A Technical Implementation Guide for Multi-site Videoconferencing

Liu, Sandy; Molyneaux, Heather; Matthews, Bradley

Publisher's version / Version de l'éditeur:

Proceedings of the IEEE International Symposium on Technology and Society (ISTAS 2008), June 26-28, 2008, Fredericton, New Brunswick, Canada, 2008

National Research Council Canada
Institute for Information Technology

Conseil national de recherches Canada
Institut de technologie de l'information

NRC · CNRC

A Technical Implementation Guide for Multi-site Videoconferencing *

Liu, S., Molyneaux, H., Matthews, B.
June 2008

* published in the Proceedings of the 2008 International Symposium on Technology and Society (ISTAS 2008). Fredericton, New Brunswick, Canada. June 26-28, 2008. NRC 50379.

Canada

A Technical Implementation Guide for Multi-site Videoconferencing

Sandy Liu, Heather Molyneaux, Brad Matthews
National Research Council Canada
{firstname.lastname}@nrc.gc.ca

Abstract

With the increased cost, time and potential risk or hassle involved in traveling, videoconferencing has become a popular alternative for meeting people from geographically distributed locations. Videoconferencing tools have also become widely available, and videoconferencing technologies have improved substantially over time. This paper provides a basic technical implementation guide for those involved in setting up videoconferencing in an organization. The technical infrastructure, the interaction between users and the technology, group dynamics, and the organization of the content of the videoconference, are the four key factors towards a participatory videoconferencing session [1]. This paper examines one of the four variables, the technical infrastructure, which is a necessary condition for a successful session. This paper aims to provide a practical guide for those who are given the task to obtain, set up or improve a multi-site videoconference system. It outlines a list of required technical components and potential issues that need to be addressed when setting up a multi-site videoconference. This paper starts with a checklist of requirements, followed by an introduction of different types of videoconferencing systems, the basic technical components, and some related issues in selecting and implementing a multi-site videoconferencing system.

1. Introduction

Videoconferencing can bring people together from distributed locations for synchronous communication. Multi-site videoconferencing is used by various groups of people for a variety of purposes, from business meetings to fostering community development. Videoconferencing is a productivity enhancement tool if it is implemented and used properly. However, "participatory" multi-site videoconferencing does not automatically occur. Four variables can either enable or constrain participatory videoconferencing: group dynamics, the organization of the content of the videoconference, the interaction between users and the technology, and the technical infrastructure [1]. This paper focuses on the technical infrastructure, which refers to more than the videoconference system itself; it also includes videoconference system itself; it also includes the hardware and software, the potential for multi-site bridging, and the availability of a network with the required bandwidth.

The type of videoconference and the attributes of the technical infrastructure affect the quality of the videoconference, the number of people who can communicate using the technology, and the capacity of the system to incorporate new features and elements. The technical infrastructure is crucial to multi-site videoconferencing. Even though the equipment itself does not ensure successful communications, clear audio and visual signals can enable greater participation by increasing the quality of the auditory and visual cues.

The options available today for multi-site videoconferencing can be overwhelming. The goal of this paper is to provide a practical and comprehensive guide for selecting and implementing the right videoconferencing system to fit the organizational needs.

2. Requirements and Constraints Elicitation

There is no single videoconferencing solution that can fit all usage scenarios. To find the right system, it is important to identify the basic needs and constraints within the organization; these factors can directly or indirectly affect the choice of the systems.

To identify these needs, one can start by examining some common usage scenarios. From these scenarios the following information can be extracted:

- **Number of sites involved**. This determines whether or not a bridge (see section 4.3 for details) is required; whether or not to choose a solution with multicast capability, consequently, the network may require end-to-end multicast support.
- **Number of participants at each participating site.** There are different solutions targeting for personal, small group, and large group videoconferencing. The number of participants at each site also determines the room setup (in order to provide sufficient space), the camera setup (to be able to "see" everyone), and the microphone setup (so everyone can be heard clearly). Some features such as voice activated capture may be a consideration.

