including serious audio and video frame loss.

The solution here is to choose or set a specific call speed. This author has generally found that 198 Kbps or 256 Kbps works well and gives sufficient video quality for a standard meeting. Small pictures work best. A framerate of 8-15 fps can even be supported.

Anything else I should know?

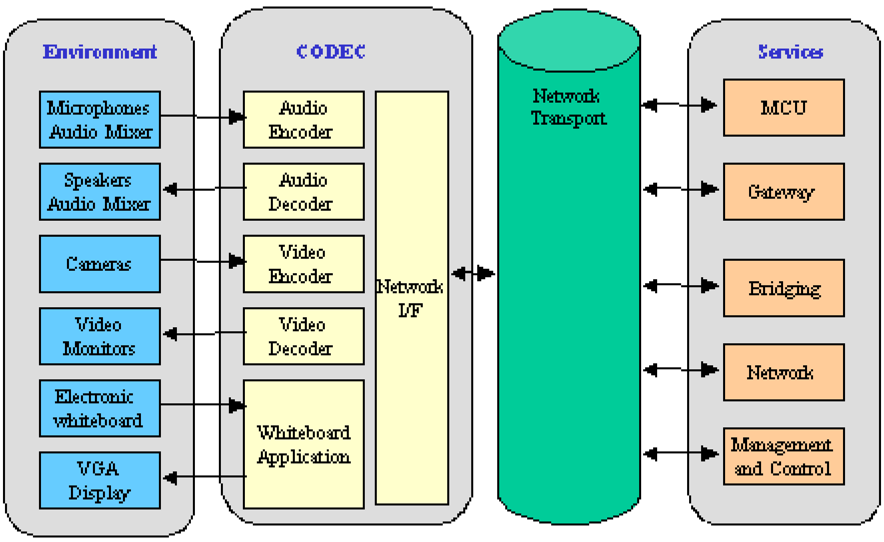
Many people are now setting up home networks for the family with economical hubs. These are excellent solutions for web browsing and email and, as mentioned above, can even provide some security behind the associated firewall-like services. But most videoconferencing end stations or MCUs expect to find each other with a specific, unique address. A residential hub sets up a private subnet behind the provider's ip connection. This means that, if you are on a private subnet, others will not be able to dial you directly.

Several solutions have been implemented on both ends. Hub suppliers are enabling port forwarding for registration and special media streams (see above.) Some hubs also come with a DMZ setting so that all firewall services can be disabled in order to allow the media streams to reach the appropriate computer. And some H.323 end station clients have built in Network Address Translation (NAT) in order to operate properly on the hubs.

Cable modem and DSL have enabled excellent telecommute environments. Providers have begun to see the value of and demand for carrying media streams. We can anticipate continued improvements and bandwidth in years to come.

Advanced Videoconferencing Components and Management

Managing Videoconferencing Services



As we have discussed throughout this book, the H.323 standard defines a videoconferencing endpoint, or "terminal", for making simple point-to-point video calls. The model above presents one view of the context in which a call occurs while also illustrating options for additional related communications.

The left two boxes of the model really combine to represent a typical endpoint. The codec itself is made up of several sub-components that handle the audio, video, and data content of the call and also includes an network interface of some type to transport the compressed data away from the sender and off to the receiving endpoint. The "environment" elements included within the first box are those most immediate to the codec, specific input and output devices that serve as interfaces for the end-user. Since the endpoint doesn't operate in a void - its primary purpose being to enable interactive two-way communication - some form of network transport is the next necessary element, providing a means to connect to other endpoints. This is true whether endpoints are relatively close to each other (between buildings on a campus, for example) or very far away (for instance, between two researchers on opposite sides of the globe.)

Up to this point in the Cookbook, little has been said about what other services might be available to an endpoint in addition to simple point-to-point dialing. The Services box in the far right of the model introduces the most typical of these. The H.323 standard itself defines three additional videoconferencing components to extend or improve access to videoconferencing functionality. These components are gatekeepers (or "call servers" if considered more generically), as one means of management and control, and also gateways and MCUs (multipoint conferencing unit). In addition, middleware and advanced network services are on the horizon and evolving to provide H.323, along with other network-based applications, a more secure and yet easier to use networked environment.

Some explanation of the need for advanced network services and a preview of a few services that are emerging "near term" are offered as part of Emerging Collaborative Technologies. Middleware for videoconferencing as of this writing is yet to be defined to the degree that it will impact videoconferencing on a large scale for what is likely to be the next couple of years. Both advanced network service and middleware however, will, perhaps drastically, affect how videoconferencing over IP can and should be implemented and how it might be used. In future editions of the Cookbook, we will address these in greater detail but wise videoconferencing practitioners (end-users as well as support staff) will begin tracking developments in these areas as soon as possible. Links to significant development efforts and educational tools for "keeping up" in each area are offered in the appendices.

Within the Services area, however, there are some components that provide extended functionality today, particularly related to H.323. These are the gatekeeper, the gateway and the MCU. As with other multi-user network-based components, understanding and proper management of these devices is critical for their successful use. Each component is capable of offering services that are readily deployable (although their functionality may still be evolving) and that, when matched with end-user needs, quickly become indispensable parts of an H.323 service offering. In the sections below, we take a closer look at each of these components - what it is intended to do, key aspects of its current functionality, and a preview of its potential role and functionality in the future.Following that, we provide some pointers on determining your users' needs followed by an examination of some key issues involved in deploying H.323 on a campus today.

Gatekeepers

An H.323 gatekeeper is assigned control of a particular set of videoconferencing resources (terminals, gateways, MCUs) and functions somewhat like a videoconferencing "traffic cop". In this role, the gatekeeper can provide or facilitate several services that enable H.323 conferencing to be more reliable and more secure. If a gatekeeper is present on the network, the H.323 standard requires that H.323 compliant terminals register themselves with the gatekeeper and allow the gatekeeper to identify them to others and control their activities within the zone. Also, if a gatekeeper is not present, the standard allows for the terminal to control its own calls, placing them via IP address with no gatekeeper registration or intervention is required. In practice, however, gatekeeper registration behavior is somewhat unclear (how does the terminal know for sure that a gatekeeper is present?

What if there is more than one gatekeeper readily available?) and difficult to enforce (what if a terminal registers with a "rogue" gatekeeper that has been installed on the network? What resources will the rogue gatekeeper be able to provide access to?). Once a terminal is registered with a gatekeeper, the H.323 standard identifies some broadly defined key services that the gatekeeper could provide:

- Address translation - This function maps an alias or 'video telephone number' of a user to the physical IP address of a terminal. This allows for people to call each other using user-friendly identification, such as an short numeric extension or an email address. Notably, a common schema for scaleable global addressing has not been defined.
- Admissions control - This function accepts or declines a call based on a variety of criteria, including available network bandwidth or specific user authorization level. Simple gatekeepers allow all calls through. (This level of call control is distinct from control at the terminal, where the end user can decide whether or not to answer any given call).
- Bandwidth control and management - The gatekeeper can accept or deny calls based on the total available network bandwidth or based on a preset maximum number of simultaneous calls. This keeps videoconferencing calls from overloading the network. The gatekeeper may also handle requests from terminals for additional bandwidth during a call. In many ways, the bandwidth control and management functionality overlaps with the "bandwidth broker" and "policy broker" functionality under investigation as part of IP QoS (Quality of Service) development. The gatekeeper
- Zone management - Each gatekeeper sets up a zone that may include terminals, gateways, and/or MCUs. The gatekeeper controls identification of and communication between devices in its local zone, allows devices to join or leave the zone, and controls access to the local devices from H.323 devices outside the zone.
- Call control signaling - The gatekeeper can process call control signals for particular calls, or allow this information to bypass it and go directly to participating terminals. If the gatekeeper remains instrumental in call control, enhanced management and error handling are possible but with the tradeoff of additional network and processing overhead.
- Call authorization - The gatekeeper can reject calls sent to terminals in its zone. The gatekeeper can also control what call types and resources are authorized for specific terminals. However, authentication is currently based on IP address and/or alias and not tied to any user-specific authentication
- Call management and tracking - The gatekeeper can track current calls, log calls placed over time, and provide this call tracking information to other devices. Such information can be used for system administration and maintenance as well as for billing purposes.
- PBX functions - The gatekeeper can provide "PBX-like" services such as call identification, call forwarding, and call transfer. These features, in turn, can make possible applications such as a 'video receptionist' and 'video voicemail.'

Gatekeepers today are available as full-featured standalone software applications and also as scaled down "built-in" functionality included within H.323 terminals, gateways, and MCUs. The degree of video resource identification and control provided by current gatekeepers varies widely and interoperability between one vendor's gatekeeper and another vendor's gatekeeper-controlled resource can be very uneven. Additionally, inter-zone communication and resource sharing between gatekeepers is far less than what would be needed for seamless conferencing on a global IP network such as the Internet or Internet 2. The issues surrounding such implementation can be numerous and it is safe to say that discussions about standards development as well as implementation of H.323 gatekeepers often produce more questions than they answer. However, it is widely agreed that the gatekeeper is a key concept and component for enabling scaleable, Internet-based videoconferencing. Most organizations are approaching gatekeeper deployment with the mindset that gatekeepers must be deployed, even "as is", while the developers and the community work to make them what they can and should be.

MCUs

The ability for two people at separate and remote locations to shrink the impact of the geographical boundaries between them via videoconferencing is certainly exciting and valuable. The concept becomes even more powerful when several locations can be brought together into the same conference, creating a "virtual meeting room" that exists for that particular time and group configuration facilitated by the network. Such "meeting rooms" are created through the use of a Multipoint Conferencing Unit (MCU). The purpose of an MCU is to connect three or more videoconferencing systems in the same conference, managing audio and video from each participant to the others such that group communication is achieved. Data sharing is also possible between all participants in a multipoint conference though current implementations vary greatly in terms of how this is done and also how well it works.

The H.323 standard outlines two component processes that form the basis of any multipoint interaction — the MC (multipoint controller) and the MP (multipoint processor). The MP is optional and, if present, there may also be more than one. with two different ways to provide multipoint functionality overall: centralized versus decentralized.

The MC provides for overall control of the conference. This involves forming connections between all endpoints, negotiating common capabilities, and communicating to the MP regarding any necessary switching of audio/video sources. The MP handles the actual processing of incoming and outgoing audio/video streams. Audio from all sites in a multipoint conference is typically mixed and delivered back to all sites in full duplex mode. Video, on the other hand, may be handled in a few different ways:

1. Switched based on voice activation (everyone sees the current speaker)
2. Switched via manual control ("chair control", where the designated chair decides whose video is being seen)
3. Displayed together on a split screen display ("continuous presence", also sometimes called "Hollywood Squares")
4. Displayed in individual video windows, one for each site that is being received.

In a centralized MCU, the MC and MP are included in a single unit to which all endpoints connect. This forms a physical and logical star configuration with the MCU at the center. Each endpoint is, in effect, in a point-to-point call with the MCU.

In a decentralized MCU, there is no device that can readily be pointed to as "the MCU". Instead, the component processes (MC and MP) are present to some degree in the client endpoints. The MC of one endpoint will most likely be used to control the conference while each endpoint uses its own MP to send/receive streams in accordance with its own capabilities. The video/audio/data streams from each endpoint are sent one-to-many, which requires the use of IP multicast to facilitate group identification and participation.

Arguments for and against centralized versus decentralized multipoint conferencing are not unlike those surrounding the debate of centralized server-based computing versus peer-to-peer computing. However, with particular respect to H.323 multipoint, the centralized approach has a practical lead at this time given the current state of the H.323 standard. Centralized MCUs are more thoroughly defined and more readily understood; therefore they are more widely available in standardized product implementations. Still, a quick review of the pros and cons of each approach can be helpful.

Centralized functionality lends itself to improved reliability, control and management. It also allows for advanced capabilities to be introduced into one entity but made available to all, thereby reducing costs at the endpoints. Of course, cost is then shifted to the central unit - in this case, the MCU. Other functionality, such as additional transcoding or network gateways, can also be fairly readily added to a centralized MCU, extending the service capabilities further than "simple" multipoint call handling. Again, this increases the cost and complexity of the MCU while decreasing cost and complexity required for client endpoints. Another consideration is that, until quite recently, most centralized MCUs forced each conference participant to the lowest common denominator for call capabilities. For instance, if one participating endpoint could only send/receive QCIF calls at 128K bandwidth, all other participants in the same conference would be forced to send/receive the same. This limitation is changing as increased transcoding capabilities are being introduced into some centralized MCUs.

Decentralized functionality more readily supports flexibility for end-users and a more distributed load over the network. Cost can be determined and distributed based on capabilities desired for particular endpoints. Each endpoint also determines its own send/receive capabilities and does not need to adjust these based on what other participants can do. Also, in addition to providing a mechanism for group calling, support for IP multicast allows for the most efficient use of bandwidth as determined by the placement and concentration of participating endpoints within the network.

