September 1999
ECG517/0899

Prepared by:
Internet and E-Commerce
Solutions Business Unit
Enterprise Solutions Division
Compaq Computer Corporation

Introduction to ISP Hosted Multipoint Conferencing

Abstract:

This document provides ISPs and other service providers with a general understanding of services for multipoint conferencing -- communicating in real time with audio, data and video among three or more participants at different locations.

Multipoint conferencing is an important new communications tool that businesses can employ to improve their own productivity, to build better relationships with customers, partners and suppliers and to lower their own costs while decreasing cycle times. The objective of this document is to assist service providers in evaluating this new business infrastructure hosting opportunity.

This document defines multipoint conferencing and contrasts it to other familiar forms of multimedia communication. It describes market needs and opportunities for new business communication services based on multipoint conferencing. It provides four models for services and an overview of the H.323 and T.120 standards that can be used to implement them. Finally it reviews important business and technical issues to be considered in the creation of such services.

Contents

Notice

Introduction to ISP Hosted Multipoint Conferencing
Prepared by Internet and E-Commerce Solutions Business Unit

Enterprise Solutions Division

First Edition (September 1999)
Document Number ECG517/0899

Table of Contents

1 Multipoint Conferencing and Multimedia Communication

Multipoint conferencing is an emerging opportunity for service providers who wish to offer a full line of communications services to businesses or bring multimedia chat to online communities of interest.

Businesses are looking for new communications services that can help improve their operational capabilities and efficiencies. This includes support for internal operations and the growing need for constant interaction with customers, partners and suppliers, anytime and anywhere.

This section introduces multipoint conferencing and compares it to other forms of multimedia communication in use on IP networks.

1.1 What is Multipoint Conferencing?

Multipoint conferencing connects people at three or more locations via a network and allows them to interact as a group in real-time. A combination of audio[1], data, and video media is used to communicate among the participants. Examples follow.

Audio conferencing allows people in a conference to hear each other speak. A telephone conference between a team of five area sales managers and their supervisor is an example of multipoint conferencing using audio over the Public Switched Telecommunications Network[2] (PSTN).

Video conferencing allows people in a conference to see each other as well as to hear each other. Video enriches the human dimension of remote communication and can convey important non-verbal messages. Video can also carry the image of what is being discussed, for example, a document, a chart, or a machine undergoing repair. Video is specific to the capture, transport and presentation of moving images.

Data conferencing is a method of sharing information on one person's computer (or data communications device) with other people in a conference. Data conferencing is often identified with the features of Microsoft NetMeeting[3], the most widely distributed data conferencing product. (NetMeeting also includes audio and video capabilities.) Data conferencing with NetMeeting includes these features:

- Application Sharing – A person can elect to share an application running on their desktop. When sharing is enabled, the application's display window is projected onto the desktops of the other participants. This is called *view mode*. In view mode the application can be operated only by the person who has shared it. *Collaboration* is a second, optional *mode* that allows the other participants, in turn, to operate the shared application. For example, with Collaboration a distributed workgroup can collectively edit a PowerPoint presentation.
- Chat – A shared application that sends text message to everyone or to a chosen individual.
- Whiteboard – A shared application that can be typed on, drawn on or pasted onto.
- File Transfer – Transmits a file from one person to everyone else.

[1] In this guide, audio signifies voice unless indicated otherwise. Voice and music have different spectral characteristics. To gain optimal fidelity and efficiency, different signal processing techniques are often used for each.
[2] The Public Switched Telecommunications Network (PSTN) is a global network of networks that was designed originally for voice communication. It is governed by a set of technical standards and often regulated by governments. The network is primarily circuit switched and supports full interconnection across networks. The network is predominantly digital on the backbone and analog at the local loop.
[3] http://www.microsoft.com/netmeeting/ .

1.2 Characterizing the Styles of Multimedia Communication

How does multipoint conferencing compare with other forms of multimedia communication found on IP networks -- like streaming multimedia and Internet Telephony? There are three predominant *styles* of networked multimedia communication: netcasting, playback on demand, and conferencing.

Netcasting is a *TV-like* style based on a one-way transmission from a source to one or more receivers. The source controls the transmission. A receiver elects to view a netcast at its scheduled transmission time. The content of the netcast can be either live or pre-recorded.

Playback on Demand is a *VCR-like* style based on a one-way transmission from a store and forward source to one receiver. The receiver initiates playback on demand and controls the transmission using VCR-like functions, for example: pause, play, skip, rewind. Microsoft Windows Media[4] and RealNetworks RealSystem[5] are examples of products for netcasting and playback on demand. (Netcasting and playback on demand also are called *streaming multimedia*.)

Conferencing is a ***phone-like*** style based on a two-way or multi-way conversation among participants (transmitters or receivers). Conferencing is initiated by placing or accepting a call.

The following properties further differentiate conferencing, netcasting and playback on demand for networked multimedia communications:

Media Type refers to the form and method used to convey information from a transmitter to one or more receivers. There are three generic types of media:

- Audio – sound, especially voice.
- Video – moving images, usually in combination with audio.
- Data – text, graphics and images, often provided by a live application.

The transmission of audio, video or data over a network is called a flow or stream.

Flow Type: Given *a pair* of terminals in communication, the flows are:

- One-way, or unidirectional – one device transmits and the other receives.
- Two-way, or bidirectional – both devices transmit and receive.

Relationship Type: Given *all terminals* in communication, the flows are:

- One-to-one, or point-to-point – only between a pair of terminals.
- One-to-many, or point-to-multipoint – from one transmitter to many receivers.
- Many-to-many, or multipoint – among three or more terminals where each can be a transmitter and receiver.

Netcasting products are optimized for one-to-many communication and scale to audiences in the thousands. Typically multipoint conferencing products (based on the H.323 or T.120 standards) have a practical upper limit of 10 to 50 terminals per conference. In practice there are usually fewer. For example, the typical PSTN audio conference call has six or seven telephone connections and the typical videoconference has fewer.

---

[4] http://www.microsoft.com/windows/windowsmedia/ .
[5] http://www.real.com/products/index.html .

Although multipoint conferencing focuses on many-to-many, it can operate one-to-one and one-to-many as well.

Rate Type describes the bit rate of the flow:

- Stream, or constant bit rate – relatively steady and within a predefined maximum; applies to time domain media like audio and video.
- Burst, or variable bit rate – unpredictable and with no defined maximum, typical of interactive applications and file transfers. Data conferencing is bursty.

Delivery Type[6] describes the sensitivity of the application to delays in a flow:

- Asynchronous – the application is not time-based and there are no constraints on delivery time.
- Synchronous – application data are time-based and there is a flexible delta between application time at the transmitter and application time at the receiver. Usability is not impacted as long as the application time differential is constant and the application clock rates of the transmitter and receiver are identical.
- Interactive – application data are not time-based but communications should meet the needs of the human stimulus-response system. Delays in transmission may be noticeable to end users, but do not affect usability or functionality.
- Isochronous – application data are time-based. Usability is impaired if application time at the transmitter and the receiver are not identical or very nearly so. Application clock rates at transmitter and receiver also must be identical.
- Mission-critical – application data are time-based and require the same application time and clock rate at the transmitter and receiver. Delays disable functionality of the application.

In netcasting and playback on demand, media presentation at the receiver often lags the source by ten to fifteen seconds or more. This buffer in time (a jitter buffer within the receiver) provides a margin of safety against delay in the flows, but also limits the potential for interactivity. Thus netcasting and playback on demand are *synchronous*.

By contrast, the interactivity of audio and video conferencing requires end-to-end delay within a few hundred milliseconds[7]. Increasing delays create increasingly perceptible differences in the rendered quality of the audio and video. The delivery type for audio and video conferencing is *isochronous*. Data conferencing is *interactive*.

Table 1. Summary of Multimedia Communication Styles

Style	Metaphor	Media Type	Flow Type	Relationship Type	Rate Type	Delivery Type
Netcasting	TV	Audio and video	Unidirectional	One-to-many	Stream	Synchronous
Playback on Demand	VCR	Audio and video	Unidirectional	One-to-one	Stream	Synchronous
IP Telephony	Phone	Audio	Bidirectional	One-to-one	Stream	Isochronous
Conferencing	Phone	Audio and video	Bidirectional	Many-to-many	Stream	Isochronous
Conferencing	Phone	Data	Bidirectional	Many-to-many	Burst	Interactive

[6] Suggested in Internet Bandwidth Management: iBAND2[SM] Edition. (http://www.stardust.com/iband2/whitepaper.htm)

[7] The maximum delay defined for the PSTN by the International Telecommunication Union is 300 milliseconds. Analog cell phone delay is approximately 300 milliseconds. Videoconferencing strives for a delay less than 500 milliseconds.

This characterization is not intended to suggest limits in combining styles. For example, playback on demand in combination with the time-synchronized presentation of Web pages provides an effective unidirectional audio-video-data presentation. Or, an unanswered phone call that is forwarded to a messaging device, recorded and posted, so that it can be received via e-mail or a paging device, is another effective hybrid solution.

1.3 Media for Interactive Business Communications

Voice (audio) is the essential medium for real time business communications. The effectiveness of the spoken word to convey information cannot be replaced by still or moving images. Vocal inflections provide critical cues sometimes needed to properly interpret the words spoken. *Therefore the importance of high quality audio cannot be overemphasized.*

Pictures, graphics and visual aids can significantly increase the comprehension and retention of verbal information. Seeing what you are talking about is a close second to talking about it.

Video (moving images) adds valuable context to spoken language by communicating gestures and environmental information that can be used to interpret a speaker's words and underlying emotions. Video can also provide cues about how others are reacting to the speaker. Studies have shown that video has higher value when people are getting to know one another than after voice-gesture associations and related artifacts have been assimilated. Video may have its highest payoff in special situations like negotiation, persuasion, interviewing, counseling, motivational speaking; or, in supporting applications like telemedicine or instructor led distance learning.

2 Drivers of Multipoint Conferencing

This section examines some of the factors and trends that are creating the need and opportunity for multipoint conferencing services.

2.1 Cost and Time Efficient Business Meetings

Meetings are how business people come together to interact. MCI[8] reports that every day in the U.S. there are about 11 million business meetings. According to the National Statistics Council, 37% of employee time is spent in meetings. Meeting expert Roger Mosvick[9] finds that middle managers spend up to 40% of their time in meetings; and senior executives, up to 80%.

The trend is toward more meetings. MCI found that 46% of professionals are attending more meetings than a year ago, while 8% said they were attending fewer. A 3M study found that non-managers spend twice as much time in meetings as 10 years ago.