- Current **network connection capability** (i.e. bandwidth of LAN & WAN connection), especially the capability that can be used by videoconferencing, and the potential to acquire additional bandwidth. Obviously, higher definition video requires higher bandwidth to transmit the signals.
- **The meeting venue**. Some may prefer a stable setup in a specific conference room while others may prefer a mobile solution where the meeting venue can be changed from time to time.
- **Purpose of the videoconferencing** and if there is any **specific requirements**. Videoconference can be used for many purposes, such as training, lecturing, observation, business meeting, conference forum, presentation, and so on. The purpose of the videoconferencing has direct influence on the acceptable level of delay in signal transmission and the image resolution. For instance, if the videoconference is used for remote medical diagnosis, the resolution requirement will be much higher than a normal business meeting.

- **How often the videoconference devices will be used**. Some may require a 24/7 connection, while others could be once a month or once a quarter. The frequency of usage can have implication on the network Quality of Service (QoS) and whether or not to buy certain equipment vs. to obtain the service from a third party.

- **Budget** places a critical constraint on what can be obtained in reality. It is a good practice to come up with a list of must-have features, and a list of good-to-have features. Then find a system that can fit into the budget while also providing, at the minimum, the must-have features.

In the next section, we will discuss the types of videoconferencing systems available today.

3. Types of Videoconferencing Systems

Many different types of videoconferencing systems are available at different price levels and designed for various purposes. Videoconference systems used by organizations and groups include set-top and integrated or dedicated systems. Other systems can be used primarily for personal videoconferencing, such as desktop videoconferencing and 3G mobile phones.

3.1. Group videoconferencing

3.1.1. Set-top videoconference systems. Set-top videoconference systems are complete systems that sit on top of computers or television monitors. The two most popular brands of videoconferencing equipment are Tandberg and Polycom. Sometimes set-top systems are located in small located in small boardrooms but often set-top systems are placed on carts and can be moved to different rooms for different videoconferencing purposes.

Set-top systems are often used for small groups, including business and administration meetings and small group educational sessions. Set-top systems are also used in telehealth for clinical sessions. Although set-top systems can be used for one-on-one communication, they are more commonly used for group communication.

Many group videoconferencing units contain the hardware and the software necessary for videoconferencing. For example Tandberg and Polycom units provide an integrated or embedded "firmware" solution, sometimes referred to as black box design; this allows for a more turnkey system, with the technical details removed from the user experience. The software is always "firmware," non-volatile memory that is electronically programmable and erasable. However, most companies can provide maintenance services that may include the upgrade of the "firmware" when new versions are available at an additional cost.

Pre-packaged audio and video hardware include a conference room camera, with tilt, pan and zoom, a wall mount or tripod (although it is better to position the camera directly above the screen displaying the off-site locations in order to simulate eye contact), video cables, and a conferencing system that includes an audio mixer, microphone distribution box, microphones and speakerphones, XLR connectors and a controller.

3.1.2. Integrated videoconferencing. The second type of group videoconferencing system is known as integrated, dedicated or room-based systems. In integrated systems all components are incorporated into a single piece of equipment and are not meant to be moved. The wiring and processor for the system is centralized and the camera is mounted on the wall of the room. This provides more stability and allows for the greatest potential functionality. When the wiring, cameras and system control are permanently installed, they can be optimized for ease of use and performance.

Integrated systems are ideal for large group communications and commonly used for administrative meetings, distance education - including medical education and academic and medical conferences. Most integrated systems are capable of connecting to both H.320 and H.323 end points on ISDN and IP based networks.

3.2. Personal videoconferencing

Personal videoconferencing is largely dominated by different desktop systems, where videoconferencing is carried out through a desktop or laptop computer. They are popular because of their low cost and easy installation.

However, desktop videoconferencing systems do have their drawbacks for multi-site videoconferencing. Most desktop systems cannot accommodate group communication and interact with multiple sites. In addition, many videoconferencing network managers will not allow, or do not like to allow, personal videoconferencing units to be included on a multi-site videoconference with set-top units.