Another consideration for the implementation of an H.323 MCU is hardware versus software-based. Again, the factors influencing the decision are not exclusive to a discussion of H.323. Hardware implementations tend to be more expensive and are likely to contain a variety of proprietary components but are likely to be faster and are also prone to be more reliable. Software implementations are more portable, more flexible, and less expensive but may suffer performance issues due to their reliance on the operating system and resources of the computer they are running on. Each type of implementation is available on the market today in a variety of forms. A careful matching of performance requirements to cost variables should be combined with a broad comparison of available products within each implementation type before a final buying decision is made.

There are a few different hardware-based MCU configurations that are available as of this writing. One type features a modular chassis that holds one or more power supplies and a number of other interface cards. Connection "ports" are included on some of these interface cards with the number of ports available corresponding to the number of sites that can be participating in conferences at the same time. Other hardware-based MCUs are based on more streamlined units that do not feature pluggable modules but instead are ordered with the desired number/type of ports built in. In either case, multipoint conferences involving specific numbers of endpoints (e.g., a three-point conference, a six-point conference, a 20-point conference, etc.) are "brought up" on the MCU and encumber as many actual ports as necessary for the number/type of connections and the amount of time required. Some MCUs include scheduling capabilities that allow conferences to be configured/scheduled in advance and brought up automatically. Others only allow ad hoc use of available ports on a "first come, first served" basis.

Software MCUs operate in much the same way as hardware-based MCUs but consist only of a software package running on a powerful server/computer. Software MCU manufacturers usually limit the number of simultaneous connections by a license key which is purchased by the customer. However, there are technical limits to the number of sites that can be connected together at one time based on the processing power and speed of the server.

Both hardware and software-based MCUs can be connected together to allow larger numbers of sites to be conferenced together simultaneously. This is termed "cascading" and is a functionality that is described in the H.323 standard. MCUs from different vendors should therefore be able to be cascaded together quite readily. In order to do this, one of the ports on each of the MCUs are used to "call into" the other.

Audio and video mixing/switching should still operate as if there is only one MCU involved; the cascading is transparent to the participants.

Gateways

A gateway provides transcoding services such as address translation, network protocol translation and audio/video coding translation between dissimilar conferencing technologies. This concept means that a user using one type of service - H.323 for example - could expect to connect to and communicate with another user that is using perhaps even a radically different type of service. This potential ability not only provides a bridge between differing technology that might be present at any given time but also technology as it has changed over time, extending the useful life of technology that was built on previous standards.

One of the most common gateways for use with H.323 is designed to transcode between H.320 videoconferencing with its ISDN transport and H.323 as it travels over IP. Since many campuses have already invested heavily in H.320 and have mature H.320 applications, it is advisable to consider a model for H.323 deployment that includes such gateways as a complement to the IP-based service. This could also permit existing and future communications with areas that do not have high-performance IP networks available or where ISDN may be a more affordable option. A secondary use for the H.320-to-H.323 gateway could be to provide redundancy for a LAN-based MCU service. Should a network break occur, a conference could be routed alternately from one MCU, across a local LAN, through an H.320 gateway over ISDN, back through a second gateway and onto the LAN local to the second MCU.

There are also common communication scenarios that call for the inclusion of traditional voice calls over the PSTN (public switched telephone network) in an H.323 communication. This may be as simple as needing to communicate with someone who has not yet implemented H.323 to extending H.323 services to users when they are mobile (e.g., using a cellular phone.) To enable this type of interconnection, an H.323/VoIP/voice gateway may be deployed. Other gateways - H.320-to-H.321 (ATM), H.323-to-H.321, and H.323-to-VRVS (Virtual Room Video Service, http://www.vrvs.org) also exist and should be included where needed.

Because gateways function between protocols and not within a single protocol, some special configuration may be required. In particular, the RAS (registration, admission and status) section of the H.323 specification, which permits dynamic conference ID registration, has no functional equivalent in the H.320 specification. The result is that if a gatekeeper is present, the conference ID must be pre-defined for multipoint calls. Point-to-point calls not using a gatekeeper do not require special treatment.

A second configuration issue to be careful of is that IVR (interactive voice response) systems often use the asterisk (" * ") to signal request for operator. In such an environment predefined groups intended for use with gateways shouldn't include asterisks. Unfortunately, this requirement conflicts with the trend among H.323-only users to utilize the asterisk as a delimiter.

Some CPU intensive audio transcoding can cause significantly delayed audio, resulting in an objectionable lack of audio/video synchronization. H.323 systems use G.723 and G.711 while H.320 systems use G.728 and G.711. G.711, the protocol in common, provides toll quality audio but uses 64Kbps. Disabling transcoding minimizes the audio delay due to the transcoding but would leave only 64Kbps available for video in a 128Kbps single circuit ISDN call. Enabling G.728-G.711 transcoding would reduce the audio bandwidth requirement to 16Kbps and free an additional 40Kbps for video. In a 384Kbps triple circuit bonded ISDN call, minimizing the audio delay might be deemed worth the minimal video degradation. Whether or not to permit audio transcoding should be decided on a call-by-call basis.

Supporting gateways can be operationally complex. Some service providers have recommended that users implement dual-technology codecs. For example, many group conferencing systems support h.323 and h.320. By recommending these systems, a service provider can centrally support the more desirable h.323 IP-based technology while allowing the user to manage (and pay for) their own dedicated ISDN BRI lines for h.320 compatibility. This has then benefit of ensuring the user has access to legacy technology if needed, but at a cost that encourages migration to more modern protocols.

Determining Your Users' Needs

The concept of "videoconferencing" is not monolithic or homogenous. It actually covers a large number of possible scenarios and is used in different contexts, to do different things. Standards such as H.320 and H.323 enable products to be developed that can communicate with one another but what combination of these products will work best for your end-users? Are you trying to provide videoconferencing facilities that will be used on a desktop computer? Do you need a monitor or multiple monitors that can be seen by more than one person? Are you implementing videoconferencing to enable distance learning? Are you trying to support research collaboration? The specific needs - number of participants, room layout, peripherals required - may be different for each videoconferencing installation.

This section presents some of the global issues you'll need to face when designing a system and walks you through some archetype videoconferencing setups for the most typical uses of videoconferencing: meetings, classroom applications, and collaborative activities. The functionality required within each of these can serve as a guide as you envision and design a videoconferencing system. Understanding your users' needs, along with the general issues within each collaborative context, should provide the tools you need to determine what to buy to make your videoconferencing initiative a success.

The section begins directly below with five points of discussion that are fundamental to any videoconferencing project. Following that, you'll learn how these elements come together in meetings, classrooms, and collaborative sessions, with diagrams and text to explain who's seeing what when as well as some tailored suggestions for each type. At the end of the section, we consider some additional "case studies" of videoconferencing as it is used in systems that ViDe members have designed, developed, and run day-to-day.

Videoconferencing Usage Fundamentals:

1. *Environment of the participants* - Potential environments include desktop, conference room, or auditorium. Will conferees be meeting from their desktops or will a few, small conference rooms be combined into a larger, "virtual" conference room?
2. *Role of voice/video/data* - Is voice transmission sufficient? If video is required, must participants be able to see detail or is video used only to maintain presence? Will graphical materials be exchanged and must participants be able to collaborate using the same application?
3. *Degree of interactivity* - How will the meeting be held? Will there be a single main speaker, as in a classroom environment, or will a productive meeting require that the participants be able to speak to other conferees freely and spontaneously?
4. *More than two participants* - Conferencing between greater than two participant sites will require use of a multipoint control unit (MCU) or multicast network services. In addition to the technical factors there are also human factors that change the way a videoconference works

when it involves more than two parties.

5. *Need for Gateways* - Is there a need to talk with people using endpoints on different protocols? For example, is one endpoint coming in via H.323 (TCP/IP) and another on H.320 (ISDN)? If so, conversion from H.323 to H.320 (and vice versa) will require a gateway to translate the signals.

Environment of the participants

Videoconferencing can happen from all sorts of places, from classrooms to offices to conference rooms to auditoriums. The environment in which users will confer impacts the complexity and cost of the endpoints that you'll need.

In office environments, one to three individuals conferencing from a PC desktop unit with an inexpensive camera will suffice. Close proximity to the camera is the key here.). If there is a need to show other objects in the room (such as a whiteboard) or to control the camera from a distance, then a pan/tilt/zoom (PTZ) camera can make a big difference. A camera that can auto-focus and be controlled remotely will enable details of objects to be increased as needed. For conference rooms and auditoriums, a PTZ camera is a must.

A second, and equally important, consideration for the environment is the set up of microphones and speakers. Echo cancellation is a feature that is quite important but may not be included. For systems with built-in echo cancellation, using the supplied microphone and speakers will be sufficient. For systems that have poor or no echo cancellation, a noise-canceling speakerphone can usually be purchased as a separate item and is well worth the cost if more than one person will be in the conference or if hands-free operation is desired. Another audio option for use with a desktop setup is an "operator-style" headset - or "headphones-and-boom-mic". These headsets are an inexpensive way to eliminate the echo problems that can disrupt a videoconference.

For conference rooms, multiple microphones can greatly enhance the quality of the conference. Having several desktop microphones for a conference table of several people will prevent the need for shouting during a conference or passing the microphone from speaker to speaker. Such microphones vary in functionality in that they may or may not require the user to activate the microphone to speak. More permanent ceiling microphones are also an option for a group conferencing space. See "Best Practices and Etiquette" and also the appendix "Developing a Productive Videoconferencing Room" for more information on this topic.

Role of voice/video/data, and the degree of Interactivity

Voice quality should, and usually does, take precedence over video during a conference. A conference is unsuccessful if one of the voices is not clearly audible. The bandwidth allocated to the audio portion of a videoconference is usually 16Kbps to 64Kbps, depending on the audio codec used. 64Kbps transmission typically gives a higher quality of voice since the level of compression is not as significant and because it has a wider dynamic range. However, this is not always the case as improved 16Kbps algorithms are now in production.

If the video is being used only to maintain presence or if a single speaker is featured in a largely one-way mode of speaking (often called "talking heads"), then an inexpensive camera can suffice. However, as mentioned before, when video detail is required or a high degree of interaction is necessary, a pan/tilt/zoom camera is a must.

The other side of this coin is the display of the video. If your system captures video using high-resolution PTZ cameras but then displays it on a small monitor, the effect of the quality video is largely lost. When video quality matters, or when participants will be sitting any distance away from the monitors, larger monitors should be used. The 32" or larger TV monitors can be used but they do not scale well to large conference rooms or auditoriums. Plasma display and digital projectors (rear-projection or "standard" front-projection) are a more expensive but likely more satisfactory option.

In addition to audio and video in a videoconference, data of varying types may be exchanged. Supporting the exchange of graphic and textual material requires the use of application sharing features. This form of exchange has been generally implemented in videoconferencing endpoints according to the requirements of the ITU standard, T.120. The ability to exchange data is sometimes even termed "T.120 support" after the standard itself. We will describe data exchange in the remainder of this section per the T.120 standard. For information on newer technologies and data sharing in general, see "Emerging Technologies."

There are two main types of exchange: presentation (application sharing) or collaboration (data collaboration). Generally, application sharing refers to one-way sharing, where one site is able to share slides, text, or even physical items (using a document camera) with others. In contrast data collaboration signifies the process whereby all participants in the videoconference can share control of a given application (for example a whiteboard application or PowerPoint) as if they were sitting in front of the host machine.

For application sharing, there are two basic options: 1) run the application sharing synchronously in line with the audio and video bandwidth, or 2) transmit the data "out of band" and asynchronously using a separate application. (As a low-tech alternative, if a videoconference is to have something like a PowerPoint presentation associated with it, the organizer can email the file to participants prior to the videoconference and prompt everyone to follow along simultaneously while they scroll through their local copy of the slides. The advantage of this is bandwidth savings and some ease of use. The disadvantages include the necessity of a computer at each endpoint along with the requisite software on it.)

Data collaboration is often enabled separately from application sharing to reduce the chances of security breeches or accidental destructive behavior. It is important to remember that when data collaboration is active, all people in the conference have full control of that application, just as if they were in front of that computer; anything the local user can do with the application so can the remote participant. Data collaboration has usually been limited to a fixed bandwidth for the transmissions (similar to the audio and video streams.) Thus the transmission times for sending and receiving data should be predictable.

As a further note about equipment, some facilities that will be used for both videoconferencing and data collaboration have two monitors. In such configurations the users can see the presenter in one monitor and the data on the second monitor. This setup gives a more natural interaction between the presenter, viewers, and presentation material. If an endpoint does not have two monitors, a PIP self-view can be inserted or the view can be toggled between the data view and the camera view at either the presenters' or the viewers' discretion.

More than two participants

If there will be more than two participants in a video call then there are two choices for handling the interaction: using an MCU or using multicast.

In the case of the MCU, choices must be made about budget and network. MCUs can range from moderately priced software-based MCUs (several thousand dollars to twenty thousand dollars) and hardware-based MCUs (from fifteen thousand dollars to over two hundred thousand dollars.) It is important to remember that software MCUs rely on host computers that must be fast enough to keep up with all of the video streams in the conference and that the load on the computer increases with the number of people in a conference. Stability and support of the host system should also be taken into account.