The goals of meetings vary considerably: from information sharing and dissemination to brainstorming, strategy formulation and selling. Meetings whose purpose is to expedite action, to resolve problems, to complete deliverables or to decide issues are valued highly.

A 3M study found that in 1998, 25% of its internal meetings included at least one remote participant. The trend toward increasing numbers of remote meeting participants is expected to increase. MCI found that the cost of an in-person meeting requiring plane travel can exceed that of an audio conference meeting by more than seven times.

Table 2. Trends Driving the Need for Group Communication over Distance

Trend	Example
Reduced travel costs	Travel costs include lost productivity, time wasted in transit and deferred office tasks; plus, the hard costs of transportation, meals and lodging.
Faster cycle times	Projects are less tolerant of the critical path dead time that often results from reconciling busy calendars in order to schedule in-person meetings.
Geographically dispersed workforce	Globalization, mergers and acquisitions, SOHO, telecommuters, road warriors, regional and interdepartmental teams, intra-field and corporate-field teams are factors.
Increased partnering and outsourcing; collaborative extranets	More and more business activities are focused externally. Critical collaborative activities include project definition and project management.
Fewer subject experts	Flatter organizations are requiring greater leverage of fewer subject experts, for example, for call centers, customer support, training and consultative selling.

Finally, MCI found that 89% of professionals believe technology will make meetings easier and that 79% of those who had used video conferencing exclusively in a conference room would use it more if available from their desktop.

[8] Meetings in America, An MCI WorldCom Conferencing White Paper. http://nmc.mci.com/MeetAmerica/mtg_usa_wp.shtml (Also Meetings in the UK, An MCI WorldCom Conferencing White Paper. http://nmc.mci.com/meetuk/ukwhitepaper.html)

[9] "Another Meeting? Ughhhhh", Shelly Coolidge, Christian Science Monitor, March 8, 1999.

2.2 Audio Conferencing – An Established Business Tool

Audio conference calls on the public telephone network are a widely adopted business tool in North America and usage is on the rise worldwide. Telespan Publishing Corporation reports[10] that in 1997 North American service bureaus hosted 7.5 million conference calls, compared to only 250,000 in the U.K and fewer than 100,000 in Japan[11]. And the adoption of American-style business practices may be driving higher annual growth rates in the international arena: 100% in Japan and nearly 70% in the U.K. U.S. service bureaus report the growth rate of calls to be 20% to 40%, significantly above their 10-year average.

Comparatively the adoption of videoconferencing is in an early stage. The volume of service bureau hosted multipoint videoconference calls is just a fraction of a percent of hosted audio conference calls. Availability, cost, ease of deployment, ease of use and cultural issues are often cited factors.

Data conferencing with audio conferencing adds the ability to deliver visual information with the spoken word. Proponents of this combination claim that it delivers 80% of the benefits of video conferencing for 20% of the cost. (A common use of video conferencing is to transmit images of presentation materials via camera. In addition to interactive data conferencing, there are also one-to-many products that push presentations via the web to audiences who listen simultaneously via teleconference.)

Many new services are springing up for audio and data conferencing. Established providers of PSTN audio conferencing services like MCI WorldCom[12] and Latitude[13] have rolled out complementary data conferencing services. Other new products and services are being introduced. Lotus[14] Sametime provides data collaboration within the Notes/Domino platform. eMeeting.net[15] is a recently launched service that combines PSTN and IP audio conferencing with NetMeeting. WebEx[16] is the first service to provide integrated PSTN audio and data conferencing through a Java-enabled browser.

The established practice of audio conferencing is a platform on which richer and more effective forms of interactive communications can be built to reduce the cost and increase the impact of business meetings.

2.3 Online Socialization and Communities of Interest

The Internet is a personal place in addition to a place of business. People use the Internet to socialize, to seek out other people with similar interests, or to engage friends and family who are distant. Chat rooms are places people meet in groups to converse interactively via text messages.

IDC reports[17] that 27% of U.S. Internet users have friends they met online and that 11% of this group uses Instant Messaging or Buddy Lists regularly. *Cable World* magazine reports[18] that chat rooms are a common activity for 44% of online users.

---

[10] *The Future of Global Conference Calls: A Look at Opportunities for Growth*, A Prospectus, Telespan Publishing Corporation, April 1999.

[11] These figures include only service bureau provisioned conference calls -- not those completed directly, for example, via PBX.

[12] http://www.wcom.com/services_for_business/on_net/integrated_conf.shtml .

[13] http://www.latitude.com/products/index.htm .

[14] http://www.lotus.com .

[15] http://emeeting.net/ .

[16] http://www.webex.com/ .

[17] “Online Nation: 1998 U.S. Internet User Survey”, Jill L. Frankle, IDC Report #17684, December 1998.

[18] “Internet Users Reluctant to Access Internet through TV”, *Cable World*, November 30, 1998.

Audio and video can add new dimensions to the text chat experience. Just as there is often a progression from a public chat space to a private space there may be similar progressions for self-selecting groups who wish to move from text chat to audio or video modes of communication.

Audio and video group communication can aid in the development of communities of interest. These self-interacting groups are increasingly becoming key elements of web portals whose participants are differentiated by interest or professional affinity. In the past, user groups have been responsible for facilitating relationships among communities of users. The benefits of these relationships have ranged from idea generation, consensus building and common action to the identification, expedition and resolution of specific problems. Credit for benefits accrues in part to the group's sponsor. Conferencing, in combination with other tools for communication and collaboration, can aid the development of these synergistic relationships among widely dispersed groups of individuals.

First generation audio and video chat services are just beginning to spring up on the internet. For example, Time Warner's Road Runner is offering a desktop video conferencing[19] service.

2.4 ISP Value Added Services and Network Convergence

With decreasing margins on basic access service, Internet Service Providers are looking for added value services to expand profitability. In addition, traditional ISPs face the threat of well-heeled, incumbent telecommunication providers and emerging network service providers who have targeted data networks and services as the next growth market.

In aggregate these providers are offering access service, developing and rolling out next generation IP networks, and planning new services that include multimedia communications capabilities. A Gartner Group strategic planning assumption states, "ISPs that cannot offer the same mix of voice and data services to enterprises as offered by converged network service providers by 2001 will be relegated to consumer or business niche status by 2003 (0.8 probability)."[20]

Business meetings and audio/video chat are activities that can be supported directly by multipoint conferencing. As a communications service, multipoint conferencing can be integrated within other applications and services, for example, distance learning and call centers for customer services and support. It can be an important communications mechanism supporting the processes of major e-business systems like knowledge management, customer relationship management and the global value chain.

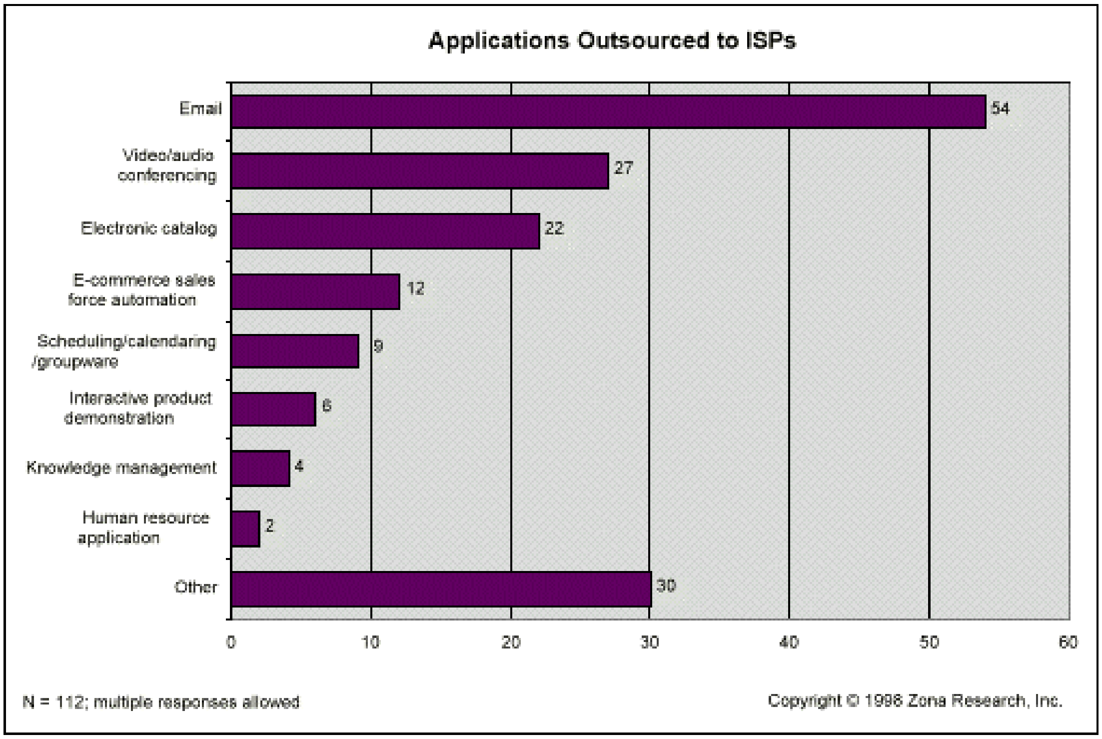
Business customers represent a major source of revenue for value added conferencing services. The ZONA Enterprise Usage Study[21] found that 27% of senior IT professionals with purchasing authority for internet communications said they would be likely to outsource video and/or audio conferencing to their ISP.

The needs and incentives are in place for an important and expanding role for multipoint conferencing in facilitating distributed business meetings, online socialization and communities of interest. The challenge for service providers is to develop truly consumable services.

[19] http://www.neo.rr.com/rr/info/videoconferencing.html .
[20] "Converged Network Services Accelerate ISP Consolidation", E. Paulak, Gartner Group, Inc., Research Note Strategic Planning, 17 August 1998.
[21] ZONA Enterprise Usage Study - Internet Communications Databook, ZONA Research, Inc., Q2 1998.

ZEUS Communications Databook – Second Quarter 1998

Q: What applications is your organization likely to outsource to your ISP?



3 Models for Hosted Multipoint Conferencing Services

This section introduces four communication models for hosted multipoint conferencing services. The models support business meetings and multimedia chat services.

The models are expressed in terms of media and networks. The focus is on IP networks and audio over PSTN. In this context *IP network* means *IP over anything*, in essence, a network that is IP at its edges. IP is the native network for ISPs today as well as the consensus universal network of the future. PSTN audio recognizes teleconferencing as the current "killer app" for conferencing, as well as the ubiquity and continuing importance of the "black phone."