It is technically possible to link a site using a desktop system such as Polycom PVX into a multi-site videoconference with other sites using set-top systems like Polycom and Tandberg; however, the quality of the audio-visual signal from the sites using the desktop system will generally be inferior to the signal from the sites using the set-top units, especially if the desktop system is using inadequate bandwidth. This can create challenges for the bridge operator attempting to maintain a high-quality videoconference for all participants. Participants at the sites using the set-top units may find linking up with sites with inferior audio and video frustrating. Degraded video may be acceptable but degraded audio makes videoconferencing difficult. New bridging technology can maintain different connection speeds from different sites, thus not every site will suffer from the presence of one slow site.

Ports used by many desktop videoconference programs are often blocked by corporate firewalls, making them difficult to use in a business environment. Integrated and set-top systems are more feasible because network administrators can create specific firewall rules for each system. Firewalls can also cause problems for videoconference systems that use a large number of random ports or are not compatible with NAT (network address translation).

However, despite the potential problems of including desktop systems in a multi-site videoconference, the need to do so will continue as multi-site videoconferencing increases in popularity. Many potential participants who may want to be connected to a multi-site videoconference do not have access to set-top systems and there is no other way to have them participate.

Desktop systems are the most popular means of personal videoconferencing. MSN Messenger reported more than three billion webcam sessions up to 2006, and MSN is only one of more than 40 different systems used to conduct a video call over a personal computer [2]. Desktop systems can be used for business as well as personal communication with friends and family. Most desktop videoconference systems are used only for one-to-one meetings, although with most desktop systems single participants can videoconference with a group.

Desktop videoconferencing systems require several basic pieces of equipment and software, first of all a desktop or laptop computer, a full-duplex audio card and a video card capable of a minimum resolution of 1,024 by 768 and 768 and displaying 24 bit colors. Auxiliary hardware includes a webcam, microphone and speaker. Some webcams have built-in microphones but if they do not then head set units with headphones and a microphone or a combination of microphone and speaker units are used for videoconferencing. Various types and sizes of screens can be used with desktop computers, such as projection screens, but computer monitors are the standard type of viewing screen for desktop videoconferencing. Special webcams that clip onto laptops are available, and the latest Apple Macbooks, and some selected laptop models have integrated webcams and microphones.

Personal videoconferencing systems are typically implemented as software applications available for download on the Internet. The most popular ones include Skype, MSN or Windows Live Messenger, AIM, Yahoo Messenger, just to name a few. Most of these systems are free to download, but most of them are designed for dual audio and videoconferencing only. The videoconferencing functions are meant to augment the text capabilities, and therefore onsite and offsite images are not normally full-screen.

iChat is a program designed specifically for Mac computers for text messaging, audio conferencing and videoconferencing. The latest iChat (on Leopard) has the capacity to videoconference with up 10 sites, as long as the other sites meet system requirements [3]. iChat displays dual videoconference images full screen, while multi-site are shown simultaneously. With the capability of connecting 10 sites, the line between personal and group videoconferencing systems blurs.

In addition, there are software solutions such as Polycom PVX (for PC) and XMeeting (for Mac) that offer the capability for individuals to join a standard H.323 session that involve other participants using integrated or set-top videoconferencing systems.

Web-based video and audio conferencing solutions such as Lotus Sametime, Cisco WebEx or Adobe Connect are software-based systems that are more firewall-friendly because they run on a central web server and use a limited number of ports to communicate. Many businesses use web conferencing software for desktop videoconferencing because they are so firewall friendly and, in addition to sharing audio and video, they allow users to share presentations, applications, whiteboard sessions, or entire desktops over the internet in web-based virtual meeting rooms.

Medium to large sized organizations can also purchase more expensive specific systems that include both hardware and software; for example, the Cisco Unified Video Advantage system, which includes a Cisco webcam, Cisco IP phone and software that connect to a desktop computer. With this system group members can conduct dual-site videoconferences with colleagues.

With the growing bandwidth availability for wireless networks and the fast growing cell phone manufacturing technologies, using 3G mobile phones and video phones are quickly becoming another option for personal videoconferencing. The debut of iPhone also makes video a trendy feature on mobile phones. Nonetheless, video phones are not yet suitable for multi-site videoconferencing today, largely due to the minimum level of service required in order to videoconference over the mobile network. There are also additional issues including the network coverage, inconsistent bandwidth, the high cost associated with the service, and the limited screen size that prevent the mobile phones from becoming a mainstream tool for videoconferencing.