Networking infrastructure is also a major concern when it comes to hosting an MCU. For every person participating in a conference on an MCU, bandwidth is taken out of the total network where the MCU is hosted. For instance, if ten people were conferencing at one time at 384Kbps, it would require a total of 7.680Mbps. This would load down a typical 10Mbps LAN and further consideration would have to be given to any participants that are conferencing from the Internet or WAN links. For a site with T-1 connectivity, there would only be room for two (at most four) Internet clients to connect. Note that some higher-end codecs from various vendors include MCU capability (up to four sites) at low or no cost.

Multicasting as a means of supporting multipoint conferencing is an appealing option from a price perspective and also a networking point of view. Multicasting eliminates the need for an MCU and frees the network from the burden of one concentrated point of multiple single connections. However, this assumes that a multicast conference can be hosted in the first place, which may not be the case. Multicasting to the desktop, as a network service, still seems to be facing a number of issues that are slowing down deployment, particularly on networks between organizations (as opposed to within an organization). In order to host a multicast multipoint conference, you will need to verify ahead of time that each participant site multicasting capability. If even one of the end points does not have multicast capability then another means of enabling the multipoint conference will have to be found (such as tunneling a unicast connection to the multicast network.) See "Related Topics: What about Multicast" for more detailed information on how multicast works.

Switching gears to the human factor side, know that a multipoint call is likely to be more complicated for your end users. There are a few different ways that interaction in a multipoint conference can be managed. One of these is "voice switching". In voice switching, all participants receive video from the current speaker, who is determined by both the length and strength of the audio coming from his/her endpoint. The amount of time someone needs to speak before capturing the audio is configurable and, once configured, the conference can, for instance, switch automatically between a presenter and an audience member who is asking a question. Another mode for interaction management is called "continuous presence" (sometimes called "Hollywood Squares" for its splitting of the monitor into quadrants). This mode allows you to see all (or a reasonable number) of end point sites at the same time. In the case of H.323, continuous presence is implemented using the same bandwidth (within the same video "window") that would normally be available for the video coming from a single endpoint. Because of this, individual video windows that are part of a continuous presence call are reduced in size. A final more of multipoint interaction is that of "chair control", where one of the participants is designated as the conference chair and can switch both the audio and video at will using some form of call integrated into either the MCU or their own endpoint application. All of these modes can be a bit disconcerting at first. Training conferees to introduce themselves by name and to state their locations before launching into their actual dialog can help alleviate this somewhat.

Need for Gateways

Gateways are a necessity when a conference is to be held between two or more clients using different protocols. For example, if one client uses only H.323 protocol and the other client uses H.320 protocol, then a gateway is needed between the two to handle the translation. The location of the gateway depends on several factors. The client using H.320 protocol may opt to have the gateway at their end since this would eliminate long distance charges on their ISDN lines (which would quickly add up for each pair of phone lines.) On the other hand, the H.323 site may have a frequent need to communicate with H.320 sites or there may be times when the H.320 sites don't have the services of a gateway available to them. While the H.323 site may incur long distance charges, it may be unavoidable and worth the cost. Gateways are also a practical way of bringing in calls from regular telephone systems (POTS). For example, if someone is on the road on their wireless phone, they can still participate in a conference via a gateway, which bridges them into the conference. See Advanced Functionality and Management -- Gateways for more information.

Campus Deployment Issues

Although H.323 use in higher education is growing, it is far from being systematically and ubiquitously deployed. H.323 enables more end-user autonomy than H.320 videoconferencing, and the technology is not prohibitively expensive, but these very attributes make the need for coordinated and judicious deployment all the more essential. Typical faculty/researcher purchase is most likely going to be limited to end-points/clients, leaving the provision of MCU/Gatekeeper services to either the central IT organization on campus, or to an external service provider. Although by no means exhaustive, the following issues and questions are some of the more critical that need to be addressed in a H.323 campus deployment plan. Others, such as security and network matters, are discussed elsewhere in this Cookbook. The significance and scope of many of these issues became apparent in the ViDe Large Scale Video Network Prototype project (http://www.vide.net/lsvnp/index.html).

Coordination and Management of Local Gatekeeper Zones

The gatekeeper function is critical for call management as well as the provision of advanced H.323 services (e.g., multipoint calling, call forwarding, call transfer, gateway connections to non-H.323 endpoints.) According to the standard, if a gatekeeper is present on a LAN, clients must be registered with the gatekeeper. In addition, gatekeepers at different sites must be able to locate each other for to enable inter-realm gatekeeper-assisted communication. Some centralized management in the establishment and operation of local gatekeepers is necessary in order to avoid the presence of multiple and conflicting gatekeepers on a single LAN and to coordinate reliably with gatekeepers at other sites.

Coordination and Management of inter-realm Gatekeeper Zones

Current implementations of H.323 gatekeepers do not have an efficient means of discovering other gatekeepers in order to fulfill call requests for endpoints that are out of zone. Most rely on sending broadcast messages over the network or on the maintenance of internal tables that contain necessary information on other known gatekeepers (similar to host tables on the Internet before there was DNS, or on "next hop" tables in a router.) Neither of these methods can scale to enable global H.323 deployment. Until the H.323 standard evolves to address this problem, the interested community needs to develop and collaborate in "workarounds" that can accomplish the immediate purpose of inter-realm dialing while also illuminating how this could or should happen in the future.

One of the most well known examples of collaboration in this area is ViDeNet (http://www.unc.edu/cavner/videnet). As of the first quarter of 2002, ViDeNet is enabling over 100 individual gatekeeper zones from nearly as many separate institutions to inter-operate with one another, approaching "seamless" H.323 inter-realm communication. The dialing schema and discovery mechanisms in ViDeNet also form the basis of the recently formed national R&E collaborative environment, the Internet2 Commons (http://www.internet2.edu/html/commons.html). As H.323 presence and capability continues to grow, those who are deploying H.323 for their organizations need to become aware of efforts such as these and align their H.323 service with them to enable the largest scale of communication possible for their users.

Management and Coordination of Directories

H.323 calls in a basic or early implementation are often based on dialing the IP address of the end-point. If calls are supported in this way, there is obvious need to publish at least the names (or descriptions) of available endpoints along with their IP address and gatekeeper assignment to enable communication beyond the community of endpoints that any given user might already know. In addition, in implementations such as ViDeNet and the Commons mentioned above, there may be additional aliases or a suggested dialing schema in place to make placing calls more "user friendly". In that case, the aliases or dial string would also need to be published for each available endpoint. In both cases, the existence of an easily maintained and accurate directory of users matched to their "call attributes" is highly desirable. At some point in the future, core middleware that offers identification and authentication services and interconnection between directories (a "directory of directories") will help fulfill this need in a secure and automated way. Until then, those deploying H.323 services for their institutions will need to manually maintain and coordinate such information. Additional information on the development and evolution of middleware services specifically for digital video and videoconferencing is available in the [list the keeping with the changes section]

Scheduling Support for Specialized H.323 Services

Components that support specialized H..323 services, such MCUs and gateways, are often a limited and shared commodity on an H.323 network. In order to use these services most efficiently, H.323 service provides may need to enable a scheduling mechanism for such services, ensuring that they are available when users need them. (The alternative to this, of course, is to over-provision the service, as long as there is some tolerance for a high degree of idle time in exchange for availability of the service during peak request times.) Many of these components today do not have scheduling built-in or may not have scheduling mechanisms that fit in well with the scope or all of the goals of shared use (particularly inter-institutional shared use.) The organization that is providing the services may want or need to build their own scheduling system to manage such services.

Provision of Assistive Technologies

Consideration should be given to whether voice-activated MCUs will support all use on campus, or will there be need for a MCU with continuous-presence features, - for our deaf and hearing-impaired constituents, for example? Are there other instances of special needs on campus that should be met?

Technical Support

Provision must be made for the same level of technical support, troubleshooting, client recommendation, installation and deployment with as any production technology. This is probably the single greatest cost factor in the deployment of H.323 services since there are many aspects to what may need to be provided. Proactive support requirements, such as training, as well as breadth of necessary helpdesk knowledge may be especially heavy because videoconferencing represents a whole new concept in computer use than most individuals and IT organizations are used to. A very low support staff-to-user ratio is best particularly in the early stages of deployment.

Cost Models

Since videoconferencing over the data (IP) network is somewhat a hybrid service between information technology and traditional telecommunications services, choosing between a charge-back cost model and a centrally funded service model can become an issue. The operational costs likely to be incurred in establishing and maintaining a H.323 network are evident throughout the sections above. Operational costs aside, however, the potential to integrate the telephone network with a H.323

network will obviously have the effect of reducing the funding stream associated with traditional telephone services on a campus. This issue, needless to say, is controversial, and demands careful consideration. It is also being even more hotly debated in connection with voice over IP (VoIP). If shifting cost models become a determent to wide-scale or manageable H.323 deployment, those in charge of such deployments are likely to find studies of VoIP implementations useful in helping to find ways to get beyond these issues.

ViDe.Net: An International Videoconferencing Network

ViDeNet is an international virtual network providing video teleconferencing, telephone and collaboration services over the Internet, Internet2 and related advanced networks. Now you can begin to offer innovative new communication services, increase your project team's communication and dramatically cut telecommunications costs to your organization. And ViDeNet is an open forum, like the Internet itself, so there are no costs or barriers to participation.

ViDeNet was formed by ViDe, the Video Development Initiative, to be a testbed and model network in which to develop and promote ViDe's goals for highly scalable and robust networked video technologies. It is quickly evolving into a global virtual network interconnecting advanced voice and video networks around the world. ViDeNet serves as a forum for discussion and an experimental testbed for ideas related to improving the state of inter-networked video and voice over IP architectures.

From a technical perspective, ViDeNet is a mesh of interconnected H.323 zones. Each zone represents a collection of users at each site that are administered by the site itself. Thus, ViDeNet provides a mechanism by which individual campuses and network providers can interconnect themselves, creating a seamless global environment for teleconferencing and collaboration. ViDeNet's underlying technology moves forward as the state of the art of voice/video communications develops, incorporating technologies such as, QOS and policy-based networking.

Goals of ViDeNet

Dramatic Cost Savings

It's no secret that the key to today's information economy is an organization's data network and the Internet that connects it to the world. But while companies are investing heavily in the internal infrastructure and personnel to support this asset, they still are dragging along huge costs for telephone services , costs which are lost forever to the telephone company. ViDeNet provides a way to eliminate this redundant network and lost revenue by carrying telephone services over the data network. Savings are realized not just on long distance calls, but also on "local" services where most of the costs are hidden.

Increased Collaboration, Communication and Organizational Effectiveness

People everyday are using ViDeNet to manage projects, teach students, troubleshoot problems, provide medical services, and talk with friends and coworkers around the world. All of these activities happen better, faster and with increased effectiveness when the people involved are in regular contact. And communicating with video brings added warmth to human interactions - people feel more in touch when they can see each other with facial expressions and gestures.

Innovation in Communication Services

The telephone hasn't changed much in the last 100 years, but computers sure have! Communication via the Internet offers possibilities for interaction just not possible until now. Imagine reading a medical record, clicking on a section of particular interest, and the person responsible for that entry automatically appears on your screen so you can ask follow-up questions! This kind of development is possible under the open ViDeNet architecture. By moving service platforms off of the telephone network and onto the IP network, ViDeNet creates an environment in which innovation is fostered at the same rate as the Internet and the Web.

A Single Voice and Video Network for Earth

What if you tried to call your friend who lives down the street, but couldn't connect to her because she used a different telephone company? Not very useful! ViDeNet provides a way for all h.323 voice/video networks to connect themselves together, so anyone can talk to anyone else! How else should it be?

Multi-Vendor Networking

ViDeNet's open technology platform allows you to choose the right mix of technology, and change platforms as your needs evolve. We currently have demonstrated interoperability among equipment from the following vendors:

- CISCO
- Ezenia
- Intel
- Microsoft
- PictureTel
- Polycom
- RADVision
- Siemens
- SUN Microsystems
- VCON
- VTEL
- Zydacron

Participation

Standardized Dial Plan

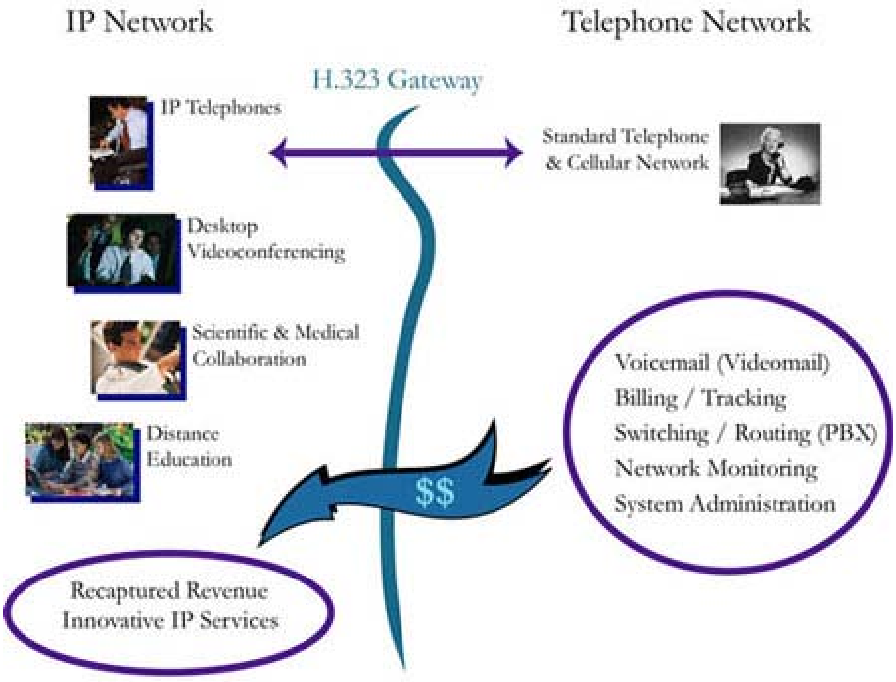
ViDeNet provides a standardized plan for naming telephones, videoconferencing stations and zones on the network. This naming scheme includes both telephone-like numbers and email-like friendly aliases. Users on the old telephone network dial your unique telephone number and never need to know you are actually on the Internet. Users at computer terminals can dial you using a friendly name like John_Doe@jupiter.edu.