Table 3. Models for Hosted Multipoint Conferencing Services

Network and Media				
Model Number	**PSTN Audio**	**IP Audio**	**IP Data**	**IP Video**
1				
2				
3				
4				

The numbered rows in Table 3 represent the four models. A shaded cell indicates inclusion of that column's network/media combination in the model.

The network / media combinations for each model are cumulative and interoperate fully across common media. For example, a video conferencing service based on Model 4 should accommodate an audio-only participant who is connecting by PSTN phone.

A conferencing "bridge" -- also called an MCU or Multipoint Control Unit -- is central to the implementation of these models. The function of a bridge is to provide a common point of communications exchange for the participants in a conference. Conference participants connect by calling the bridge or answering a call from the bridge. Typically it is a server-based device.

The location of the bridge on the network is an important design issue. For example, a bridge might be located on customer premises inside the firewall, on a "green net" or DMZ at the edge of the customer's network, near the edge router that connects a VPN, or outside the firewall at the service provider's premises. Considerations in locating a bridge include network performance and security issues.

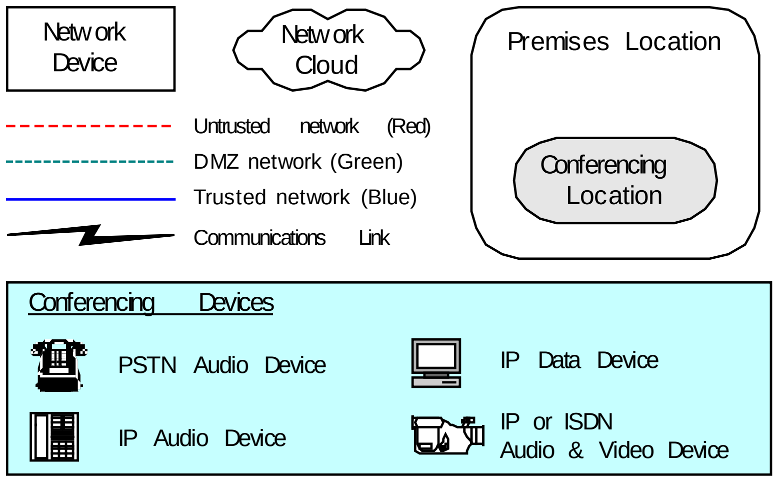
The models assume that the service provider is able to manage servers that are located on a customer's premises.

The subsections following introduce the four models. Each model is a potential entry point for a multipoint conferencing service. Each model is illustrated by an example that can be used as a conceptual point of departure for design.

Start with a practical implementation today and add more advanced capabilities as, for example, solutions for network quality of service issues are advanced.

Figure 1 provides a symbol legend for the examples in this section.

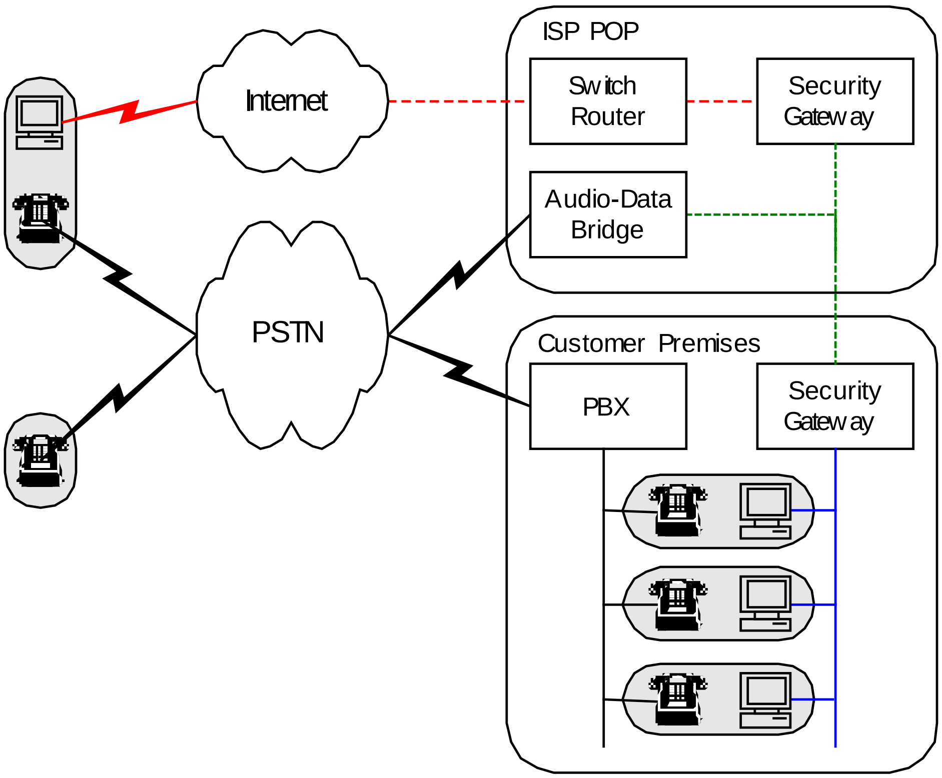
Figure 1. Legend for Model Examples



3.1 Model 1: PSTN Audio and IP Data Conferencing

PSTN audio conferencing and IP data conferencing are an attractive combination because the participants use familiar devices that already are installed on their desktops: a PSTN phone and a personal computer. This service extends the established business practice of audio conferencing to include the benefits of data sharing and collaboration. Implementations of the model should support audio-only conference participants.

Figure 2. Model 1 Example



The Model 1 example shows three conferencing locations on the customer premises. Two more are off premises. One of these has audio-only capability. When connected to the bridge the participants can use an audio device (phone) to converse and a data device (PC) to share data and collaborate. The bridge does not synchronize the audio and data streams.

The bridge connects both IP data and PSTN audio communications and therefore can control all conference and connection resources. This simplifies conference scheduling and operations. Implementation of the PSTN interface on the bridge employs an integrated gateway with a T1/E1 interface. The bridge mixes and distributes the audio streams and broadcasts the data conference.

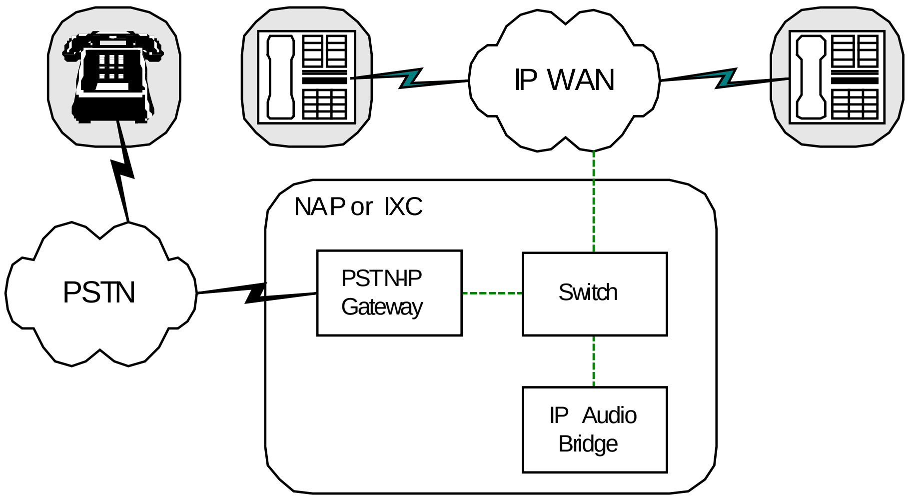
A variation of the Model 1 Example uses a standalone PSTN-IP gateway as shown in the Model 2 Example. Another variation uses independent audio and data bridges. The audio bridge could be a traditional PSTN audio bridge.

This service does not require *quality of service* (QoS) IP networks. It leverages the reliability and accepted quality of PSTN audio.

3.2 Model 2: IP Audio with PSTN Audio Conferencing

IP Audio with PSTN Audio Conferencing is the IP equivalent of the PSTN teleconference. Participants use an IP audio device or a PSTN audio device. With this service an Internet Telephony Service Provider (ITSP) can add teleconferencing to existing two-way phone services.

Figure 3. Model 2 Example



The Model 2 example shows three conferencing locations. IP audio devices connect over an IP network. PSTN audio devices connect over the PSTN and through a PSTN-IP gateway. Audio devices can be located wherever network logistics permit. The bridge is co-located at a NAP or other traffic exchange location on the network backbone. The participants can speak with one another when connected to a conference.

The bridge connects the IP audio streams, mixes them and transmits the resulting streams back to the conference participants.

A variation of the Model 2 Example uses a bridge with an integrated gateway, for example, one with a T1/E1 interface as shown in the Model 1 Example. Another variation places the bridge in a customer premises environment, for example, near an Internet Telephony “IP PBX” or “un-PBX”.

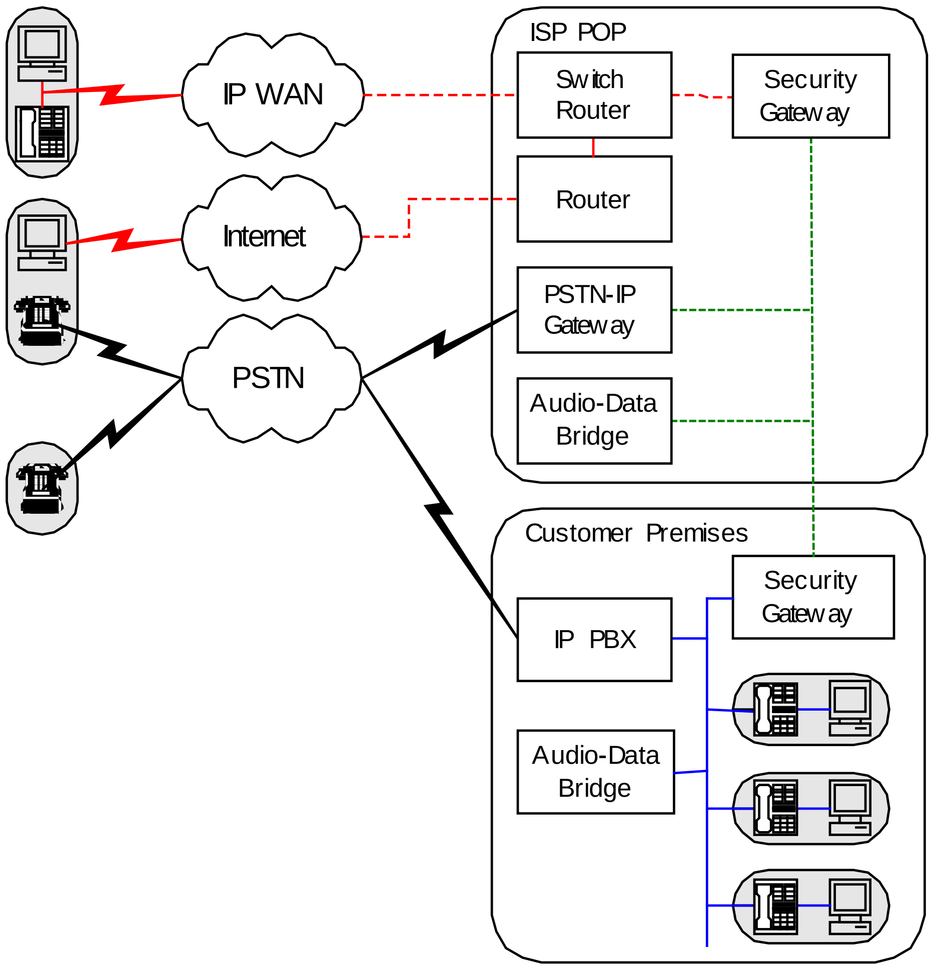
This service is well suited for existing Internet Telephony environments. Audio conferencing is a strategic service for ITSPs who ultimately must offer the entire range of services found on the PSTN.