4. Basic Technical Components

From a technical prospective, videoconferencing can be seen as a type of synchronized communication between two or more locations over some telecommunication lines. A typical videoconferencing system consists of the audio components (e.g. microphones, speakers, mixers), the video components (e.g. cameras, displays), and the computer software and/or hardware (the software can be built-in to a piece of proprietary hardware or being installed to a PC). Enabling the simultaneous communication between two or more parties at distributed locations requires the smooth integration of all of these components.

4.1 Audio components

Audio components typically include microphones and speakers. When multiple microphones are used, a mixing board is also required. The selection and placement of microphones may not be straightforward; we will discuss the selection of microphones in the following subsection.

4.1.1. Microphones. Videoconferencing rooms should ideally have ceiling-mounted microphones which are omnidirectional. A microphone's directionality or polar pattern indicates its sensitivity to sounds arriving at different angles. Omnidirectional microphones can pick up sounds from all different directions but they also pick up the unwanted noise. In contrast, a unidirectional microphone is sensitive to sounds from only one direction. A cardioids microphone is considered to be the most common unidirectional microphone. It has a heart-shaped sensitivity pattern around the microphone.

When ceiling-mounted microphones are not used, most group videoconference units use two main types of microphones: desktop and lapel. Of the many different types of desktop microphones, the two most commonly used are octagonal or round in shape. Their placement is important. Octagonal microphones pick up sound only from the front from the front and round microphones pick up sound from all sides [4]. In multi-site group videoconferences the microphone should be positioned on the table in front of the participants. A standard from Tandberg for videoconferencing is to keep the microphone three feet away from the set-top unit and not to move it [5].

Lapel, lavaliere or hands-free microphones are either clipped to clothing or worn around the neck. Lapel microphones are used during lectures and presentations. The presenter has to remain aware of the microphone, avoid unnecessarily handling of the microphone and remember to turn the microphone off during breaks [4]. Major drawbacks of the lapel microphones are very poor spectrum reproduction, ancillary noise due to interaction with clothing during movement, and large variations in volume depending on the orientation of the wearer's head. Some form of microphone which is attached to the head and maintains a fixed position with respect to the mouth is far superior to the lapel type.

In larger rooms with large numbers of participants multiple microphones may be required. Use as few microphones as possible during a multi-site videoconference because multiple microphones increase the likelihood that other sites will hear more background noises [6]. Multi-screen views are usually voice activated, therefore for multi-site videoconferences muting the microphone while not speaking is a critical factor for the initiation of the multi-screen view.

4.1.2. Acoustic echo cancellation. Acoustic echo is the reflection of a sound signal. This sound needs to be cancelled, or it will become very distracting. Unfortunately, conference phones such as Polycom's Soundstation or installed room systems which use ceiling speakers and microphones on the table often pick up echoes.

Echo can be reduced or even avoided by rearrangement of the microphone and speaker settings. Try to point the microphones away from the sound system's speakers, so the audio from the speakers is not being picked up by the microphones. Choosing a low sensitive and narrow directivity microphone, such as a cardioids microphone, can also help to reduce echo. Moreover, using blinds or curtains to cover windows, acoustic ceiling tiles, carpeted flooring, and using acoustic boards on top of old-fashioned plaster-cement walls, are all considered to be useful in reducing echoes.

There also are software or hardware solutions called "echo cancellers" that aim to suppress echo. To work, an echo canceller needs to be able to recognize the originally transmitted signal that re-appears and remove only that signal without removing the sound from the far end. This is a challenging task; when echo cancellers don't work correctly they produce a variety of unwanted side effects. However many videoconferencing systems have built-in echo

echo cancellers; and standards are presented to ensure they work properly to eliminate echo.