Tools For Getting Connected

ViDeNet provides tools for directory and registration services. It also provides tools that allow network administrators to keep in touch with one another and to stay abreast of changes on the network.

Forum for the development of h.323 technology

One of the key elements of ViDeNet is that it provides a forum in which network administrators can work together directly with researchers and developers to improve the way the network works. This has already resulted in efforts leading to improved security, ease of use, account management, billing and advanced functionality.



ViDeNet allows you to migrate services over to the IP network where costs are invested rather than lost, and innovation allows users on a variety of different telephone, desktop and conference room platforms to communicate together, without losing access to legacy telephone and cellular telephone systems.

Related Topics

Tips for Selecting and Tuning your PC

Selecting and upgrading a PC for videoconferencing depends on the following factors:

- Operating system
- Videoconferencing end station

- Other applications you plan to support on the same PC
- Processor power and memory
- Processor load from existing applications

Your choice of videoconferencing end station will be dependent on what operating system platform the application is developed for, and what operating system you are familiar with and comfortable using. Consider also the intended use for this machine — do you plan to do person-to-person conferencing? Will your videoconferencing PC screen be required to support a display of multiple users? There are many solutions available for PC users, but the market is still short on Unix and Apple solutions.

Types of desktop conferencing solution:

Video Appliance (USB connected)

PCI based solution (internal hardware codec)

Software (generic camera, microphone and sound card)

Be wary of over subscribing your processor with applications other than the videoconferencing application. The need to do data sharing, and the applications that are required to support the shared data, must be factored in when selecting an endstation system. Knowing the impact of a personal videoconferencing solution on your desktop requires careful review of product specification, including application interactions and interoperability issues.

Examples of PC system requirements to support videoconferencing applications

Video Appliance

Assumptions for this type of solution:

1. Appliance will handle all compression and decompression.
2. Appliance is USB connected.
3. Appliance may or may not have embedded microphone with echo cancellation.
4. PC is already provisioned to support the OS platform requirements

System requirements

- USB port
- Windows 2000, XP, or 98, Second Edition
- 350+ MHz Pentium-class processor
- 128+ MB RAM 4MB video memory
- 20MB available hard disk space
- SVGA display (800x600) 16-bit color or higher
- headphones, headset, or echo cancelling microphone/speakers

PCI based solution

Assumptions for this type of solution:

1. PCI card/ codec board will handle all compression and decompression

2. Card is used in conjunction with vendor provided software

System requirements:

- 400+ MHz or higher Pentium PC
- 128+ MB Ram
- 1 PCI slot
- Win NT, 98, 2000

Software solution (generic camera, microphone and sound card)

System requirements:

- Intel® Pentium or AMD® Athlon® processor running at 266 MHz
- Windows® 98, 2000, or XP
- 64MB RAM
- 70MB available hard disk space for installation of all software titles
- Available USB connector and CD-ROM drive
- Display adapter capable of 16-bit color at 800x600
- Sound Blaster or other Windows-compatible sound card and a microphone are required for videoconferencing (sold separately)
- Live videoconferencing is possible with a number of third-party applications such as Yahoo! Messenger, CUSeeMe, Microsoft NetMeeting, Web Clients, and MSN Messenger

Questions to ask yourself (or your system administrator):

Does the system configuration for my PC really matter?

The videoconferencing components are housed inside a desktop computer or workstation; if the computer system is not powerful enough to support the videoconferencing hardware and software the videoconferencing terminal will provide poor performance. Once you have selected a particular videoconferencing product, be sure to review the vendor's PC specifications. Remember that the vendor will specify the MINIMUM requirements (usually operating system, processor speed, amount of random access memory, video display board, and video memory). In short: Overbuild

Can't I just get by with the minimum system configuration?

MINIMUM requirements should be interpreted to mean that there will be no other programs active on the PC while the videoconferencing software is running. If you anticipate that the videoconferencing PC will have an e-mail program running, a web browser open, and be playing music from the CD player while you are videoconferencing, you should select a PC that exceeds the minimum requirements. Random access memory and choice of video board/video memory have the greatest impact on videoconferencing performance. The video board is important because the videoconferencing hardware displays the camera image on the computer screen by using the computer's video board. If the codec is very fast but the computer video board is slow it will degrade overall performance.

A bunch of folks at my organization frequently need to be in the same videoconference at the same *time. I suggested we all "squeeze in" around my desktop but everyone else just groaned. Don't you think that would work? Who's right -- them or me?*

Well, it's certainly possible to squeeze together in front of the desktop so it's hard to say who's right. I guess it depends on your definition of "would work"! Unless your desktop camera has a wide-angle lens, you won't be able to fit everyone in front of it and the remote site will not be able to see all of you. Also, depending on the size of your PC monitor, people at your end are likely to have to strain to share the view. Sound might work out O.K. for a small group. Microphones and speakers often "stretch" better than the video peripherals do. Still, it's probably better to admit defeat on this one. Once you have more than a couple of people who want to conference together, you either need to bring everyone into the conference from their own desktops using an MCU or set up a group conferencing area where everyone can meet around a group sized system. Remember, makin' do isn't always makin' sense ;-)!

I heard that I could spend just a little more money on the audio/video components for my PC and turn it into a group conferencing system. If that's true, why do group conferencing systems cost so much more?

Imagine you have a VW beetle but you'd like it to carry several passengers and their week's worth of camping gear and drive over rough terrain, like a Ford Explorer. You have a chance of making it work if you a) understand the design limitations of the Beetle, b) understand the design goals of the van, c) have at least some money and time to spend, and d) are handy with tools and improvisation. Same idea. If you can build it yourself, you might be able to save some money. However, if you're not the type to re-engineer something and then support what you have re-engineered, buy a group system "out of the box"!

Sometimes my PC seems so sluggish when I'm doing anything else during a videoconference. Where is processing typically done on a VC system?

Where your video processing occurs depends on the client you are using. Some clients come with a special add-on board that will offload some or all of the video work. Some clients, generally the cheaper ones, will rely on your main processor to handle the video. Therefore slower PCs may see worse performance during application sharing. The morale here is, if you want a cheaper videoconferencing client, install it on a faster PC.

What if I don't use a PC?

You are certainly in a "no-Win" situation! There is a notable gap in the H.323 market in terms of both UNIX and Macintosh video terminals. This is due to the large, general consumer market for Windows systems. Multicast tools are available for Unix; H.323 products may surface soon as well.

I have a really nice, fast PC. It has two processors in it as well. Why can't I use this system with a board assisted videoconferencing client?

Videoconferencing assist boards must go into certain PCI slots. It appears that vendors have programmed their software to address a particular range of IRQ numbers. The second processor typically fits into a slot higher up and therefore throws off the IRQ numbers for the videoconferencing board. Should you install the videoconferencing assist board in a dual processor PC, mayhem ranging from refusal to operate to constant, regular video freezes results.

Just when I thought I was handling the Information Age pretty well, the other day I got all flustered. It all started when my videoconferencing client "rang" at the same time that my email beeped, my telephone rang, and someone stopped outside my office door. Each interface seemed to demand that it be "first". I let the phone go to voice mail, left the email for later, took the video call but asked them to "hold" while I talked the in-person person. But this sudden crisis of communication left me very confused. For the rest of the day, I kept trying to do strange things - Drag and drop a phone number from the address book on my PC to my telephone. Eat a bagel that was sitting on the conference table of a remote site I was videoconferencing with. Pan my telephone handset around the room to "show" the caller my new office arrangement. I even became convinced at one point (and quite frustrated thereafter!) that I could cut and paste a good joke into a colleague's mind. Am I going crazy?

You're not nuts, just harassed as well as maybe a little ahead of your time (I mean, really, cutting and pasting into peoples' minds??!) Someday it will all come together. For now, go outside for a nice quiet sit. Don't take the cell phone, the pager, the PDA, the laptop, your pile of reading, your Dick Tracy watch, or your Maxwell Smart shoephone. Just stop. Look. Listen. Learn to be still.

I ordered one of the new USB videoconferencing clients, the ViGo. It has some really nice features, but I'm having some trouble with the audio. Do you have any advice for me?

The VCON ViGo audio quality can suffer if the audio volume is too high or if 'Automatic Gain Control' is on and there is a background noise at the same time. There are no audio problems at all if a headset is used. Strange behavior has been observed if 'Allow Adaptive Bandwidth Adjustment' is on in a conference with a ViaVideo; ViaVideo decreased continuously its Video-Bandwidth from about 300 kbit/s to 10 kbit/s. In a connection to Tandberg 1000 the ViGo reduces its Video-Signal to QCIF (as soon as the connection is established it can be switched back to CIF manually). If the mic is muted before a connection is established - the mute is cleared as soon as a call is received. (this may cause a privacy problem which a user is not aware of).

Developing a Productive Videoconferencing Room

Any current conference room can be adapted for use as a videoconference room by making adjustments based on the needs of video and audio equipment to capture signals. This is less of a concern for new construction, as these details will be an integral part of the function of the room, and will be designed in by the architect. More probable will be the conversion of an in-use conference room for videoconferencing. The advances in technology have made the concept of an in-house video studio an attainable communication tool. As in movie and television production where the sound stage is a critical part of the process, the conferencing room is a critical part of productive videoconferencing. The walnut paneled conference room is not the most conducive atmosphere, and creates a challenge for video and audio capture. There are several adaptations that will enhance the videoconference as a useful communications tool.

A basic understanding of audience and purpose are key to successful room design. Generally, most videoconference rooms are designed for 10 —25 participants, for applications such as:

- university/ executive departmental business,
- distance education as primary means of content delivery,
- distance education in the form of a single class event that involves demoing the technology or

enlisting an expert opinion or guest speaker from another educational institution or private industry.

The most difficult obstacle is maintaining a balance for the camera. The background colors and lighting will affect the view as seen by the remote participants. To be able to see all participants clearly, wall or discrete floor light sources have supplemented traditional ceiling light sources. Lighting is one of the few critical factors to successful videoconferencing.. Lighting consideration for the intended room will factor heavily into the choice of wall coverings and table surfaces.

Lighting

Lighting in the videoconferencing space is key to delivery of a good quality image. Elements such as lighting level, lighting angle, and overall color temperature should be considered. For lighting levels, 125-foot candles and above are recommended. Light levels between 75- and 100-foot candles provide good balance and camera performance. (Note: A foot candle is the luminance on a surface one foot away from a candle's light.) Most modern videoconferencing cameras will render a good quality image at theses lighting levels. For angle, keep the lighting source in front of participants at or above eye-level. Diffused light versus direct light can help with glare or excessive "whitewash"; consider installing diffuser on existing light fixtures. For temperature (expressed in Kelvin), consider that optimum color temperature of light vary across each video input and conferencing codec. The majority of today' systems will perform best using 3200k.

While there are several concepts popular with designers, one key design parameter appears throughout all the recommendations. To eliminate shadows, a combined lighting arrangement ratio of 60/40 for ceiling and wall lighting is recommended. Wall lighting should be indirect and these fixtures are readily available from a wide range of suppliers. The key in this split lighting scheme is to equalize the available light on the participants and eliminate shadows, dark backgrounds, and bright spots in the center of the conference table.

Interior Room Design

When considering the interior design of the videoconferencing space, the primary goals should be to make the room as comfortable as possible, de-emphasizing the technology in the room and making the user interface to the system uniform and predictable. Specific colors are recommended for backgrounds and wall covering to enable better recognition of attendees without straining the capture capabilities of the video camera. Recommended colors are soft, textured wall coverings, but smooth painted walls will work if colors are muted earth tones and the lighting is adjusted to suit. When considering furniture and walls be aware of color and physical characteristics that may make your video or audio input have to work harder.