For business quality audio this service must operate on IP networks that have been engineered for quality of service.

3.3 Model 3: IP Audio, PSTN Audio and IP Data Conferencing

Model 3 combines Model 1 and Model 2. IP audio capability allows a conference participant to utilize a communications device with integrated IP audio and data conferencing features (such as Microsoft NetMeeting). A participant can also connect in audio-only mode.

Figure 4. Model 3 Example



The example shows three conferencing locations on premises with IP audio and data capability. Three more are off premises. One of these (top left) has the same capabilities as on premises locations. Assume the second (middle) external location does not have the network quality of service required of audio and therefore connects via the PSTN for audio while using IP for data conferencing (Model 1). The third (bottom) external location has only PSTN audio capability.

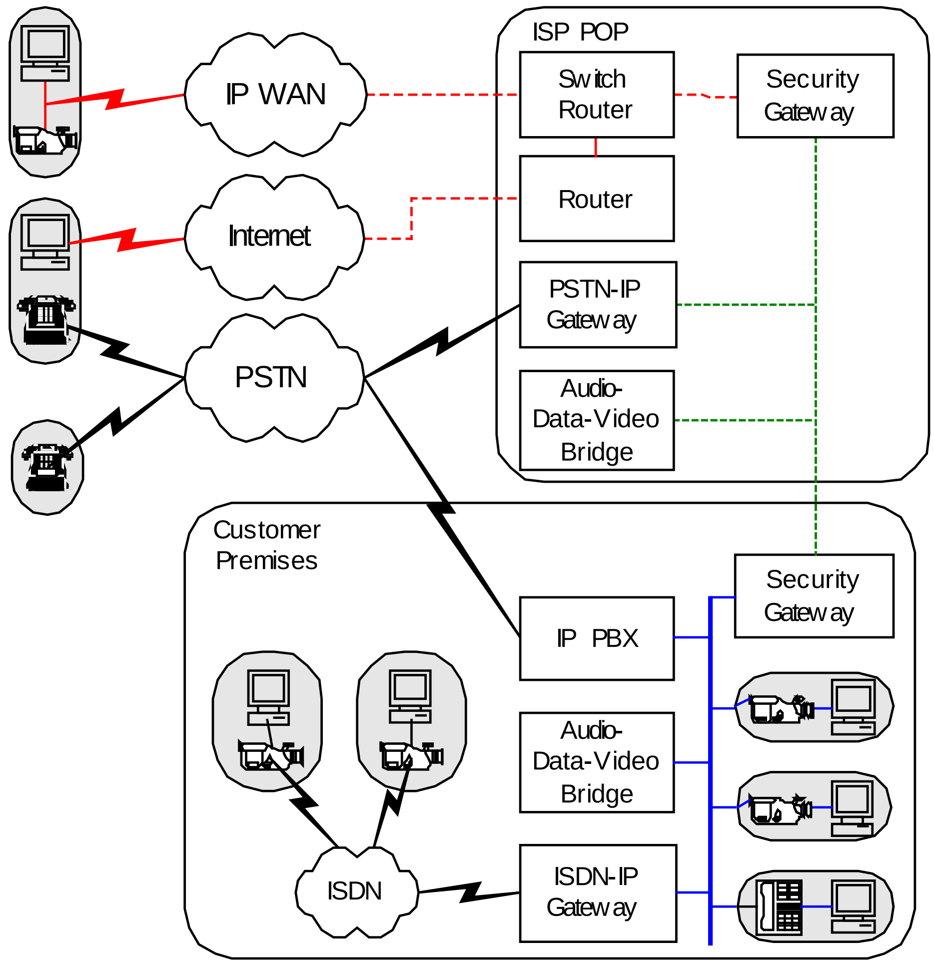
When connected in conference the users can speak with one another via audio and share data and collaborate using a data device. The audio and data streams are not synchronized.

Each bridge in the example connects IP audio and IP data streams. The gateway and IP PBX convert between PSTN audio and IP audio formats. A conference can be hosted by a single bridge or by the two bridges in tandem. A bridge could also be located in the POP outside the security gateway (red net) or at a traffic exchange location on the network backbone. Model 3 requires a network that has been engineered to provide IP audio quality of service.

3.4 Model 4: IP Audio, PSTN Audio, IP Data and IP Video Conferencing

The fourth model adds video capability to Model 3. Conferencing locations that are not video-equipped can conference in audio-only mode or with audio and data.

Figure 5. Model 4 Example



The example shows, on premises, three conferencing locations on the IP network and two other ISDN locations. The ISDN locations are connected to the IP network through an ISDN-IP gateway. The two ISDN and two of the three IP locations are enabled for audio, video, and data conferencing. The fifth has only IP audio and data capability. There are also three external conferencing locations. One has IP audio, video, and data capability. The second has PSTN audio and IP data. And the third has only PSTN audio capability.

When connected in conference, participants can speak to one another via audio and, depending on equipment available at the conferencing locations, transmit and receive video images and/or share data and collaborate.

The bridges connect IP audio and PSTN audio (via the gateway), IP data and IP video streams. In addition to mixing audio and routing data streams, the bridge also processes and routes video. Only audio and video streams are synchronized. The bridges can work separately or in tandem.

The ISDN-IP gateway connects the two ISDN videoconferencing locations to the IP audio-data-video bridges. An alternative location for this gateway is at the ISP POP. This gateway is used only if support for ISDN based conferencing is required.

This is the most challenging service to implement because it requires high bandwidth and low packet delay on the IP network and the most sophisticated equipment and training for users.

4 Technical Components of Hosted Conferencing Services

This section looks at a commercially available, standards-based technology that can be used to create multipoint conferencing services.

4.1 The H.323 Standard

H.32X is a comprehensive family of multimedia teleconferencing standards. These standards, called "Recommendations", have been developed by the Telecommunication Sector of the International Telecommunication Union[22] (ITU-T), an agency of the United Nations, formerly the CCITT.

Table 4. ITU-T Multimedia Teleconferencing Recommendations

Recommendation	Applicable Networks	First Ratified	Last Revised[23]	Media
H.320	ISDN, circuit-switched	1990	May 1999	Audio and video
T.120	All listed here	1996	May 1999	Data
H.323	IP, LAN, packet-switched	1996	May 1999	Audio and video
H.324	POTS, V.34 modem, low bit rate	1996	September 1998	Audio and video

H.323[24] is the umbrella standard of a protocol suite for conferencing over IP (packet-based networks). H.323 is also used with Internet Telephony. And it has been designed to interoperate with conferencing over ISDN[25] (H.320).

The H.323 umbrella includes T.120, a protocol suite for data conferencing. T.120 may also be used independently. Microsoft NetMeeting is based on the H.323 and T.120 standards.

H.323 specifies use of Real-Time Protocol (RTP) and Real-Time Control Protocol (RTCP), both Internet Engineering Task Force (IETF)[26] standards[27]. RTP provides packet sequence and time stamp information for audio and video flows. RTCP allows the transmitter and receiver to communicate status of the media flows, for example, network delay and dropped packet information. RTP and RTCP do not reserve network bandwidth or guarantee quality of service.

H.323 conferencing over IP is implemented with TCP and UDP.

- TCP is used for call control and data flows. TCP provides reliable delivery and is "congestion aware". If a packet is dropped, for example by a busy router, the receiver will request that it be resent. On receipt of requests for retransmission, the sender reduces its transmission rate to alleviate presumed congestion on the network.
- UDP is used for audio and video flows and related control. It underlies RTP and RTCP. It does not guarantee delivery and is unaware of dropped packets or network congestion. For audio and video media there is no value in a packet that arrives after its presentation time. By

[22] The ITU's web site is http://www.itu.int/ .
[23] The ITU-T Study Group 16, Multimedia Services and Systems, met in Santiago, Chile from May 17-28, 1999 and approved changes to the T.120 and H.323 Recommendations. The text of the revisions should soon be available on the ITU site. The so-called H.323 version 2 Recommendation (February 1998) is the reference standard for this white paper.
[24] ITU-T Recommendation H.323 (1998) "Packet Based Multimedia Communication Systems."
[25] ISDN, *Integrated Services Digital Network*, is a suite of communications services for Public Data Networks first defined by the ITU-T in 1984.
[26] The IETF's web site is http://www.ietf.org/ . An introduction to the IETF can be found at http://www.ietf.org/tao.html .
[27] The RTP and RTCP protocols (RFC 1889 and 1890) have been developed by the Audio/Video Transport (avt) Working Group within the IETF Transport Area, see http://www.ietf.org/html.charters/avt-charter.html .

the time a packet can be identified as "late" there usually is not enough time for retransmission.

Table 5. H.323 Umbrella Specification

<table>
<tr><td colspan="2">Audio & Video Applications</td><td colspan="3">Terminal Control and Management</td><td>Data Applications</td></tr>
<tr><td>G.7XX</td><td>H.26X</td><td rowspan="2">RTCP</td><td rowspan="2">Terminal to Gatekeeper Signaling</td><td rowspan="2">Call Signaling, Setup and Connect, Capability Exchange, etc.</td><td rowspan="2">T.120</td></tr>
<tr><td colspan="2">RTP</td></tr>
<tr><td colspan="4">UDP (Unreliable Transport)</td><td colspan="2">TCP (Reliable Transport)</td></tr>
<tr><td colspan="6">IP</td></tr>
</table>

Today commercial products based on H.323 and T.120 represent the most widely deployed standards based approach for conferencing over IP. But the future might hold viable alternative approaches. Refer to Appendix A for more information and perspective on the evolution of multipoint conferencing standards development.