4.2 Visual components

4.2.1. Camera. Most videoconference systems require at least one camera at each endpoint. The type of camera used in a videoconference system directly affects the video quality although other factors such as the network speed also play a role. The quality of the video produced by the camera is determined by the image sensor, which in turn, determines the resolution of the image. Typical resolutions are 640x480 pixels for webcams, 720x480 pixels for NTSC cameras, 720x576 pixels for PAL cameras, and1080x720 pixels for high definition cameras. The technology used for the image sensor affects the cost camera and the quality of the image. Webcams often use CMOS sensors that are inexpensive but cannot render color and brightness signals very well. Better cameras often incorporate CCD sensor technology that can capture professional quality images.

In addition to image quality, there are other features on a camera that may add up the cost. For example, the ability to pan, tilt, and zoom, wide angle lens, manual & auto-focus, manual & auto-iris, auto-tracking, remote control, RS-232 control, etc. [7].

4.2.2. Display. The display shows the incoming video from the remote sites. It influences how clearly the remote site can be seen and also how many people at the receiving site can easily see it. The screen size and the resolution are the two main selection considerations. Besides computer monitors and projectors, now large size (40" and up) high definition LCD TVs and Plasma TVs have become less expensive, and thus have become popular options for use in small to large group videoconferencing. The recommended viewing distance of a display is about 2-6 times the diagonal size of the display [7].

Generally a normal computer screen has much higher resolution (1024x768 pixels) compared to the typical image sensor resolution. Enlarging the video window does not give richer image information if the incoming image is not at the same resolution as the display. The "extra" pixels will be automatically filled in by the system to look similar to their neighbors; consequently the video quality will be degraded [7].

4.3 Network and bandwidth

The network provides the connection or link between the end points. The bandwidth refers to the amount of electronic information which can be transmitted each second between two end points. Larger bandwidth allows for higher quality videoconferencing.

Videoconferences can be held on a variety of bandwidths. bandwidths. There are three general quality levels for multi-point videoconferencing: business quality, enhanced definition and high definition. Business quality bandwidth is from 128Kbps to 256Kbps and is generally the accepted quality for meetings and collaborations. Enhanced definition, between 384Kbps and 768Kbps, is a good visual and audio quality for meetings and customer interactions, especially when showing visual material. The high definition resolution requires 1Mbps or more, and shows a high visual level of detail needed for some types of clinical consults and product developers [10].

For interactive videoconferencing, it is important to have symmetrical upload and download speed, so that all participants can communicate equally.

Videoconferencing was quite expensive in the past as the cost of using the telecommunication lines was high. Many were run on T1, ATM, or ISDN lines. Now TCP/IP networks are prevalent, and the cost of videoconferencing has been reduced substantially where these networks are available. The following two subsections provide a close look at the ISDN connection and the TCP/IP connection.

4.3.1. ISDN. ISDN lines are usually offered by telephone carriers and transmit data over pre-existing telephone lines. There are some advantages to using ISDN lines. They are not connected to the internet and therefore are more secure than TCP/IP, and they use digital rather than analog transmission so no analog conversion is needed, resulting in a clear transmission. ISDN lines also have guaranteed bandwidth and can thus ensure a continuous QoS. This is important in medical uses of videoconferencing where high image quality and secure continuous transmission is vital. The bandwidth can be increased by adding more ISDN lines.

There are several problems with using ISDN lines for multi-site videoconferencing. ISDN lines cannot communicate with IP lines without a gateway, a device that can connect different networking technologies together (H.320 -ISDN and H.323 –IP). ISDN lines are also very expensive; while the bandwidth can be increased by adding new lines, each additional line is an extra cost. Each time ISDN lines are used there is a cost involved for the site making the call, which can be considerable for a high-quality ISDN multi-site videoconference lasting a few hours. In addition, ISDN is difficult to configure and manage. It is subject to service interruptions – as ISDN users must dial in to a provider that offers ISDN internet service – which means that the call could be disconnected. Also, video calls on ISDN cannot be put on hold nor be forwarded.

4.3.2. TCP/IP. An alternate to ISDN and more commonly used type of network is the TCP/IP (Transmission Control Protocol/Internal Protocol) network. Unlike ISDN, TCP/IP can be used over a heterogeneous network – it can connect

connect computers using different types of networks. TCP/IP is generally cheaper to use than ISDN because once the network is in place it can be used for no additional cost.