Acoustics and Audio

Audio technology has developed to a level where only the obvious interference from air conditioners, telephones and other extraneous noise sources would factor into microphone placement. While most conferencing systems use the audio out path to the speakers installed in the monitors, there are separate speaker systems available to meet the needs of larger room sizes. Modifications to rooms need not be expensive. Employing simple techniques like wall-to-wall carpet, acoustical panels in the ceiling and walls, and full coverage window covering are very effective in optimizing system performance. Many new systems come with omnidirectional microphones, with echo cancellation, capable of picking up sounds from all sides. The room application lends itself to the complexity of the audio in and out of the system. Larger spaces will require more microphones, audio mixing resources, and amplified speakers for coverage

Room selection

New videoconference rooms that are fortunate enough to be conceptualized in design from the start, and are dedicated specifically to the application, are easy to implement. Due to the growing acceptance of the technology, videoconferencing can be found in universities, K-12 classrooms, hospitals, private industry, and courtrooms. The challenge here is in the conversion and adaptation of these legacy spaces to suitable for videoconferencing. Consider

The next concern is the size of the room based on the available space. The videoconference is directionally oriented by the visual focus capabilities of the camera and factors in to room layouts. Allowances must be made for furniture, additional wallboards etc. The size of the attending group is not only dependent on the actual room size. A room layout will determine how many participants may attend. The actual seating arrangement is then defined to allow the participants to see and be seen through the conference. There is a minimum distance required for the camera to capture all of the attending participants and must be factored into a layout. Furniture manufacturers have developed conference tables specifically designed to allow meeting attendees to see and be seen by the video equipment. There are several sources available for specialized video equipment including custom conferencing tables and matching cabinets. The best capture angle for the video camera is a "down the table view" with the end seat closest to the camera empty. This avoids having an attendee in that seat, who can neither see the monitor nor be seen by the camera, and permits the assembled group to view the remote part of the meeting. This arrangement also creates a clear walkway into and around the table, and creates an aperture distance for the camera without unnecessary waste of available floor space.

Videoconferencing equipment does require room. There is a monitor required for receiving a conference, and where the H.323 terminal's system does not provide a screen in screen option, a second monitor is needed for the "self" image portion of the videoconference. Some conferencing equipment uses additional equipment requiring space. Most conferencing cabinets allow for the housing of this equipment in the base and placement of the monitor at an easy to be seen height on top of the cabinet. Additional cameras for enhanced teaching situations, with an additional monitor would also factor into the space considerations in planning a videoconferencing room. These design factors are dependent on the requirements of the manufacturer and the available space in and near the proposed conference room. There is a sliding scale for required space. Most equipment can be located within the room and installed within a finished cabinetry. An ideal case scenario would be an adjoining "mechanical" room to house the associated equipment, leaving the monitors, camera and microphones the only physical presence in the room. With the advance of the flat screen monitors, this presence will be diminished in the future as these new monitors will require less space

For the issue of reliability and "clean" power, most manufacturers recommend individual service circuits for the equipment. Due to the influx of sensitive high technology equipment in business, most commercial real estate space has "clean" power lines available for computer, communications and operational equipment. The actual link is made over ISDN (digital telephone lines) or the computer network and these lines are currently commonplace in 99% of commercial real estate dedicated to business use.

Attending to a few critical details will develop a modern videoconference room and there are designers who specialize in these concepts. These concepts permit development of a comfortable, functional videoconference room that meets the physical needs of the equipment and accommodates interior design tastes in the intended work environment.

What About Multicast?

To understand how multicast can possibly help videoconferencing in general, it is useful to understand the transport background and mechanisms used by "normal" or "unicast" applications such as H.323.

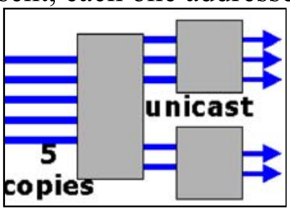
How H.323 traffic travels the Internet

H.323 videoconferencing sessions travel across the network on top of a network transport layer known as IP. The H.323 standard uses two types of IP transport: TCP and UDP. TCP is designed to guarantee that the data arrives in full, in its original condition. UDP is designed to get most of the data to the destination most of the time. When you purchase something online on the Internet, you are using the TCP transport. Both you and the vendor want to be sure that your order is received completely and accurately. You would not be happy if a few bits were changed in the amount charged to your credit card. In the event that some error occurs during the data transfer, the transaction can be repeated over and over again until it is done correctly. In contrast, suppose you are watching a live broadcast of a sports event. If there is a glitch and a few frames get dropped, you probably don't care. If you had the last frames sent again and again until successfully delivered, what would you do with them? See them out of order? Stop live recording and wait? This type of transport is known as UDP.

The H.323 standard requires the use of both TCP and UDP transport. TCP is used for control and data sharing such as file transfer. You do definitely want to be sure that sessions are set up correctly; you do want to guarantee that there are no errors in the transmission of a document. UDP is used when sending video, audio, and status information. Most of the time, most of this type of data arrives correctly; when they do not, we don't care unless the percentage of missing information becomes large enough to be noticeable.

Efficient Network Transport

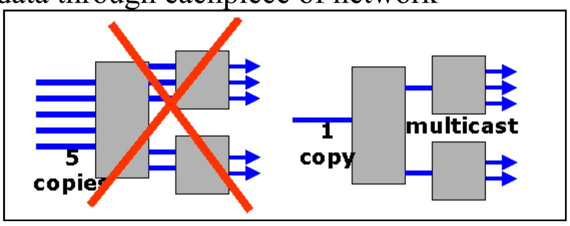
H.323 sessions are typically "unicast", meaning that one copy of the transmission is created for and addressed to each unique end-user. If there are 5 users participating in a session, 5 copies must be sent, each one addressed to a different end user. The data is deconstructed into packets, each of which



carried an end-user IP number as the destination address. Since IP numbers are unique, a separate copy must be sent to each end-user. In any stream of data, including a stream of video data, the actual encoding of the data can be separated from the details of the transport mechanism used. Today, the H.323 standard calls for TCP transport of control and data; therefore, unicast is required. However, it is possible to substitute a different transport mechanism for TCP. Substituting one transport mechanism for another should have no noticeable effect (from the end-user's point of view) on the quality or the videoconferencing session; what the network sees, however, are two entirely different traffic patterns.

What is IP Multicast?

IP multicast is a bandwidth efficient way of delivering data, such as video and voice, to multiple recipients using a single copy for all rather than one copy each. The network can more efficiently transport the information by sending a single copy of the data through eachpiece of network equipment. Rather than addressing a unique device, multicast packets are addressed to a special set of IP addresses, known as Class D addresses (the block of IP numbers from 224.0.0.0 to 239.255.255.255). This "multicast group address" serves as a virtual channel; the end-user selects the channel by selecting the appropriate



address, the network (hopefully) configures itself to deliver the multicast traffic, and the user then receives the data stream.

For the network to "configure itself" for multicast traffic delivery requires the use of some multicast specific protocols. These are outlined below.

Multicast makes efficient use of the network by making sure that few, or no multicast packets are sent to a router unless some end user behind the router has made a request to send or receive. Then, only the requested programs are allowed to pass through. Selecting a multicast broadcast is known as joining a multicast group. When you join a multicast group, your request is sent back through your router; the router sends a request towards the source(s). These actions build a delivery "tree" through which a single copy of the multicast is delivered. The end user may experience a pause between requesting to join a multicast group and the start of the requested stream of data due to the time it takes to build the delivery "tree".

Why are universities, as well as other ISP's, interested in conserving bandwidth? Even if universities are able to provide "unlimited bandwidth" on campus, off-campus connections are usually arranged through some commodity internet service provider, and that connection is an expensive one. Access to the Internet at large, even to Internet2, is usually a bottleneck in the network (the point with the most limited bandwidth). Videoconferencing can be used across campus, but the more typical application is to communicate with colleagues at a distance - off-campus. Anything that can be done to conserve bandwidth at the bottleneck is going to be cost-effective.

Multi-point Sessions without an MCU

Earlier in this cookbook we described point-to-point and multi-point sessions. We stated that multi-point sessions require an MCU to receive and re-broadcast the session to each individual participating in the multi-point conference. However, IP multicast makes it possible to engage in multi-point conversations without the use of an MCU. Since MCUs are very expensive, it is easily apparent that multicast offers a more cost-effective approach to multi-point videoconferencing by using the existing router infrastructure.

In a multicast conference, a Class D address is assigned in advance. Since there is no global repository of who has used what address, there are some interesting issues in defining how you know whether an address is "unused" or not. There are various address allocation schemes in use, but none have achieved global acceptance. Some tools, like the Session DiRectory (SDR), have a built-in semi-random mechanism that works quite well. Given the size of the address space, and the few current users of multicast, the chance of collision is small — and if it does happen people usually realize very quickly and pick a new address.

Ok, suppose you want to create a multicast session. Everyone participating in the videoconference sends out network packets that are addressed to the same Class D address. When using an MCU, each participant's VC is transmitted to the MCU that acts as a server, re-broadcasting the data to all connected participants. In multicast, each user's data is broadcast directly from each user's VC system to all other participants, without need for a central server.

As an end-user of multicast, these details are hidden from you. In some applications what you see is a user-friendly interface that presents you a list of available broadcasts, much like a TV guide. You see which sessions are currently running; you see which sessions are scheduled in the future. If you want to create or announce a new session, you click a button and fill out a few fields. To join a session, you click on one of the entries in the "TV Guide".

Without an MCU, or any central authorization mechanism, participants are free to add or remove themselves from multicast conferences, without having to be pre-authorized through a gatekeeper. You currently cannot create a truly private yet global multicast conference, in particular you cannot restrict receivers or multicast "join" requests in any globally scalable fashion. There are a variety of research efforts in this direction but they have some ways to go. Some commercial tools that use multicast also provide a password mechanism to access a directory service, but that security is weak and is usually implemented in a proprietary fashion.

If a session requires security, whoever establishes the session can establish the security and send passwords, effectively encryption keys, to selected participants. Other people can join the conference, but cannot decrypt the content. The security is then as strong as the encryption chosen and the encryption key.

What hardware, software, and network infrastructure are required to support IP multicast?

Gee, if IP multicast saves bandwidth and eliminates central administration, why doesn't everyone just use it? There are 3 major problem areas: the wide-area network, the campus/local network, and the tools.

In the wide area only a few commercial carriers offer multicast services, and in some of those you have to find the right person to talk to! Conversely, most of the global research and education networks offer a multicast service.

On campuses, networks can be made multicast-capable, but few have done so. There is a general wariness of multicast, most (but not all) unfounded, and campuses don't often have the more modern network equipment required (see below). Very few campuses can deliver multicast traffic "anywhere" but the numbers are slowly growing.

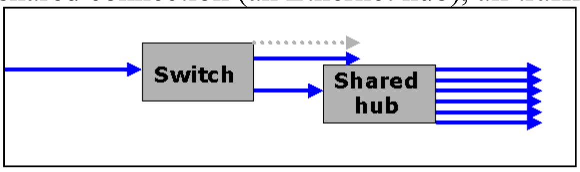
Many campuses that do offer multicast services "everywhere" currently do so using proprietary network protocols. Just as proprietary H.323 protocols allow you to achieve sophisticated levels of H.323 videoconferencing at the cost of losing inter-operability, proprietary multicast implementations allow you to deliver multicast traffic, but at the cost of losing inter-operability. Ideally, a network implementation of multicast support would be vendor-neutral. Large campuses purchase network equipment over a period of time, often from different vendors, and thus rely on adherence to standards to assure cross-vendor compatibility at some predictable level. Even if your campus has managed to achieve single vendor, single generation equipment purchases, you will be communicating with colleagues at other institutions who may have standardized on some other vendor's equipment. Adherence to standards may limit you to lowest common denominator performance, but it's predictable and reliable.

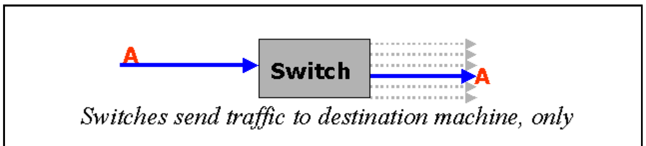
Presuming you've got the first two solved, and there is plenty of help around to do that (see below and elsewhere), that leaves the tools. These have mostly come from the University community in the past, and reflect a great deal of engineering effort — except in their user interfaces, documentation, stability, … Fortunately the commercial world has started work on multicast tools and developed something you can comfortably offer users. Most of these products do not use H.323, although some do interoperate with H.323 products. Some vendors, including VCON (http://www.vcon.com/) and Lucent (http://www.lucent.com/enterprise/ipapps/conferencing/), sell H.323 systems that support IP multicast. These offerings have all the gotchas and limitations that you would expect in Version 1 of any product, but they will improve in the future. A free version of the IP/TV multicast viewer, with 1 year license, is available to Internet2 member institutions at http://netaid.uoregon.edu/. IP/TV allows you to watch and listen in on multicast conferences, but it doesn't allow you to contribute; it can operate at higher resolution than the public domain clients can. A nicely bundled set of MBONE tools for Windows (MASH) is available from UC Berkeley at http://bmrc.berkeley.edu/bibs/download/index.html. Another bundled set known as Shrimp is available from http://www.ja.net/development/video/shrimp/. The Unix/Linux community will find tools at University of Oregon Video Lab and the Internet2(tm) Networks Multicast Trial and Setting up MBone Tools for Windows95/NT, Macintosh and Unix.