4.2 Networks

4.2.1 Networks

Communications for conferencing can be carried over many kinds of networks, including packet-switched, such as IP, circuit-switched, such as ISDN or the PSTN, ATM[28], or hybrids, such as IP over ATM.

4.2.2 Bandwidth

Bandwidth utilization and variability on the network differs considerably by media. For audio and video the maximum bit rates are determined by the methods used to encode the media into digital formats. Actual bit rates vary downward from the maximum.

Business quality videoconferencing typically runs at 128, 256 or 384 Kbps per (unidirectional) stream. Higher bit rates produce better quality by allowing larger or higher contrast images delivered at higher frame rates. Bit rate variability from the maximum rate depends largely on the amount of motion in the captured frames. Typical variations are from 25% to 100% of the maximum rate.

Data flows are not governed by a maximum. Data bit rates are typically limited by the capacity of the source device to deliver data to the network interface and the network to accept it. One result is that bursts of data traffic can disrupt lower bit rate audio and video flows by causing packet delays or queues to overflow in the routers and switches that handle the aggregated traffic.

Data traffic is extremely bursty. Variability in the bit rate of data traffic can be greater than an order of magnitude.

[28] ATM (Asynchronous Transfer Mode) is a connection-oriented, cell-switching technology for broadband signals. For more information visit the ATM Forum at http://www.atmforum.com/ .

Table 6. Bit Rate Variability by Media

Media/Operation	Typical maximum bit rate range (per one-way stream)	Bit rate variability
Audio Flow	6 Kbps to 64 Kbps	Very Low
Video Flow	56 Kbps to 1.5 Mbps	Moderate
Data Flow	22 Kbps to 925 Kbps*	Very High
Network File Copy	2.75 Mbps*	Extreme

*Maximum rates measured on a LAN for NetMeeting data conferencing and for network file copy. Unlike audio and video whose maximum rates are determined by digital signal processing algorithms, these operations have no fixed maximum.

4.2.3 Quality of Service

Specially conditioned IP networks are critical to business quality audio and video service. The networks *must* be able to deliver audio and video packets at prescribed bandwidths and within predefined packet delay and packet loss limits. Business quality audio and video are generally not achievable over the "best effort" Internet at this time.

Significant work in the IETF and other organizations is currently directed toward developing the controls necessary to support "quality of service" guarantees for a full range of applications including audio and video. For an introduction to network QoS issues and standards initiatives refer to Internet Bandwidth Management: iBAND2[SM] Edition[29].

While inter-domain QoS for the internet may be at least several years away, essential tools for building high performance QoS networks within an Enterprise or service provider domain have already begun to roll out. Solutions from network vendors are due starting in 2000. QoS is a significant issue but it need not deter pilots or bounded implementations of conferencing services.

4.2.4 Network Security

Traversing security devices like firewalls, proxy servers or network address translation devices deployed at administrative boundaries imposes challenges for H.323 protocols and for conferencing services based on H.323.

For a technical statement of the problem and keys to a solution see The Problem and Pitfalls of Getting H.323 Safely Through Firewalls[30].

At this time few firewalls support H.323. Devices that do offer potential solutions are Check Point's FireWall-1, Cisco's PIX Firewall and Cisco's Multimedia Conference Manager. Investigation is recommended to ensure that one of these can meet specific requirements. Tunnels and VPN devices also offer a potential solution to this problem.

[29] http://www.stardust.com/iband2/whitepaper.htm .
[30] http://support.intel.com/support/videophone/trial21/H323_WPR.HTM .

4.3 Components of an H.323 Network

H.323 defines four logical components that interoperate to form the nucleus of a multipoint conferencing system.

Table 7. The Four Major Components of an H.323 Network

Component	Required?	Description
Terminals	Yes	Client devices that place and answer calls and communicate in real time via audio, video and data.
Gateways	No	Devices that enable communication between H.323 Terminals and terminals on other networks.
Gatekeepers	No	Devices that manage a collection of H.323 components called a zone
Multipoint Control Units (MCU)	No	A bridge device that enables conferencing among three or more Terminals.

Note: In H.323, an "endpoint" is a Terminal, Gateway or MCU. An endpoint may originate and terminate calls. Gateways and MCUs are endpoints because their functionality includes that of a Terminal. Gateways and MCUs need Terminal functionality to communicate transparently with actual Terminals.

Platform considerations are outside the scope of H.323. In practice Terminal platforms range from hand-held devices and appliances with embedded software, to PCs and workstations with appropriate media peripherals installed, and off-the-shelf software. Gateways, Gatekeepers and MCUs most often are implemented as software running on a server under Windows NT or UNIX, sometimes with specialized add-in hardware. Less often they can be found as an appliance.

Neither does H.323 specify how components may be combined. There are products, for example, that combine a Gatekeeper with an MCU, a Gateway with an MCU, or an MCU with a Terminal.

An overview of the capabilities and interworkings of the major H.323 components follows. For a more technical discussion of this material refer to Demystifying Multimedia Conferencing Over the Internet Using the H.323 Set of Standards[31]. Additional references to supplementary material are provided in Appendix B.

4.3.1 Terminals

A Terminal is a real time communications device on the network. An H.323 Terminal must support audio and may optionally support video and/or data communications. When a call is setup, Terminals exchange media capabilities and may request a preferred mode for receiving audio or video streams. Bandwidth is requested based on bit rates of the media streams to be used for the call.

4.3.2 Gateways

A Gateway transparently interconnects Terminals on an H.323 network with those on a switched circuit network. Functions include translating between communications protocols and procedures and transcoding audio between formats. H.323 does not standardize the functions of Gateways. Examples are Gateways that interconnect to phones on the PSTN and Gateways that interconnect to ISDN videoconferencing systems (as illustrated in Figure 5).

[31] http://developer.intel.com/technology/itj/q21998/articles/art_4.htm .

4.3.3 Gatekeepers

A Gatekeeper is the "operations manager" for a logical collection of Terminals, Gateways and MCUs called a "zone". The specific makeup of the collection is left to the network designer. H.323 does not prescribe any relationship between a zone and network topology. A zone is defined simply by the collection of endpoints that are registered to a Gatekeeper. An endpoint may register with only one Gatekeeper at a time. Once registered, the Gatekeeper performs services for the endpoint and the endpoint is obliged to take procedural direction from the Gatekeeper.

Table 8 summarizes the required and optional services and functions of a Gatekeeper. Note that H.323 provides for interposition of local policy for admissions control, call authorization, and bandwidth management. For example, Radius Authentication may be used to authorize calls.

Table 8. Gatekeeper Services and Functions

Function	Required?	Description
Address Translation	Yes	Process requests to translate alias addresses into transport addresses. Process requests to update address translation table. Other methods of updating the translation table are also allowed.
Admissions Control	Yes	Process requests for admission to network. Confirm or reject based on call authorization, requested bandwidth or other policy mechanism. Confirms request if no policy mechanism is implemented.
Bandwidth Control	Yes	Process requests to modify bandwidth usage. Confirm or reject based on bandwidth management policy. Confirms request if no policy mechanism is implemented.
Zone Management	Yes	Provide the functions listed above to the endpoints registered to its zone.
Call Control Signaling	No	Allows the Gatekeeper to process call control signals on behalf of an Endpoint in a two-way call. Called Gatekeeper Routed Calls, this allows the Gatekeeper, for example, to auto-forward a call for an Endpoint.
Call Authorization	No	Allows the Gatekeeper to authorize calls via authentication or other policy-based mechanisms.
Bandwidth Management	No	Allows the Gatekeeper to implement bandwidth management policy; for example, to reject calls if network bandwidth becomes oversubscribed.
Call Management	No	Provides an "active call" list that Gatekeeper can reference to determine if a Terminal is "busy", or use as resource for bandwidth management function.

4.3.4 Multipoint Control Units (Bridges)

A Multipoint Control Unit (MCU) is the key component of a multipoint conferencing service. It functions as an intelligent "bridge" between the endpoints in a conference. An MCU consists of two components: the Multipoint Controller (MC) and the Multipoint Processor (MP). An MCU contains exactly one MC and zero or more MPs. In implementation the MC and MP elements may be distributed across servers.

The MC performs control functions. It manages the negotiation of media capabilities among conference endpoints and with the Endpoints determines which media streams are to be used, including multicast streams. It does not handle the media streams directly.

Note: H.323 multicast refers to the process of transmitting a media stream from one source to many destinations. The actual mechanism can be IP multicast, if available, or multiple unicast.

The MP mixes, switches and processes audio, video and/or data streams as determined by the MC.

As part of mixing and switching the MP may convert between audio, video or data formats and bit rates to better support endpoints with different media formats or bit rate capabilities. It can create customized streams for different endpoints. The MP supports lip synchronization by resolving any time base discrepancies in the audio and video streams. The MP may distribute the audio and/or video streams via IP multicast, if available.

H.323 Video processing

The MP processes video by either switching or mixing. Switching selects an output video stream from the input streams. The MP may switch among the input streams based on a change of speaker as sensed by audio. Video mixing formats more than one video source into the output stream, for example, to implement a "continuous presence" feature by creating a two by two picture matrix for display by the receiver.

H.323 Audio Processing

The MP processes audio to generate *n* audio outputs from *m* inputs by switching, mixing or a combination of these. Audio mixing requires that input streams be decoded into a pulse code modulated format or into analog signals, then combined and recoded into the appropriate output formats. In combining the signals the MP may eliminate or attenuate signals to manage noise or other unwanted signals. Each audio output may have a different mix, for example, to facilitate private chats within a conference. To prevent echo the speaker's voice should be eliminated from the stream returned to the speaker.

Simple switching dynamically selects one of the input streams to be the output stream, usually the loudest signal. This technique does not yield audio quality competitive with PSTN services.

T.120 Data Processing

Data conferencing, standardized by the T.120 Recommendation, is an optional component of an H.323 implementation. T.120 also may be used independent of H.323. In this mode a Terminal supports data conferencing only.

H.323 assigns the task of processing T.120 data to the MP. Happily, many implementations of MCUs support T.120 both ways: with or without H.323. The approach recommended by H.323 is to use H.323 call setup whenever possible: simply turn off audio and video if data conferencing only is desired.

T.120 provides conference controls, such as the announcement of persons entering and leaving a conference. It also provides the controls needed to share applications and other data resources within a collaborative framework so that control can quickly be switched among the participants of a conference.

Centralized and Decentralized Conferencing

H.323 supports centralized and decentralized conferencing. In the centralized model a point-to-point connection is established between each endpoint and the MCU. All communication among the endpoints passes through the MCU.