There are several types of high-speed IP connections available. These include DSL, cable modem, satellite, wireless broadband and fibre optic. Most businesses today are connected to the internet through wired Ethernet. The service can be delivered via a fiber optic connection or through the telephone network. The connection speeds generally range around 10, 100, or 1000Mbps. The wired Ethernet is fairly reliable and has low latency. Organizations and communities can also obtain dedicated connections via, for instance, T1 or ATM services. T1 is a dedicated connection supporting data rates of 1.544Mbps, while T3 lines are operating at 45Mbps. ATM (Asynchronous Transfer Mode) is also a dedicated connection switching service that organizes digital data into equal sized 53byte cells over a medium using digital signal technology (DSL). ATM speeds are ranging from 155Mbps to 2488Mbps.

There are potential problems with IP lines related to bandwidth sufficiency, quality of the network, consistency of the connections, reliability and, of particular interest in telehealth, the problem of keeping the transmissions secure and confidential [9]. Firewall configuration is always an additional problem when using a TCP/IP network for videoconferencing. Since H.323 applications use quasi-random dynamically allocated sockets for audio, video, and data channels, the firewall must be configured to allow the traffic through in real-time.

4.4 Multi-site videoconferencing bridge

Multi-site videoconferencing can be established via a Multipoint Control Unit (MCU), a concept introduced by the H.323 standard. The MCU thus becomes the congregated point of contact for all the participating units. An MCU unit is also commonly referred to as a "bridge."

Videoconference systems can be linked in four ways: the first and most traditional way is via a stand-alone bridge device; the second is with MCUs embedded into integrated videoconference units; the third way is through a multipoint service provider (MSP); finally, some new desktop videoconference software has the capacity to link a limited number of sites. Thus an MCU could be implemented purely as a software solution or as a combination of software and hardware.

A good selection of MCU systems can bridge IP and ISDN-based videoconference sessions. Dedicated stand-alone MCUs can act as a gateway between different communications media, accommodate hundreds of concurrent connections, and utilize varied protocols depending on the "option" package purchased. For example, standards such as SIP allow these units to integrate into an integrate into an organization's VoIP system. A selection of set-top units will also provide limited bridge functionality to multiple endpoints.

A stand-alone MCU bridge is a central point of contact for all audio and video endpoints. The bridge supports many simultaneous sessions so it is used to interconnect videoconference units and other end points in one or more multi-party conferences. The integrated MCU in a set-top videoconferencing unit can be used to increase the total number of supported sessions by joining everyone connected to that unit.

There are different methods of selecting which audio/video stream(s) to return back to the sites. A common method used is called audio switching – the site has on-screen presence when someone there is talking. This method could be problematic when multiple sites are talking simultaneously if only one site can be captured by the MCU. Another alternative is called chairperson control, which allows the chair of the meeting to control which site is active. Some newer MCUs use combined methods allowing the chair of the meeting to have the continuous on-screen presence while the other sites are selected to show on-screen by using the audio-switching methods.

Purchasing or renting MCU services can be too costly [8] and therefore it may be more cost-effective for an organization or community to own its own MCU rather than pay a fee for the bridge service. However managing a stand-alone MCU is a major commitment that requires dedicated expert staff for ongoing maintenance and day-to-day operation.

Typically the cost of an MCU is determined by the following aspects:

• The number of concurrent calls it can handle (# of ports for video and # of ports for transcoded audio) and the bandwidth range for each call;

• The number of supported protocol standards (e.g. H.323, H.320, SIP, H.243, H.264, H.281, H.235, H.239, SDP…); an MCU can act like a gateway bringing systems using different protocols together;

• The number of audio transcoding codecs it supports (e.g. G.711A/μ Law, G.722/G.722.1, G.723.1, G.728, G.728, g.729 A…);

• Compatibility with other proprietary solutions (e.g. Radvision's Gatekeeper, Cisco's H.320-H.323 Gateway)

• Other functions, such as interoperability with VoIP systems, using an on-board audio transcoder to accelerate the process, scheduled or real-time conference control via a user controlled web interface support for simultaneous display of the presenter (continuous presence layout) and presentation material, voice-activated video switching, ad hoc multipoint calls, and encryption schemes used for the data/audio/video streams.