I want multicast on my campus — what do I need?

You need to talk to your network staff.

Start with the local subnet. Use of switched rather than shared Ethernet connections is preferred. In a shared connection (an Ethernet hub), all traffic that enters the hub is broadcast to EVERY connected device. If you have incoming or outgoing video traffic, it is going to be sent to every connected station, which will have to decide whether to keep or toss the transmission. A switched network connection is better; traffic is sent only to the destination PC, not to every connected device. There is one "small problem", however. At the Ethernet layer multicast addresses are "unknown" to the switch, so it broadcasts the traffic to "everyone", so multicast traffic can in effect turn your switch into a shared hub!



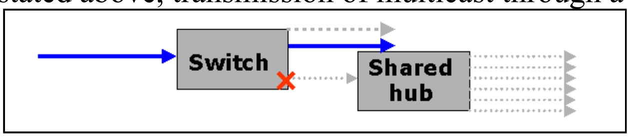


Switches send traffic to destination machine, only

Multicast broadcasts can turn your switch into a shared hub

To avoid this problem, the IGMP (Internet Group Membership Protocol) protocol is used. IGMP allows an end-user's PC to request to JOIN a multicast session or to LEAVE a multicast session. If the switch supports IGMP, it will know to send multicast traffic only to ports where the end user has requested a JOIN, and the switch will ignore ports that have not joined, or that have left, a multicast session. In short, the switch must be "IGMP aware" to be truly useful. IGMP is currently at version 2 in general deployment. A new form of multicast called Source-Specific Multicast (SSM) has led to the development of IGMPv3, but no Ethernet switch vendors support it at this time.

Switched Ethernet connections may have both end-users and shared hubs plugged into them. As stated above, transmission of multicast through a shared hub may prove to be unmanageable, so it may be desirable to allow multicast traffic to pass through a switch to individual end users, but be blocked from passing through to the hub. Ability to control multicast at the port level is a desirable switch feature.



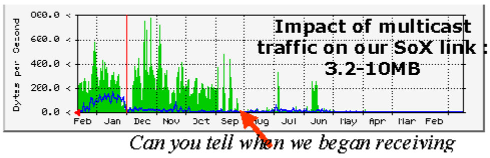
At the edge of your campus network, the routers need to interface to the wide-area networks and exchange multicast traffic, and then distribute it across campus in an effective fashion. Unless you have a completely flat Ethernet network on campus, your routers will also need to handle multicast traffic. The current standards recommend the use of MBGP/BGP4+ and MSDP at the edge of your campus network, and PIM, in particular PIM-sparse mode (or PIM-SM) across the campus.

MBGP or BGP4+ is the routing protocol used widely across the Internet today with enhancements for multicast. It allows network engineers to identify appropriate paths for multicast traffic to take, and for the router to understand it needs to handle multicast traffic correctly. MSDP provides effectively a "channel discovery" service between management domains (or "autonomous systems/AS"), which is not provided by the more rudimentary local-area multicast protocols. PIM-SM (and friend "Dense-

Mode" PIM-DM) actually works on the delivery of multicast traffic across the network, using information provided by MBGP, MSDP and IGMP at the very edge.

Recommendations from the Internet2 Multicast working group and NLANR are to configure your campus edge router to accept PIM-Sparse Mode traffic (only requested broadcasts transit your campus). Recommended campus settings include both PIM-Sparse Mode and PIM-Dense Mode. PIM-DM is suitable for campuses where end-users densely populate the network and there is plenty of bandwidth to spare. To the left, you will find a graph indicating the impact of multicast traffic on the University of Alabama at Birmingham (UAB) campus backbone under PIM Dense Mode. PIM-DM is a configuration which allows ALL available multicast traffic to flow through ("FLOOD") for several packets (one packet per source), then any connections not being watched within the campus are stopped ("PRUNE") for several minutes. The net result on a bandwidth utilization graph appears to be a continuous flow of 3-10Megabits per second of traffic for an Internet2-connected network. PIM Sparse Mode sends multicast traffic only to end-users who have "joined" a session. PIM-SM is recommended when there are only a few multicast receivers and bandwidth is to be conserved. Very excellent information on IP multicast and associated network architecture issues can be found at the NLANR engineering site (http://www.ncne.nlanr.net/faq/multicast.html).

Impact of uncontrolled multicast traffic could be severe.



Can you tell when we began receiving native multicast traffic? YES!

Broadcasting and Archiving Videoconferences Using Video Streaming

By combining video streaming and videoconferencing, it becomes possible for any number of people to view a videoconference — either live or recorded for on-demand access at a later time. A streamed videoconference can be viewed on desktop PCs using standard streaming media players, such as Real, Windows Media or QuickTime.

Streaming can dramatically increase the utility of a videoconference. For example:

- The audience you wish to include in a meeting, conference, or class may be so large that it's not feasible to include them all in an interactive multipoint videoconference, but it is necessary to have a number of the key individuals participate in videoconference.
- You may wish to save the proceedings of a meeting.
- You may wish to make the content of a videoconference-based class available on-demand for student review.
- In fact, there may be times when you simply want to record and stream an event, lecture or meeting — you don't need to videoconference — but you do have a videoconferencing appliance in the room and wouldn't it be nice if you could employ that equipment for video streaming!

A streaming participant cannot interact in a live videoconference. Streaming is a one-way delivery of media. Various methods can be employed to provide a streaming participant the means to indirectly participate, including e-mail and chat rooms.

There are four methods to stream a videoconference:

1. *All-in-one box* solutions. These devices act as an H.323 device and a streaming server. In operation, the H.323 face of the device is brought into a point-to-point or multipoint videoconference. The device removes the H.26x/G.7xx video and audio from the H.323 envelope and repackages the H.26x/G.7xx into the envelope of a streaming format such as QuickTime, Real or Windows Media. The streaming face of the device then sends the content via unicast or multicast to any user with the proper popular streaming player installed on their computer. Some devices can also record and store, providing on-demand access to the content. Examples include:

   STARBAK TorrentCE; http://www.starbak.net

   First Virtual Conference Server with Streaming Support; http://www.fvc.com

   Virtual Rooms Videoconferencing Service (VRVS); http://www.vrvs.org

2. *Combination videoconferencing terminal and streaming encoder*. In this approach, two standard devices — a videoconferencing terminal and a live streaming encoder are interconnected such that the analog A/V output of the videoconferencing terminal is fed directly to the input of a live streaming encoder. The encoder can be of the standard Real, Windows Media or QuickTime variety, or can be of a specialty type such as high-bandwidth MPEG-1 or MPEG-2. Unicast or multicast network transmission can be employed. The stream is viewed using a standard streaming player appropriate to the encoding choice. Content can be recorded and stored on the streaming server for on-demand access.
3. *Videoconferencing terminals with built-in* H.26x multicast streaming capability. Some high-end videoconferencing terminals, such as the Polycom FX and the Tandberg 880, have the ability to multicast stream a videoconference that the unit is participating in. These devices do not provide the capability to record a conference.
4. *Application service providers* (ASP). Many videoconference bridging service providers provide a streaming option. The technology that a particular ASP employs will be either of the solutions #1 (*all-in-one box*) or #2 (*combination terminal and encoder*) described above. ASP solutions are not included in the Comparison Guide presented below because the service characteristics will depend on the technology employed. Investigate which technology an individual ASP utilizes and evaluate according to the Comparison Guide.

There are some pros and cons with these methods:

- *Video and audio quality*
  - Con: Solutions #1 (*all-in-one box*), and #3 (*terminals with built-in*) repackage the H.26x/G.7xx compressed video and audio. Although H.26x compressed video can be streamed, it's not optimized for streaming. The video quality of H.26x is much more susceptible to degradation from network congestion and packet loss than the made-for-streaming formats such as Real and Windows Media. Also, the H.26x formats don't provide dynamic bandwidth adjusting capabilities such as Real SureStream and Windows Media Intelligent Streaming.
  - Pro: Solution #2 (*combination terminal and encoder*) decodes the H.26x of the videoconference to baseband video and audio, and re-encodes to a made-for-streaming format. Solution #2 provides video of a more consistent quality and can take advantage of advanced streaming features such as dynamic bandwidth.
- *Networking*
  - Con: Solution #3 (*terminals with built-in*) requires IP multicast network transmission. Multicast is sparsely supported in the commercial Internet. Support in university and commercial sector intranets is varied. Internet2 supports multicast. *Terminals with built-*

in is only feasible if your audience is on multicast-enabled networks.
  - Pro: Solutions #1 (*all-in-one box)* and #2 *(combination terminal and encoder*) support unicast and multicast network transports.
- *Operation and Management*
  - Con: Solution #2 (*combination terminal and encoder*) can be complex to manage if your institution doesn't already support live media streaming and have people managing that infrastructure. Videoconference scheduling systems don't support viewing the pair of devices (videoconference terminal and encoder) as a single logical entity, and none of the popular scheduling systems currently control a Real, Windows or QuickTime encoder. Some scheduling system vendors are considering support for this configuration.
  - Pro: Solution #3 (*terminals with built-in*) is very easy to operate and manage.
- *Features*
  - Con: Solution #3 (*terminals with built-in*) doesn't have the ability to record a videoconference.

Comparison Guide

	#1 all-in-on-box	#2 combo terminal and encoder	#3 built-in
Initial set up	complex	complex	simple
Ongoing management and operation	moderate	complex	simple
Network transmission	unicast or multicast	unicast or multicast	requires IP multicast
Quality of video and audio during network congestion	fair to poor	good	fair to poor
Can record conference for on-demand viewing	yes	yes	no
Cost to support a small number of concurrent conferences	high	moderate	low
Cost to support a large number of concurrent conferences	moderate	moderate-to-high	moderate-to-low

ViDe Members Favorite Recipes

Mary Fran's Hot & Sour Vegetable Soup

- 2 T. cooking oil
- 1 medium onion, slivered
- 3 carrots, cut into thin diagonal slices
- 3 cloves of freshly minced garlic
- 1 T. fresh minced ginger (I admit to using an equivalent amount of
- powdered ginger here)
- 4 cups defatted chicken broth
- 1 cup water
- 2 T. soy sauce
- 2 cup white mushrooms, thinly sliced
- 1 bunch fresh watercress w/stems removed (can use chopped spinach or
- swiss chard)
- 1/2 lb. fresh snow peas
- 1 cup fresh bean sprouts
- 1/4 cup rice vinegar
- 2 t. sesame oil
- Dash of chili oil to taste (white pepper works too!)

Heat 2 T. cooking oil in heavy soup pot over medium heat. Add onion & carrots; cook, stirring constantly for approx. 3 minutes. Add garlic & ginger; cook, stirring, for approx. 1 minute more. Add broth, water, soy sauce and bring to a boil; bring to a boil and boil partially covered for approx. 2 minutes. Add mushrooms & watercress; boil one minute more. Turn off heat and add snow peas & bean sprouts. Cover & let rest 2 minutes. Stir in vinegar, sesame oil, & chili oil to taste. Heat through one minute and serve immediately.

Markus' Australian Galah Soup

The Australian Galah is a medium-small parrot, grey on the back, pink breast, white crest, and are found in vast numbers all across Australia. While being highly intelligent they are also the craziest birds on the planet, making noise far beyond their size, and flying upside down whenever the mood strikes them. Hanging upside down from branches and powerlines (alive!) is a reasonably common way to see them.

>They are not renowned for good eating, but there has been centuries >of research into the topic. The following recipe is its result

- 1 large saucepan, filled with cold water
- 1 galah, plucked and cleaned
- 3 riverstones of reasonable size
- seasoning to taste
- a heat source, such as a campfire

Place galah, stones and seasoning into cold water. Place saucepan onto fire, and bring to the boil. Continue to boil, replenishing water as necessary. Boil until stones are soft and tender. Throw out the galah and water and eat the rocks.

Marc's Roasted Almond and Tangerine salad

- 6 cups Mixture of cut or torn greens, lightly packed
- 10 oz. Canned mandarin orange segments, drained
- _ cup Sliced almonds, toasted in 350degF oven for 5 to 8 minutes
- 6 Bacon slice, cooked crisp and crumbled
- Dressing:
  - 3 tbsp. White vinegar
  - _ cup Granulated sugar
  - _ tsp. Prepared mustard
  - _ tsp. Paprika
  - 1 tbsp. Cooking oil

Place first 4 ingredients in a large bowl. Stir vinegar, sugar, mustard and paprika together well in a small bowl until sugar dissolves. Drizzle cooking oil over greens, toss well, add dressing then toss again.