In the decentralized model each endpoint multicasts one or both of its audio and video streams to the other endpoints. The endpoints provide their own audio mixing, video switching and lip synch for streams received from other endpoints. In any case, H.323 requires an MC for control function and an MP if data conferencing is used.

The decentralized model adds complexity and is quite possibly more problematic for service providers. Consider these factors: (1) Decentralized multipoint capability is an optional feature of Terminals and may not be found in "thin" Terminals like IP phones. (2) It places a disproportionate processing load on the Terminals. (3) Without IP multicast the decentralized model can consume more bandwidth than with the centralized model. (4) Increased complexity makes operating procedures more difficult and costly to design, provision and support.

Cascading

Cascading allows a single multipoint conference to be distributed across multiple MCUs. Cascading can be a useful tool for managing bandwidth on low bit rate or expensive network segments like WANs. For example, assume that two MCUs are deployed, one on each of two LANs. The LANs are connected over a lower bit rate WAN. To establish a conference, terminals connect to the MCU on their respective LANs where bandwidth is plentiful and inexpensive. The two MCUs connect to each other over a single WAN connection. (Alternatively each Terminal on one of the LANs would connect with the MCU over the WAN.)

4.4 H.323 Audio and Video Encoding

In H.323 audio and video compatibility is standardized through codecs. Codecs are responsible for coding (compressing) the media stream at the transmitter and reciprocally decoding (decompressing) it at the receiver.[32] Codecs can be implemented in hardware or software.

Table 9 lists the codecs specified by the H.32X standards.

Table 9. H.32X Codec Compatibility

Media	H.320	H.323	H.324
Audio	G.711 (R) G.722 G.728	G.711 (R) G.722 G.723.1 G.728 G.729A	G.723.1 (R)
Video	H.261 (R) H.263	H.261 (R) H.263	H.261 (R) H.263 (R)

(R) = Required; others are optional

Table 10 provides information on required and optional video formats.

Table 10. Formats of the Videoconferencing Codecs

Format*	Image Size (in pixels)	H.261**	H.263***
Sub-QCIF	128 x 96	Optional	Required
QCIF	176 x 144	Required	Required
CIF	352 x 288	Optional	Optional
4CIF	704 x 576	N/A	Optional
16CIF	1408 x 1152	N/A	Optional

*Common Intermediate Format (CIF) is compatible with the NTSC, PAL, and SECAM analog formats.
**H.261 frame rates can be 7.5, 10, 15 or 30 fps. Bit rates range from 40 Kbps to 2 Mbps.
***H.263 bit rates range from 8 Kbps to 1.5 Mbps.

[32] For a complete discussion of video codecs, refer to the ActiveAnswers White Paper, *Video Streaming Technology*, ECG068/0798.

Table 11 lists the characteristics of the audio codecs. More complex compression algorithms require greater computational horsepower and can increase end-to-end signal latency.

Table 11. Audio Codec Characteristics

Codec	Bit rate (Kbps)*	Signal bandwidth (KHz)
G.711	64	3.4
G.722	64	7
G.723.1	6.3 / 5.3	3.4
G.728	16	3.4
G.729A	8	3.4

*Bit rate is for one unidirectional channel.

4.5 H.323 Security

The H.235 (H.secure) Recommendation[33] was added in version 2 of H.323. It addresses security issues related to authentication, integrity, privacy and non-repudiation. Authentication ensures that the participating endpoints are who they say they are. Integrity ensures that data received have not been tampered with en route. Privacy provides for the encryption of data between endpoints so that it cannot be viewed en route. Non-repudiation means that a participation in a conference cannot later be denied.

H.235 can utilize IPsec[34] as a supporting protocol. H.235 has not been widely implemented to date.

4.6 H.320-H.323 Videoconferencing Gateways

H.320 videoconferencing systems have been in use over six years and products are relatively mature. IDC estimates[35] that at year end 1998, approximately 80% of the installed commercial videoconferencing market operated on H.320. While H.320 delivers good quality, its cost and complexity of operation are inhibitors to growth. Further, it is doubtful that ISDN will ever become a widely deployed access technology.

In contrast the H.323 videoconferencing market is new and growing steadily, especially on LANs. Microsoft has provided a ubiquitous client in NetMeeting. Vendors like Intel, PictureTel, VCON and White Pine also have developed popular videoconferencing clients. Lack of network bandwidth and quality of service guarantees remain as major technical inhibitors to mass deployment. As these roadblocks are removed, the widely anticipated transition from H.320 to H.323 conferencing should accelerate.

Over the next few years those who would provide H.323 video services to enterprise businesses or others with investments in H.320 videoconferencing systems also should be prepared to address the need for compatibility with H.320 videoconferencing.

H.320-H.323 gateways answer this need with a modular approach to transcoding between the two protocols and their respective formats, as illustrated in the Model 4 Example. In addition, certain vendors provide integrated H.320-H.323 interoperability within single products.

---

[33] ITU-T Recommendation H.235 (1998) "Security and Encryption of H series (H.323 and other H.245 based) Multimedia Terminals."

[34] IETF "IP Security Protocol" working group. See http://www.ietf.org/html.charters/ipsec-charter.html .

[35] "The Evolving Videoconferencing Markets: Desktop, Set-Top, and Compact Videoconferencing in Focus", Randy Giusto and Maureen McManus, IDC Report #19785, August 1999.

4.7 Summary and Comments

H.323 is a comprehensive and complex specification that contains conditional and optional provisions.

MCUs and Gatekeepers are the two key components of an H.323 conferencing service. MCUs are “bridges” that interconnect the users in a conference. A Gatekeeper administers calls within its zone, providing services such as the translation of aliases to transport addresses, bandwidth management, authentication and the collection of information for billing. Gateways are an optional component that may be used to conference with PSTN phones or H.320 Terminals. Terminals are necessary components. Typically these would be provided by customers.

Interoperability is not guaranteed by the Recommendation but is achieved through methodical interoperability testing. The International Multimedia Teleconferencing Consortium, Inc.[36] (IMTC) focuses on the stabilization of multimedia teleconferencing standards through the promotion and facilitation of interoperability testing events.

The feature set of H.323 is rich and flexible. It is good practice to double check that all needed or expected features of selected products are present and that the products have proven interoperability across a target feature set. Before deploying widely, testing is recommended to ensure that all components interoperate as a system.

[36] The IMTC’s web site is http://www.imtc.org/ .

5 Creating Conferencing Services

While components based on the H.323 standard can provide a foundation for building a commercial conferencing service, the scope of H.323 is not sufficiently broad to address all aspects of a service, for example its use and operation. Important elements falling outside the H.323 specification include:

- Audio and video functions and devices for Terminals such as voice activation, microphones, speakers, mixers, amplifiers, acoustic echo cancellation, cameras and monitors and controls for split screen video.
- Data applications, special devices and associated user interfaces that utilize T.120 or the data channel.
- Management of fixed resources, such as allocating the dynamic connections, called ports, of the MCUs and Gateways.
- The IP network(s) over which an H.323 system runs and all issues related to the design and provisioning of such networks.
- Directories related to finding, scheduling and calling people and conferences.
- Policies, procedures and best practices related to use and operation of conferencing resources and networks.

Further, how do the features of H.323 relate to creating a conferencing service that can be offered commercially by a service provider? Which features and functions are most important?

This section provides context for this discussion and introduces some important considerations for service providers who wish to create a commercial conferencing service.

5.1 Service Provisioning and Operations

The scope of H.323 does not extend to the provisioning, scheduling or administration of conferences and the generation of usage records for billing. Vendors of H.323 components address functions like these through “value added” extensions to MCUs and Gatekeepers. The web is a popular interface for accessing such extensions. Some vendors may offer an SDK or toolkit to allow customization or development of the extensions.

From a service provider’s prospective, efficiency of operation is an imperative. Therefore extensions that provide customer-provisioned conferencing and efficient administration and management of conference operations via a Web interface should be highly valued. (Customers of PSTN audio conferencing services are already requesting web-based provisioning and conference controls.)

For purposes of discussion consider the following four roles (levels of authority and privilege) in the provisioning, use, and operation of conferencing services:

- **Operator –** the service provider personnel who operate the conferencing service, manage its servers and hold responsibility for quality of service.
- **Administrator –** personnel who are authorized to incur charges and/or to schedule conferences, thereby reserving ports on MCUs and/or Gateways. For example, an administrator might be a designated representative of a customer to whom services are provided.

- **Conferee –** is a person who participates in a conference (a customer or an invited participant of the customer.)
- **Chairperson –** is a conferee who has additional conference-time privileges used to run a conference. Also, a chairperson may share certain administrator privileges, but only for chaired conferences.

Consider the following concepts and issues (presented as a series of questions that a customer needs to consider) when determining requirements for provisioning and operating a conferencing service:

5.1.1 Privacy and Security

What level of privacy do customers require? For example, will customers allow information about their conferences or conference participants to be seen by other, possibly anonymous customers? Within an organization, is there a need for conferences that are hidden from ordinary view? What levels of security are required for authentication, communications integrity and encryption, and non-repudiation?

5.1.2 Account and User Management

What information is needed to manage a customer account? How are accounts identified? For each account which services are provided? What is the name of the primary business contact? The root administrator? What are the contract terms and “bill to” information? And what are the defaults and restrictions to be applied in scheduling conferences.

What information is needed to identify and manage users? Should a user be affiliated with one or many accounts? What information is needed for identification and authentication? How are permissions granted to users, such as designating a user as an operator, administrator, conferee or chairperson. Who may create a new user? Who may modify, delete or view user information? Who is allowed to create and schedule conferences? May there be multiple administrators for one account? Do all administrators have the same permissions?

Is account and user management web provisioned?

5.1.3 Conference Templates

A template is a predefined description of a conference that is used to create an actual conference. Should templates be used as a shortcut mechanism to create conferences? If so, what information should be included in a template:

- The communications protocols (H.323, T.120) and media (audio, video, data) to be used for the conference?
- The resources to be allocated, for example, the MCUs or Gateways and number of ports to be reserved?
- The bandwidth management policy to be used?
- The Gatekeeper to be used?
- Is IP Multicast to be enabled?

- Is the conference is to be private, hidden or require a password?
- The identification of the designated chairperson for the conference?
- Is the management of templates web-provisioned?

5.1.4 Conference Scheduling and Management

How are actual conferences created from templates and who has the privilege to do so? Which template parameters may be overridden? How are conferences scheduled: one-time use in advance? On-demand? For a specific time and duration? Recurring on regular intervals, for example on same day of week, at the same time? Who may modify, delete or view scheduled conferences? Are the resources needed for a conference guaranteed to be available? How and when can conferees be invited to a meeting? How are they notified? How do they accept or reject an invitation? Is there an interface to external calendar or e-mail systems? Can a directory service, for example, an LDAP-enabled database or Microsoft Exchange, be used to locate or identify conferees? Is conference scheduling and management web provisioned?