Multi-site videoconferencing is not something that only IP systems can do. Typically ISDN systems also have built in multi-site capabilities providing the ability to have four sites in one call without the need for an external bridge.

Multi-site videoconferences can still be possible without the presence of a MCU. One alternative is to enable IP multicasting at each participating site. In this case, each site sends a stream of packets to all the participating peers. Therefore, the bandwidth requirement for IP multicasting is lower than using a MCU, where copies of the same packets are sent to each of the receivers. Nonetheless, end-to-end multicast support is required to make this possible but not every network supports IP multicasting, and even when the network has the capability, the network administrator may choose to disable it because of bandwidth or security concerns. The challenges include a lack of interoperability between vendor-specific implementations of multicast protocols; network administrators' lack of exposure to and experience with multicast systems; and downstream service providers not accommodating multicast for a variety of reasons, including but not limited to a desire to charge a premium for the extended functionality.

5. Related Issues

Typically there is a critical step involved in videoconferencing. It is the conversion of voice and image data stream from their native analog format to digital format and from digital to analog at the other end of the communication line. This is the function of the Codec, which is short for "Compressor-Decompressor", "Coder-Decoder", or "Compression/Decompression algorithm". Codecs are often used in videoconferencing and streaming media applications. For example, a video camera's analogue-to-digital (ADC) converts its analogue signals into digital signals, which are then passed through a video compressor for digital transmission or storage. A receiving device then runs the signal through a video decompressor, then a digital-to-analogue converter (DAC) for analog display [11].

In the encoding and decoding processes, the continuous signals are converted into ones and zeros, packaged in discrete packets to go through the digital communication lines and then unpacked and decoded at the other end. Depending on the compression ratio, various degrees of loss of clarity and delay are introduced in these encoding and decoding processes. Generally, higher compression ratios result in smaller data size, which in turn requires less bandwidth for transmission but has a higher degree of loss of audio and video quality. Compression also introduces delay.

5.1 Delay or latency

Real-time communications such as videoconferencing are sensitive to delay and variation in packet arrival times. Delay or latency is inevitable as it takes time for packets to travel across the network. Compression also introduces another layer of delay. The delay budget for a reasonable two-way conversation is about 150 milliseconds round-trip (RT); more than 150RT can be annoying to users.

Another disturbing factor is jitter, which occurs when packets arrive early, late, or out of sequence. Excessive jitter causes the users to experience quality degradation during a call. Some videoconferencing systems employ uncompressed audio and video in order to present users with a jitter-free low-latency high quality videoconference experience; this, however, requires both high bandwidth and priority QoS implemented in the network to accommodate these uncompressed data streams.

5.2 Security concerns and QoS

For TCP/IP videoconferencing for telehealth, for business, and for linking remote communities with limited bandwidth, good practice includes a consistent implementation of QoS and some level of network performance management to ensure optimal performance.

Single or multi-party videoconference sessions in a business or telehealth context are usually private in nature and may contain sensitive information. Videoconferencing security policies must address issues of reliability, integrity and confidentiality. Message reliability, when connecting remote endpoints across a TCP/IP wide area network, can be particularly problematic if session information must compete for bandwidth with other network traffic.

In this scenario, QoS techniques should be implemented within the LAN to properly classify traffic, and Gatekeepers must be deployed for bandwidth and priority queue congestion control. In IP-based networks, videoconference sessions are susceptible to the same integrity issues as other traffic. Standard IP network attacks like spoofing, eavesdropping and denial of service must be considered when designing security to protect data/audio/video streams. A firewall is one example of a typical technology deployed by organizations to protect inside systems from unauthorized access. However, perimeter security systems like firewalls do not address security within the local area network (LAN). Since many videoconference sessions occur within a LAN or between trusted wide area network zones via leased telecommunication lines, additional measures must be taken to ensure the integrity and confidentiality of videoconferences.