Tyler's H.323 Noodles with SIP Sauce

- 4 teaspoons fresh ground ginger
- 6 cloves garlic
- 8 teaspoon white vinegar
- 4 teaspoon black pepper
- 6 teaspoon sugar or honey
- 8 tablespoons tahini
- 5 tablespoons vegetable oil tabasco to taste soy sauce to taste pinch cinnamon

Puree all ingredients in a food processor. Add water to maintain a thin syrup consistency. Adjust tabasco and black pepper to control spicy/mild. Serve over thin noodles.

Mary's Favorite Marinated Vegetable Salad

Salad ingredients:

- 1/2 head of cauliflower, washed and cut into flowerets
- 1 cup cooked green beans, cut into thirds
- 1 cup cooked carrots, cut into rounds
- 1/2 cup sliced green pepper
- 3/4 cup thinly sliced celery
- 2 large green onions, thinly sliced
- 1/2 cup sliced mushrooms
- 1 can (3 oz) black or green pitted olives, drained
- 1 can (2 oz) sliced pimientos

Dressing:

- 1/3 cup oil
- 5 to 7 tablespoons lemon juice
- 2 tablespoons cider vinegar

- 1 teaspoons sugar
- 2 teaspoons dried oregano leaves, crumbled
- 2 teaspoons salt
- 1/4 teaspoon black pepper
- 1/4 teaspoon dill leaves
- 1/8 - 1/4 teaspoon cayenne pepper

Combine in a large bowl all of the salad ingredients, mixing thoroughly but gently. Prepare the dressing in a separate bowl, mixing all the ingredients until well combined. Pour the dressing over the prepared vegetables, mixing well. Cover and chill overnight, stirring once or twice. Drain dressing if too liquidy, transfer to salad bowl (on a bed of lettuce) and serve well chilled

Ed's Cheesy Bacon Grits

Note: This is not breakfast food, but a side dish for dinner. I like this with a blackened steak, or served under a spicy beef stew.

Adapted from Emeril Lagasse's "Emeril's TV Dinners"

- 4 oz precooked bacon, chopped (or 1/2 pound raw bacon chopped and cooked)
- 4.5 cups milk
- 1.5 teaspoons Kosher salt
- .5 teaspoon cayenne pepper
- 1 tablespoon butter
- 2 cups quick cooking white grits
- 8 oz White Cheddar Cheese, grated (I like Cabot Farms Hunter's Cheddar)

In a medium saucepan combine milk, salt, cayenne, and butter. Bring to a boil, stirring frequently to keep the milk from sticking. Add the grits and reduce the heat to medium. Stir for 30 seconds, then add the cheese and stir until the cheese melts. Add the bacon and stir until it is mixed in. Cook, uncovered until the grits are tender and creamy, about 4-5 minutes.

Ed's Vegetarian Curry Puffs

Note: The filling should be very spicy. Once it gets wrapped in puff pastry, the heat will become balanced. Inspired by 2 recipes on www.epicurious.com

For filling

- 3/4 cup minced onion
- 1/4 cup minced canned green chilies
- 4 teaspoons minced peeled fresh gingerroot
- 1 3/4 teaspoons curry powder (I like spicy foods, so I use Vindaloo or Madras Curry powder)
- 1 teaspoon chili powder
- 1/2 teaspoon ground cumin
- 1/4 teaspoon ground cloves
- 1/4 teaspoon cinnamon
- 1/2 teaspoon salt
- 1/4 cup vegetable oil
- 2 large boiling potatoes (1 pound), peeled and minced (about 2 cups), reserved in a bowl of cold water

- 1 small tomato, peeled, seeded, and chopped fine (1/3 cup)
- 2 tablespoons chopped fresh coriander
- 2 tablespoons plain yogurt
- Pepperidge Farm pastry sheets (2 sheets to a box. 1 box makes 18 puffs).

Make the filling:
In a large skillet cook the onion, the chilies, the gingerroot, the curry powder, the chili powder, the cumin, the cloves, the cinnamon, and the salt in the oil over moderately low heat, stirring, until the onion is softened. While the onion is cooking, in a large saucepan of boiling salted water cook the potatoes for 3 to 5 minutes. Stir in the tomato, the coriander, the yogurt, and black pepper to taste, cook the mixture, stirring, for 1 minute (do not let it boil), and let it cool. (The filling may be made 2 days in advance and kept covered and chilled.)

Make the puffs:
Thaw pastry sheets according to box. Cut pastry sheets into 9 squares. Fill each square with curry mixture. Fold the square into a triangle. You will probably have extra potato mixture left over. Put triangles on pan and brush the top of each with a beaten egg (to give it a nice yellow color). They can rest at this step for a while in the fridge before baking. If you chill them, give them a half hour on the counter to warm up before baking.

Bake at 400°F for 30 minutes.
Serve immediately.

Mary Fran's Cabbage Supreme

Sauce (really just a white sauce):

- 1/2 cup margarine, 1/2 tsp salt, 1/4 tsp pepper, 3 Tbsp flour,
- 1-1/2 cups milk

Melt margarine over low heat. Add flour, salt & pepper and stir until well blended. Add milk and bring to a boil slowly, stirring frequently to constantly until mixture thickens.

- Medium cabbage
- 1 tsp salt
- 1 cup parmesan cheese
- 1 cup cracker crumbs

Boil cabbage until tender; set aside. Make sauce. Create layers of cabbage, then sauce, then parmesan cheese, then cracker crumbs, repeating layers to fill a casserole dish (dish better higher than lower in height).

Mary Fran's Quick & Easy Cheddar Skillet Dinner

- Approx. 1 lb. ground beef (or meatless ground product, e.g., Boca crumbles, Quorn, etc.)
- 2 medium onions
- 3 - 4 medium potatoes
- 1 can condensed cream of mushroom soup
- 1 can condensed cheddar cheese coup

Slice potatoes in "planks" and cook. (Remember: leaving those skins on adds flavor and vitamins!)

Set aside. Brown ground stuff w/onions. Add potatoes and soup; also salt & pepper to taste. Heat through and serve.

Bob's Mandarin Double Chocolate Chip Cookies

Buy two bags of Nestle's Chocolate Chips. Follow the recipe on the bag, EXCEPT:

1. Put in TWO bags of chips instead of one.
2. Add 1-2 teaspoons of Orange extract (adjust to taste)
3. Do not bake as long as specified. Check the cookies carefully during cooking and take them out while they are still a little moist and above all not crisp and hard.

   Eat some while still warm. Store the rest in a container with a slice of fresh bread, so they do not get dried out.

Mary Fran's Tomato Soup Cake

(350 degrees, 30 minutes)

Frosting: Mix 3 oz. cream cheese, 1 c. confectioners' sugar, and 1 . vanilla.

Cake:

- 2 cup flour
- 1 cup sugar
- 1/2 t. salt
- 1 t. cinnamon
- 1 can condensed tomato soup
- 1 cup raisins
- 1/2 cup melted shortening
- 1 t. baking powder
- 1/2 t. nutmeg
- 1/2 t. ground cloves
- 1 t. baking soda
- 1 cup chopped nuts (optional)

Sift dry ingredients. Dissolve baking soda in tomato soup (note: may foam up) and add to the dry ingredients. Add melted shortening, raisins, and nuts. Bake in lightly greased & floured bundt pan. Cool, frost, eat!

Glossary of Terms

A - B - C - D - E - F - G - H - I - J - K - L - M - N - O - P - Q - R - S - T - U - V - W - X - Y - Z

A

antialiasing

A method for smoothing the jagged edges (stair steps) often seen in graphics or video. The method reduces the jagged edges by placing intermediate shades of color or gray around the steps.

ASF

Active Streaming Format. A Microsoft file format for digital video playback over the Internet, or on a standalone computer. Kind of a wrapper around any of a number of compression types, including MPEG. Part of Netshow, a proprietary streaming media solution from Microsoft. Biggest competitor is Real Networks. While this 'wrapper' support many standard formats, ASF files are themselves proprietary.

AVI

Audio Video Interleaved. A Microsoft format for digital audio and video playback from Windows 3.1 Somewhat cross-platform, but mostly a Windows format. Has been replaced by the ASF format, but still used by some multimedia developers.

B

banding

The presence of extraneous lines.

bandwidth

A measure of the amount of data that can fit on a network. Measured in Hertz or bits per second. For example, a regular Ethernet line has a bandwidth of 10 Mbps (10 million bits per second.) Note that bandwidth is sometimes measured as the bandwidth in one direction, and sometimes as the total in both directions. Be sure you understand which it is for the system you are looking at.

bit rate

The number of data bits per second that can flow in a communications circuit.
Some common speeds are 56K (for dial up modems), 384K (for most video conferencing) and 10 Megabits (for a low-speed Ethernet).

C

CIF

A video format that supports both NTSC and PAL signals. CIF is part of the ITU H.261 videoconferencing standard. It specifies a data rate of 30 frames per second (fps), with each frame containing 288 lines and 352 pixels per line.

CODEC

Stands for Coder/Decoder (a telecommunications term) or Compressor/Decompressor (a computer term). A codec is a piece of hardware or software that compresses and decompresses digital audio and/or video.

chrominance

The portion of a video signal that specifies what color each portion of the picture is to be. See also Luminance, S-Video and Composite Video.

Composite video

A method of carrying video information, which combines chrominance and luminance on a single wire, resulting in lower video quality than S-Video

D

decoder

A piece of hardware or software that is used to convert video or audio (typically) from the digital form used in transmission or storage into a form that can be viewed.

digital audio

Audio that has been encoded in a digital form for processing, storage or transmission.

dithering

Giving the illusion of new color and shades by combining dots in various patterns. This is a common way of gaining gray scales and is commonly used in newspapers. The effects of dithering would not be optimal in the video produced during a videoconference.

E

echo suppression

Echo suppression is a crucial portion of all video conferencing systems. If echo is not suppressed, the speaker hears his own audio coming back from the other end of the circuit, after a small time delay. Th fault always lies with the far end, although they do not perceive any problem.

F

full duplex

Sending data in both directions at the same time. Usually higher quality but requires more bandwidth. In video conferencing, full duplex will be much more natural and useable. Cheap speakerphones are half duplex, whereas more expensive ones are full duplex.

G

G.7xx

A family of ITU standards for audio compression.

gatekeeper

In the H.323 world, the gatekeeper provides several important functions. First, it controls access to the network, allowing or denying calls and controlling the bandwidth of a call. Second, it helps with address resolution, making possible email type names for end users, and converting those into the appropriate network addresses. They also handle call tracking and billing, call signaling, and the management of gateways. They also handle call tracking and billing, call signaling, and the management of gateways.

gateway

Gateways provide a link between the H.323 world and other video conferencing systems. A common example would be a gateway to a H.320 (ISDN) video conferencing system.

H

H.261

ITU standard for video coding for videoconferencing. H.261 is a discrete cosine transform (DCT) based algorithm for video in the 64kb/s to 2mb/s range. All H.323 compliant video conferencing system are required to support this codec.

H.263

ITU standard for video coding within videoconferencing. H.263 offers better compression than H.261, particularly in the low bitrate range used by modems.

H.320

ITU standard for videoconferencing over ISDN and fractional T1 lines.

H.323

ITU standard for videoconferencing over networks that do not guarantee bandwidth, such as the Internet. H.323 is the standard that this cookbook is recommending that most users in the education community should be using. For more detailed information on this and the other ITU standards see the bibliography of this document.

H.324

ITU standard for video conferencing over standard phone lines.

half duplex

A telecommunications system where data can only flow in one direction at a time. Cheaper speakerphones are a good example of this, where only one person can talk at a time.

I

IETF

Internet Engineering Task Force. This is a group that develops and publishes new standards for use on the Internet.

IGMP

Internet Group Management Protocol. This protocol is used in multicasting.

IP

The Internet Protocol. IP is the basic language of the Internet. It was developed by the government for use in internetworking multiple computer networks together.

IP Multicast

A system for sending IP transmissions out only one time, but allowing for multiple users to receive it. This would reduce the bandwidth required for audio and video broadcasting over the Internet, but it is not widely used yet.

J

jitter

The change in latency with time. This is a network problem that is very important to video quality. Significant jitter destroys video.

K

Kerberos

Kerberos is a network authentication protocol developed by MIT. It is designed to provide strong authentication for client/server applications by using secret-key cryptography.

L

latency

The length of time it takes a packet to move from source to destination; delay.

lossless compression

Refers to data compression techniques in which no data is lost. For most types of data, lossless compression techniques can reduce the space needed by only about half. Only certain types of data can tolerate lossy compression. Lossless compression technique when compressing data and programs.

lossy compression

Refers to data compression techniques in which some amount of data is lost. Lossy compression technologies attempt to eliminate redundant or unnecessary information. Most video compression technologies, such as MPEG, use a lossy technique.

luminance

The portion of a video signal that specifies how bright each portion of the picture is to be. See also Chrominance, S-Video and Composite Video.

M

MBONE

Multicast Backbone. The MBONE is a system of transmitting audio and video over a multicast network. Mostly available at universities and government facilities, the MBONE can be thought of as a testbed for technologies that will eventually be promulgated across the larger internet. The MBONE has been replaced on the vBNS and Abilene by native multicast support.