5.1.5 Conference Time Controls

How do the invited conferees join a scheduled meeting? Are both meet-me (dial-in) and call-out modes supported? How are impromptu meetings convened? How are conferees authenticated? How is the scheduled duration of a meeting extended? Who has permission to do this? Is there a designated chairperson? What responsibilities and controls does the chairperson have? For example, can the chairperson mute an unruly speaker or drop a conferee from the conference? What controls are available to the individual conferees? For example, in a videoconference can a conferee choose to continuously view the video of another conferee? Are there provisions for side-meetings (meetings within meetings) or private chats? Is a list of the invited conferees available? Does it show those who have joined the meeting? How can a conferee drop out of a conference? Are there to be conference operators to assist with technical logistics? How are they summoned? Who may do so? Which conference-time controls are Web-provisioned? Which are DTMF-provisioned for conferees who participate via telephone?

5.1.6 Systems Operation and Management

What capabilities are provided for systems operators? What tools are provided to configure and manage the conferencing system: MCUs, Gatekeepers, Gateways? What real time monitoring functions are provided: a list of active conferences and number of conferees in each; a list of ports reserved and ports in use; a list of statistics on bandwidth and packet usage. Can an operator monitor a live conference? Under what conditions? What provisions are there for creating and processing customer data records for billing? What options exist for billing: by the port-hour, by the packet, by bandwidth usage? Are there non-reputable records of scheduled and active conferences? Are there records of conferee participation in conferences? Is there an SNMP interface? Are SNMP MIBs provided? What tools are available to monitor or measure network performance? What diagnostic tools are provided? Are systems operation and management functions accessible via the web with appropriate security procedures?

5.2 Evaluating MCUs

Commercial MCUs have different functional characteristics depending in part on the vendors' respective approaches to the marketplace. Consider the following criteria and features when evaluating MCUs for conferencing services.

What is the capacity of the MCU?

What is the maximum number of concurrent connections (ports) across all conferences? What is the maximum number of concurrent connections for a single conference? (Note: Terminal software also could limit the number of connections in a conference.) What is the number of concurrent conferences? Can connections for a conference be spread over multiple MCUs (cascaded)? Are controls provided to prevent overscheduling ports?

What methods for audio processing are supported?

Switching? Mixing? Both? What is the maximum number of concurrent mixed audio output streams? How does the quality of the mixed audio streams compare to traditional PSTN voice bridges? How much delay is introduced by audio processing? How is the size of the jitter (input) buffer determined? Is audio processing handled by hardware or software? How does this affect performance and cost?

Which video processing methods are supported?

Voice activated switching? Continuous presence? Video mixing? Which video formats are available? How much delay is introduced by video processing? What is the aggregate video throughput supported? Is audio-only participation in a videoconference supported? How?

What are the dependencies on a Gateway for connecting to telephones via PSTN or a PBX?

Does the MCU support an integrated connection to the PSTN, for example, via a T1 or E1 interface? Can DTMF signaling and interactive voice response be used to support telephone participants?

Which audio and video codecs are supported?

Are the codecs compatible with the codecs in the Gateways you plan to support? (IP Telephony Gateways might not support video.) For video, across what frame sizes and across what frame and bit rates do the codecs operate? What is the quality of the lip synchronization?

Is T.120 data conferencing supported?

Is it supported within H.323 (H.323 is signaled first)? Are T.120 data conferences also supported independent of H.323? (Signaling is also dependent on Terminal software.)

Is IP Multicast supported?

IP Multicast must also be enabled on the network.

Does the MCU include an integrated Gatekeeper?

This should not preclude use of an external Gatekeeper. What requirements does the MCU have of an external Gatekeeper? Are the MCU and Gatekeeper able to work together to provide address translation? Is admission controlled via an authentication policy? Is bandwidth enforced via policy?

Does the MCU support netcasting, recording and playback on demand for conferences?

Does the MCU integrate with Microsoft's NetShow or RealNetwork's RealServer. Can conferences be recorded, stored and made available for playback?

Can the MCU be easily deployed and operated?

Which conference management features are provided out-of-the-box? Is there an effective method for managing MCU port resources? Can auto-provisioned services easily be created and managed? Are services and operation functions provisioned through a Web interface? Is there an SDK or API through which value added front end services can be developed and customized?

What is the extent of interoperability testing?

Does the MCU interoperate with Terminals to be supported by the service?

What security features are provided?

What methods are supported for authentication, data privacy, data integrity and non-repudiation?

Vendors of MCUs and their delivery partners have developed a considerable range of experience and expertise in the implementation of conferencing systems and services. Working with a vendor that can support the creation and deployment of your services should be a strong consideration in selecting an MCU vendor.

What are the price of the MCU and the price per port?

5.3 Networks

Network performance is a critical factor when deploying audio or video communications on IP networks. QoS parameters for bandwidth, packet delay, jitter and loss need to be carefully considered. It is useful to understand how much delay can be introduced by passing through devices on the network, for example, routers, proxy servers, firewalls, Gateways, and of course, MCUs.

The location of MCUs and Gateways on the network is important. Proper positioning can positively impact bandwidth utilization and packet delay. The location of MCUs and Gateways should be determined through careful network design.

5.4 Terminal Equipment

The multimedia communications experience is delivered to the user through his or her terminal equipment. Gaining direct experience with the terminal equipment that is anticipated to be commonly encountered or directly supported by your conferencing service is a worthwhile exercise on the road to providing superior support services.

A desktop or laptop PC can be the foundation of a meeting room system or individual conferencing system. Room-based systems, set-top systems and communications appliances are pre-integrated packages. The following discussion focuses on personal computer based terminal equipment.

5.4.1 Audio Conferencing

A full duplex sound card is needed to speak and hear at the same time. A headset with a boom microphone is recommended to avoid echo and environmental feedback. Look for a mute switch with the microphone to eliminate unnecessary noise when listening to a presentation. In a meeting room, or if a headset is unacceptable, use a microphone/speaker combination with built-in echo cancellation. IP phones are also a great option.

5.4.2 Video Conferencing

A video camera is the critical element. A video capture card with on-board codec(s) is recommended to reduce load on the processor. Make sure the camera, capture card and codecs work well as a unit and interoperate with terminal software. Curtail use of other applications when using a PC for video conferencing. Users should avoid running a heavy load of network or compute bound applications on their PC while conferencing. Background applications can interfere with video or audio quality by hogging the CPU or causing congestion at the network stack and interface. (Future QoS support for PCs should alleviate this problem.)

Cameras for a meeting room need to handle variable distances to their subjects and a range of lighting conditions. There are cameras for meeting rooms that automatically find and focus on the speaker. Large screen monitors are most often used in meeting rooms. LCDs must be qualified to ensure display characteristics support video to the desired level of quality.

5.4.3 Data Conferencing

At a minimum, conferencing without audio and video requires only a PC with keyboard and mouse. However, if the whiteboard is an often used tool, a drawing device with greater speed and control than the mouse could be a welcome addition.

A large monitor or LCD projector is an appropriate viewing device for a meeting room. Compatibility problems between PC display formats and interfaces and LCD capabilities are not uncommon. In NetMeeting the shared application's window is transmitted to the other participants in a conference. Clipping results if the display resolution of a recipient's PC is less than the sender's.

5.4.4 Software

The conferencing software running on the PC is the final component. These programs recommend a minimum PC configuration. Microsoft NetMeeting is the most widely available IP-based conferencing tool, but there are others worth considering, especially if video conferencing is the goal.

Integration of software, PCs and peripherals is an important topic. Develop a set of "standard" or "known working" configurations for both desktop and meeting room environments. Include set-tops and appliances. This should make conferencing services easier to sell and support.

5.5 Making It Work: Human Factors, Services and Training

A conferencing service *must* be dependable and easy to use. Audio conferencing service bureaus know that the last thing a business customer needs is for a meeting to fail because of a technical glitch. A presenter who is intently focused on positioning a new product strategy at an internal review finds only aggravation in intricate procedures for operating the conferencing system.

Audio conferencing service bureaus understand that great service drives repeat business and customer loyalty. Great service often wins against lower priced competition.

Successful adoption of conferencing depends on technical and human factors. Conferencing offers new tools for working together over distance. To be an effective tool, it must be widely deployed and accepted. In the process certain user behaviors must be carefully reshaped.

Success is a win-win situation for you and your customer. While your customer may have the desire to succeed they might not have the necessary knowledge or skills. This requires training.

Recommendation: Help customers assess needs and create benchmarks for success

Help map your customer's needs to the benefits of your service to establish value. Then provide a benchmark so that your customer can measure success in the context of his business operation. For an example visit the ROI Calculator[37] at www.intel.com. Help the customer identify changes in process that will need to occur. Take this opportunity to establish a close relationship with the champions of your services and other key people within the customer's operation. This exercise will also allow you to see how your conferencing service could grow within your customers' operations and assist in planning infrastructure build-out.

Recommendation: Establish a blueprint for a conferencing awareness campaign

Business customers need tools to jump start the usage of your conferencing service within their user community. Potential users need to become aware of the benefits of using conferencing services and how to take the first steps. Consider developing some or all of the following collateral for use in a co-branded promotional campaign: e-mail notices, FAQ sheets, quick-start guides, posters, online open houses, online help, testimonials and case studies -- preferably from the same business customer who will run the promotion.

Recommendation: Don't force unnecessary changes in behavior

Conferencing from the desktop has not yet displaced the traditional meeting room. Whether it's a social fact of life or a temporary discomfort with new technology, many people still prefer to meet together in rooms whenever possible. Therefore conferencing should be provisioned everywhere meetings or conferencing takes place: meeting rooms, desktops, home offices, and "on the road". Customers should avoid forcing their users to change meeting rooms to conference.

Recommendation: Use training supplemented by online help to reshape behaviors

Conferencing requires modification to user behaviors. For example, in the case of data conferencing, presentations do not need to be e-mailed or faxed, received and processed in advance of a meeting. A presenter can make changes until meeting time. If the presentation is not predistributed, the presenter can engage the audience on his or her terms.

[37] http://www.intel.com/proshare/conferencing/roi/index.htm .

The audience does not have the option to rummage through the materials and to get ahead or behind the speaker's train of thought. Remote participants do not have to guess which slide is on screen in the meeting room. The presenter can use the universally visible cursor (from the shared application) to call attention to details on a slide or in a spreadsheet. This, however, requires that the conferencing device be located at the podium or that the presenter sits down with it, a somewhat unnatural act for a seasoned speaker.