For example, AAA (authentication, authorization and accounting) policies should be implemented to prevent unauthorized participation in videoconferences and to track

track session history on a per-user basis. This can be particularly important when using multicast technologies as the potential audience can be large with little way to confirm group member participation. Finally, confidentiality security policies must address live session traffic and saved sessions stored on digital media for future viewing or redistribution.

Encryption is a common technology deployed to secure data/audio/video streams or digital files. Encryption standards like DES, 3DES and AES are commonly used in virtual private network systems to establish secure IPSec tunnels across public TCP/IP networks. Encryption may also be available within the videoconference MCU and can be used to secure session information between endpoint participants.

A videoconference on a TCP/IP network can run into bandwidth shortages which adversely affect the performance of not only the videoconference but also other services like email and file transfers sharing the same bandwidth on the network. Gatekeeper software, on a switch, router or server, is used to provide quality of service (QoS) - to regulate the videoconference so that it does not use up all the bandwidth and block the transmission of other data.

QoS is complex to configure and must be deployed within the LAN at the layer 2 and 3 before audio/video traffic traverses call zones. Bandwidth control via a Gatekeeper is only effective if the traffic has been classified so that strategies like class-based weighted fair queuing (CBWFQ) and low latency queuing (LLQ) can work. Additional QoS strategies are often deployed to prioritize voice traffic by keeping video out of the priority queue.

5.3 User interface, system setup, and its maintenance cost

Each videoconferencing system comes with a user interface of some sort, which may include a console window to initiate/terminate a conference session, an address book, and so on. Whether or not the interface is user-friendly affects the utilization of the system.

Other factors to consider when selecting a videoconferencing system include how easy the system can be setup or installed and how difficult and how much to maintain its functionality over time.

6. Conclusion

It is not an easy task to select and implement an effective technical infrastructure to support multi-site videoconferencing. A careful requirement analysis of organizational needs and a good survey of available products and implementation options are required. Having a good understanding of the basic technical components and their related issues for videoconferencing provides a starting point starting point for success, which is what this paper is aiming for. It introduced different types of videoconferencing systems, the underlying technical components, and also discussed some potential issues or problems corresponding to different types of system implementations.

7. Acknowledgements

Thanks to Susan O'Donnell, Bobby Ho, and David Scobie for providing their expertise and valuable input.

8. References

[1] Molyneaux, H., O'Donnell, S., Liu, S., Hagerman, V., Gibson, K., Matthews, B. et al. (2007). Good Practice Guidelines for Participatory Multi-Site Videoconferencing. Fredericton: National Research Council. ERB-1151. NRC 49869.

[2] Gough, M. (2006). Videoconferencing over IP: Configure, Secure and Troubleshoot. Rockland MA: Syngress.

[3] Wikipedia. (2007). "*iChat.*" Retrieved December 9th, 2007, from http://en.wikipedia.org/wiki/IChat.

[4] Sharer, S. (2004). "Skills for Effective Use of Videoconference." Savanna, Georgia: Communication Design Group.

[5] Tandberg. (2008). "Tips" Retrieved March 25, 2008, from http://www.tandberg.com/media/tandberg_tips/FaultFinding/Simple%20faultfinding/player.html

[6] PictureTel. (1998). "TIPS on Videoconferencing." Retrieved August 7, 2007, from www.tipsnews.org/newsletter/98-09/TIPS_on_video.html.

[7] Trawner, M., & Yafchak, M. (2007). "Videoconferencing Cookbook. version 4.1." Retrieved May 31, 2007, from http://www.vide.net/cookbook/cookbook.en/

[8] River Valley Health. (2006). Benefits Evaluation and Future Planning Strategies for the River Valley Health Tandberg Multi-Conference Unit Fredericton NB: Park Consulting Group: for River Valley Health.

[9] Masters, S., ed. Telehealth Handbook: Centre for Telehealth @ Mheccu.

[10] Tandberg. (2007). "See: green." Retrieved August 8, 2007, from http://www.seegreennow.com/

[11] Wikipedia. (2007). "Codec." Retrieved October 1, 2007, from http://en.wikipedia.org/w/index.php?title=Codec&oldid=160744371