MIDI

Musical Instrument Digital Interface is a standard for connecting electronic musical instruments and computers. MIDI files can be thought of as digital sheet music, where the computer acts as the musician playing back the file. MIDI files are much smaller than digital audio files, but the quality of playback will vary from computer to computer.

MPEG

MPEG (Moving Picture Experts Group) is a series of ISO standards for digital video and audio, designed for different uses and data rates.

MPEG-1 - The initial MPEG standard, designed to encode full motion video so it could be played back off of a CD (150 kb/s). The bit rate of a standard MPEG1 is 1.5Mbps. MPEG-1 has a frame size of 352x240 pixels, which gives a picture quality slightly better than VHS videotape. MPEG-1 included three audio standards, most video systems use MPEG-1 layer 1 or layer 2 audio. MPEG-1 layer 3 audio (commonly known as MP3), is being used widely for audio on the Internet.

MPEG-2 was a follow-on standard supporting higher data rates, and thus higher quality. MPEG-2 is the standard used in DVD video players, most digital satellite systems in North America, and in the new North American Digital TV system.

MPEG-3 was abandoned, as its planned functionality was included in MPEG-2.

MPEG-4 is a draft standard that will be better suited for use on the Internet. MPEG4 delivers video at comparable quality to MPEG1 at a much lower bit rate. MPEG-4 also supports a wide variety of elements that can be transmitted separately and combined to form the video frame, such as a talking head in one stream and the background in another. That is, MPEG4 allows manipulation of objects within the video stream (addition, subtraction, object manipulation, etc.). If you don't like where a chair is in the video, you can move it (providing the chair has been coded as a moveable object, of course). Approval is expected in the first half of 1999.

MPEG-7 is a developing standard for the description of multimedia objects. Not a video encoding format, it is a way to describe elements in a multimedia stream so that they can be accessed via database. For example, it would be useful to be able to search a multimedia database for instances of 'red wagons.'

Multipoint Conferencing Server (MCS) (also MCU)

A hardware or software H.323 device that allows multiple video conferencing (or audio or data) users to connect together. Without an MCS typically only point-to-point conferences can take place. Commonly supports voice activated switching, where whoever is talking is broadcast to all users, but new systems support "Hollywood squares", where multiple windows show each participant. ITU-T standard H.231 describes the standard way of doing this. Many current systems only support H.320 (ISDN) but many vendors are working to upgrade their products to support H.323 (LAN, Internet) as well. In the H.320 space, this functionality is referred to as a multipoint control unit (MCU). Sometimes these terms are used interchangeably, although they refer to somewhat different implementations.

P

packet

A unit of information sent across a (packet-switched) network. A packet generally contains the destination address as well as the data to be sent.

Q

QCIF

A standard related to CIF, QCIF (Quarter CIF), transfers one fourth the amount of data and is suitable for videoconferencing systems on slower connections or telephone lines.

QuickTime

A file-format and architecture developed by Apple for use with digital audio and video. Available on most computing platforms. A future version (Quicktime3) will support streaming.

R

RealAudio

A proprietary system for streaming audio (and now video) over the internet. Before Real Audio, users had to download an entire audio file before they could listen to it. Also supports real-time broadcast of audio and video programs. Many radio stations now broadcast on the internet using Real Audio.

real time

A transmission that occurs right away, without any perceptible delay. Very important in video conferencing, as much delay will make the system very unusable.

S

streaming media

Sending video or audio over a network as needed, such as Real Audio/Video or Microsoft NetShow, instead of forcing the user to download the entire file before viewing it. Typically a few seconds of data is sent ahead and buffered in case of network transmission delays. (Although some data is buffered to the hard drive, it is written to temporary storage and is gone once viewing is complete.)

S-Video

A method of carrying video information on a cable that separates luminance and chrominance on separate wires, thereby providing higher video quality than composite video. See also Chrominance, Luminance and Composite Video.

T

T.120

T.120 is an ITU-T standard (International Telecommunications Union) for document conferencing. Document conferencing allows two or more people to concurrently view and edit a document across a network.

T.120 is the commonly used name to refer to a family of distinct standards. Many video conferencing companies were developing their own implementations of this until Microsoft released its free NetMeeting software. Now, many companies are using NetMeeting, while perhaps enhancing it in some way.

Teleconferencing

Two or more people who are geographically distant having a meeting of some sort across a telecommunications link. Includes audio conferencing, video conferencing, and or data conferencing.

Terminal End Station

A terminal end station is the client endpoint that provides real-time, two-way communications. This is often shortened to just terminal.

Transcoder

A device that does transcoding. See below.

Transcoding

Converting a data stream from one format to another, such as MPEG 1 to H.263, or an H.320 videoconferencing session to H.323.

Truespeech

Truespeech is a codec used for low bandwidth encoding of speech (not music). It was created by the

DSP Group. It is available on Microsoft Windows 98 among other systems.

U

unicast

Sending each user their own copy of a video (or other data) stream. As opposed to Multicast, where one copy is sent and whoever wants it listens to that copy. It is the most commonly used method for video conferencing and video on demand today. Multicast, which is much more efficient, is slowly gaining ground, but requires Internet Service Providers to support it.

V

ViDe

Video Development Group. Currently consists of the University of Alabama at Birmingham, Indiana University, Georgia Institute of Technology, the University of North Carolina, Chapel Hill, and the University of Tennessee, Knoxville, Australian National University, Southeastern University Research Association, Ohio State University, CANARIE, and SURFNet.

video-on-demand

Being able to view any of a number of videos when you want to. Used on the internet and at hotels, cable systems, etc.

video server

A computer server that has been designed to store large amounts of video and stream it to users as required. Usually a video server has large amounts of high-speed disks and a large amount of network bandwidth to allow for many users to simultaneously view videos.

voice activated switching

Automatically switching the video feed to whoever is speaking in a multipoint videoconference. Usually a function of the MCU (multipoint conferencing unit)

Appendices

H.323 and Related Specifications

The information listed here was not written by ViDe.It is not contained on the ViDe Cookbook server. These materials are listed solely for the purpose of directing you to more detailed standards information should you be interested in such topics.

Several excellent primers which describe the standards are:

- Videoconferencing over IP -
  A Primer on the H.323 Series Standard
  Trillium H.323 Tutorial
  Trillium H.323 Tutorial Self Test
  Packetizer's H.323 Version 3 Overview
- Data Collaboration -
  A Primer on the T.120 Series Standard
- Markup Languages -
  W3C Synchronized Multimedia Integration Language (SMIL) 1.0, Specification

References

- *Open Source http://www.opensource.org/*
- *OpenH323 http://www.openh323.org/*
- *GnomeMeeting http://www.gnomemeeting.org/, http://www.ucl.ac.be/*
- *Computer Science Engineering, Universite Catholique de Louvain Department of Computing Science and Engineering*
- *Transversal of IP Voice and Video Data through Firewalls and NATs http://www.radvision.com/papers/C1_Traversal_of_IP_Voice_and_Video.pdf*
- *Deploying H.323 Applications in Cisco Networks http://www.cisco.com/warp/public/cc/pd/iosw/ioft/mmcm/tech/h323_wp.htm*
- *RFC 2663 - IP Network Address Translator (NAT) Terminology and Considerations http://www.rfc-editor.org/*
- *RFC 3103 - Recommended Internet Service Provider Security Services and Procedures http://www.rfc-editor.org/*
- *Trowt-Bayard, Toby, and Jim R. Wilcox, Videoconferencing The Whole Picture, Flatiron Publishing, Inc., 2nd Edition, March 1997.*
- *AXIS Design Group, http://www.axisdg.com, accessed December 1998, general images and information provided under "Specialty", "Space Solutions", with specific study of: http://axisdg.com/Specialty/Space_Solutions/Experience/People/President/_Clients _Projects/ Enlarged_Briefing_1/enlarged_briefing_1.htm*
- *Bellwether Design, http://www.bellweather-design.com, accessed December 1998, with specific study of : http://www.bellwether-design.com/designtech.htm*
- *EPA Audio Visual Inc., http://www.epaaudio.com, accessed December 1998, with specific study of: http://www.epaaudio.com/design.html*
- *Accuwood Inc., http://www.accuwood.com, accessed December 1998.*
- *Lighting a Videoconferencing Environment (Part 1), California Community Colleges Newsletter, Vol 2, Issue 8, September, 1998, http://www.tipsnews.org/newsletter/98-09/lighting1.html.*
- *East Carolina University's Telemedicine Center, http://www.telemed.med.ecu.edu/*
- *The US Veterans' Administration, http://www.va.gov/telemed/about.htm*

- *Georgia Tech's Biomedical Interactive Technology Center, www.bitc.gatech.edu*
- *National Library of Medicine, www.nlm.nih.gov, (http://collab.nlm.nih.gov/tutorialspublicationsandmaterials/telesymposiumcd/Contents.html)*
- *CyberCare Technologies, www.cybercare.com*
- *American TeleCare, www.americantelecare.com*
- *The Health Information Portability and Accountability Act of 1996, http://www.hcfa.gov/hipaa/hipaahm.htm*
- *UKERNA (United Kingdom Research and Education Networking Association), http://www.ja.net/development/video/vip/reports/south3.pdf*
- *The Office for the Advancement of Telehealth in the US Department of Health & Human Services' Health Resources and Services Administration, http://telehealth.hrsa.gov/*
- *The Office for the Advancement of Telehealth, http://telehealth.hrsa.gov/pubs/tech/technome.htm*
- *Telemedicine Information Exchange, . http://tie.telemed.org/*
- *American Telemedicine Association, . http://www.atmeda.org/*
- *Ed1VidConf list, ED1VIDCONF@A05VM.RVR.IL.AMERITECH.COM*
- *World Bank's Global Development Learning Network, http://www.gdln.org/*
- *Virtual Room Videoconferencing System, http://www.vrvs.org*
- *Web-based discussion forum software, http://www.strom.com/places/wc.html*
- *SIP IETF Working Group, http://www.ietf.org/html.charters/sip-charter.html*
- *SIP Resource site, Computer Science Department at Columbia University http://www.cs.columbia.edu/sip/*
- *Internet Assigned Number Authority (IANA) Protocol Numbers and Assignment Services*
- *IANA port number*
- *RFC 1700*
- *Shrimp, http://www.ja.net/development/video/shrimp/*
- *University of Oregon Video Lab and the Internet2(tm) Networks Multicast Trial*
- *Setting up MBone Tools for Windows95/NT, Macintosh and Unix.*

Keeping Up With the Changes

Products Information

- *PolyCom*
- *VCON Telecommunications Ltd.*
- *VTEL Online*
- *RADVISION H.323 Homepage*
- *Accord Homepage*
- *VideoServer Homepage*
- *Global Videoconference Network*
- *Cutting Edge Technologies*
- *Tandberg*
- *StarValley Systems*
- *Pegasus Imaging Corp*
- *Access Grid*
- *Teleportec*
- *Reality Interface*
- *Aravox*
- *Cisco*
- *Check Point Software*
- *PhonePatch*
- *Open H.323 Proxy*

- *Lucent*
- *University of Oregon IP/TV Multicast Viewer*
- *UC Berkeley MBONE Tools*
- *STARBAK TorrentCE*
- *First Virtual Conference Server with Streaming Support*
- *Virtual Rooms Videoconferencing Service (VRVS)*

Relevant Listservs

MegaCon - megacon@lists.acs.ohio-state.edu This list was started to support participation in the Megaconference (http://www.mega-net.net/megaconference) but has grown to encompass lots of H.323 technology as well as application discussion. Subscribe via email request to Bob Dixon, dixon.8@osu.edu

VIDENET-L - vidnet-l@listserv.uga.edu This list includes video practioners in the commercial world as well as some R&E. To subscribe, send a message to listserv@uga.edu and include the following in the body of the message: subscribe vidnet-l your_first_name your_last_name

VIDEOIP - videoip@nysernet.org This started as a list for the NYSERNet Video-IP project but the list went public a few years ago and others have subscribed since then. To subscribe, send a message to majordomo@nysernet.org and include the following in the body of the message: subscribe videoip your_first_name your_last_name

ViDeNet lists - There are various ViDeNet lists, related to H.323 as used across and within ViDeNet. See more on the ViDeNet project at http://www.unc.edu/cavner/videnet. Lists available are shown on under the Contacts menu. Contact Tyler Johnson (see the Join ViDeNet section) for additional information on subscribing.

For Further Reference

- *Center for Advanced Video Network Engineering and Research*
- *Large Scale Video Network Prototype*
- *Videoconferencing Guide*
- *Videoconferencing Categories and Terms*
- *NetMeeting Overview and Download Site*
- *International Multimedia Teleconferencing Consortium*
- *International Telecommunications Union (ITU)*
- *The Internet Engineering Task Force (IETF)*
- *PictureTel Standards Page*
- *Welcome to the OpenH323 Project*
- *Multimedia Streaming, University of Wisconsin - Madison*
- *TERENA DEVICE PROJECT, Desktop Videoconferencing - Current Products and their Interoperability*
- *W3C Synchronized Multimedia*