Ensure that instructor-led training is provided in best practices and in the operation of the terminal equipment. No matter how simple the conferencing application or changes in process might seem, do not assume that online help alone will be sufficient. A new participant in conferencing may have a positive attitude toward conferencing but lacks the time or technical enthusiasm to self-train. And it is difficult to conference by yourself! Here are some reasons given for not using existing videoconferencing systems: "I never used it before." "I'd have to go down the hall." "It's not simple enough." "I'm afraid I won't be able to make it work." "It'll disrupt my meeting." Instructor-led training can be provided via the "train the trainer" format or as a pay-for service. A meeting skills refresher is something many companies can use and represents added value. Train for the differing levels of skill: administrators, conferees, chairpersons, presenters and technical support.

The primary goal of a meeting service should be more effective and productive meetings.

6 Checklist for Service Creation

Considerations in planning a multipoint conferencing service:

- Identify the target users of the service. Identify the requirements and compelling reasons for them to use the service. If video conferencing is offered, will H.323 and H.320 be supported? Identify the critical characteristics of a service that meets these needs.
- Create a business plan. Will the service be a general one-to-many service or will it be hosted using dedicated equipment for specific customers. If the service is to be billed, how will the service be metered and billed? Determine the cost and scale of implementation. Determine marketing costs. How will the service grow over time? Determine metrics for success.
- Create a service description. Determine which communications model(s) will be used. Define features and functions to be provided to end users. Scheduling, security and the quality of audio are key features. Define features and functions required for internal support. Define an internal operations model. Define what roles your customers and your operations staff will play in provisioning, administration and operations. What is the role of the Web in provisioning, operations and support?
- Create a network plan. Based on the communications model and features and functions of the service, what are the network requirements? How will the network be provisioned? What is the cost? What are the network security issues? How will they be resolved? What are the network quality of service issues? How will they be resolved? How will the service scale across the network?
- Create a service deployment plan. What components are to be used: MCUs? Gatekeepers? Gateways? What terminals are to be supported? Which vendors and products are to be used? Cost? What development or customization work is needed (after installation of standard products)? Will development and/or customization work be done in house? By the vendor? A third party? Cost? How will the components scale as usage increases?
- Create an operations plan. Who will be responsible for the ongoing internal operation of the service? What new processes and procedures need to be established? Which current ones require modification? How does this impact head count?
- Create a customer service and training plan. What will a customer need to do to realize the full benefits of your service? What should a customer need to do (internally) to roll out your service? What is your role in this? What services and training will be provided to a customer? Under what terms and conditions?
- Create a marketing plan. How will you reach target customers, communicate the value of your service and sell your service?

7 Summary and Recommendations

Multipoint conferencing is a new opportunity for service providers who wish to offer a full line of communication services to businesses or to bring multimedia chat capabilities to online communities of interest.

The incentives exist for businesses to utilize new communications services to improve their operational capabilities and efficiencies. A successful multipoint conferencing service requires an approach that integrates business, technical and human factors. The service must be easy to use and reliable. The quality of the delivered media must be sufficient for its intended purposes and comparable to alternative technologies. Startup services can provide valuable assistance to customers in the integration of new communication tools with business processes, so that the full benefits of the conferencing service can be realized.

Audio conferencing over PSTN is a mainstream business communications tool in North America. With IP Telephony poised to become widely adopted there is a new opportunity to provide IP audio conferencing converged with data applications. Video conferencing has yet to achieve the same level of acceptance as audio conferencing. Video is also the most demanding of the three conferencing media.

There are four models for combining audio, video, and data conferencing over IP networks and PSTN networks using a standards-based approach:

1. Audio over PSTN and data over IP
2. Audio over IP and PSTN
3. Audio over IP and PSTN and data over IP
4. Audio over IP and PSTN, data over IP and video over IP.

Consider the interoperability, integration and ergonomics of the terminal equipment to be used. The benefits of conferencing are delivered through this equipment. Users want tools that are reliable and easy to use. They have little patience for those that are not. Left to chance, the potential for annoying or disruptive incompatibilities is great. It just has to work!

Leverage the simplicity and ubiquity of the browser to drive adoption of your conferencing services. Provide Web access to services, including conference scheduling and operation.

Social communities of interest are a fertile ground for innovation and experimentation with multimedia group conferencing. This audience has been weaned on text chat, is generally young and technically adept, and is more tolerant of imperfection than the business audience. Multimedia chat might be positioned to differentiate and enhance market share or to enable a professional community of interest.

The most significant technical inhibitors to the development of multipoint conferencing services over IP networks are the following:

- Lack of available bandwidth, especially for video
- Lack of QoS controls for audio and video, especially across network domains
- The traversal of network security devices like firewalls.

These issues are being actively worked in the ITU-T, the IETF and by the networking industry. Solutions leading to QoS guarantees should be available for the enterprise networks and via overlay IP networks starting as early as 2000. But anticipate that it will be at least several years before inter-domain quality of service is widely available on the Internet.

H.323 is a significant and evolving standard that is now being driven in large part by the adoption of IP Telephony. Watch it closely to ensure that H.323 continues to embrace and resolve issues introduced by competing initiatives.

Partner with leading suppliers of H.323 components. Choose suppliers whose products are standards compliant, who participate in IMTC interoperability testing, who have a proven track record with service providers, and who can apply their planning and deployment expertise in the creation of your conferencing services. Systems test your service before deploying widely.

A. The Evolution of Conferencing Standards

Standards are critical to open communications systems. Standards drive the interoperability of devices supplied by different vendors. We enjoy the benefits of standards every time we turn on the radio, pick up a telephone, play a CD or access the Internet.

The most significant and comprehensive multimedia teleconferencing standards are the H.32X and T.120 series. These Recommendations have been developed by the Telecommunication Sector of the International Telecommunication Union (ITU-T), an agency of the United Nations, formerly CCITT. Its goal is the worldwide standardization of telecommunications.

The following table is repeated from Chapter 4 for ease of reference.

Table 12. ITU-T Multimedia Teleconferencing Recommendations

Recommendation	Applicable Networks	First Ratified	Last Revised	Media
H.320	ISDN, circuit-switched	1990	May 1999	Audio and video
T.120	All listed here	1996	May 1999	Data
H.323	IP, LAN, packet-switched	1996	May 1999	Audio and video
H.324	POTS, V.34 modem, low bit rate	1996	September 1998	Audio and video

Each recommendation cited in Table 12 is the umbrella standard of a series which, in aggregate, comprises the complete standard. The ITU-T does not require a working implementation before ratifying a Recommendation. Its approach to standards development can be characterized as top-down, evolutionary and comprehensive. Details of implementation are often left to the first users.

The ITU-T is organized in part by Study Group. Study Group 16 holds the ITU-T charter for multimedia services and systems. At its recent meeting in May 1999 it approved revisions to the umbrella standards. The May 1999 revision of H.323 is called H.323 version 3.

The International Multimedia Teleconferencing Consortium, Inc. (IMTC) focuses on the stabilization of multimedia teleconferencing standards through the promotion and facilitation of interoperability testing. It is a non-profit consortium of over 150 international members, many of whom develop products or services based on H.32x or T.120.

Since 1996, the IMTC has held over 40 interoperability test events worldwide. An event runs for a week or more and includes structured testing of released and unreleased products from multiple vendors in systems configurations under live conditions. The results of testing are not disclosed publicly in order to achieve maximum cooperation from the vendors who are often competitors. Testing results in comments that are fed back to ITU-T Study Groups for further refinement of the Recommendations. The need for and value of the IMTC's interoperability work is clear.

The original development of H.320 was initiated by the teleconferencing industry within the (then) CCITT with the goal of achieving cross-vendor interoperability among videoconferencing systems. Realizing later that ISDN would not be widely deployed to corporate desktops, work began on a LAN-based specification, H.323, intended to interoperate with H.320 and to bring videoconferencing to the volume desktop market. Intel and Microsoft lent their support to H.323. Microsoft developed its now widely deployed NetMeeting client.

From 1995 until 1997 IP Telephony was in a very formative state. As plans for commercial IP Telephony were laid, the participants, with support from the IMTC's Voice over IP Forum, discovered that H.323 was a real and usable protocol that met initial needs. By 1998, H.323 had become the de facto standard for IP telephony.

Another important standards body is the Internet Engineering Task Force (IETF). In contrast to the ITU-T, standards development at the IETF focuses on specific problems for which solutions can be demonstrated by running code. The IETF is a loosely federated technical community whose work is accessible via the Web.

Within the last year IP Telephony has become the driving force behind voice and data network convergence. Its proponents have assumed prominent roles in the development of multimedia communications standards within the ITU-T and the IETF.

The first implementations of IP Telephony using H.323 produced a list of issues. Many are rooted in (a) large scale deployment across global IP networks through multiple service providers; and (b) interoperability with the PSTN, especially its intelligent network (IN) and advanced intelligent network (AIN) functions like Signaling System 7 (SS7).

As a result of these experiences and the high degree of innovation and competition in the market, standards initiatives sprang up within the IETF that are to a degree competitive with H.323. Media Gateway Control Protocol (MGCP, work in process)[38] and Session Initiation Protocol (SIP, RFC 2543)[39] are significant efforts worth watching closely. Both initiatives include support for multipoint conferencing.

MGCP targets higher interoperability with the PSTN and more closely models the centralized call control features of the AIN, including support for metered billing. It is a combination of two earlier proposals, SGCP and IPDC. Although only in draft, MGCP has garnered strong support from major carriers, communications equipment vendors and network service providers, making it likely to become an important protocol.

In response, the ITU announced in December 1998 that work had begun in Study Group 16 on a new Recommendation called H.gcp, intending to resolve differences between MGCP and H.323. In June 1999 after collaboration with the IETF's Megaco Working Group, the ITU announced[40] the definition and first stage approval of H.gcp. Work on H.gcp is to continue with final approval anticipated at Study Group 16's next plenary meeting in February 2000. Collaborative work is to continue with the IETF. The Megaco Working Group is scheduled to complete its HGCP RFCs in July 1999. The IETF plans to comment on H.gcp in the fall, based on experience with running code.

Work on SIP (RFC 2543) dates to 1996. SIP competes more directly with H.323 than does MGCP. SIP and its sibling protocols Session Description Protocol (SDP, RFC 2327) and Session Announcement Protocol (SAP, work in process) provide a lighter weight, layered, less complex protocol alternative to H.323 for signaling, describing and announcing multiparty multimedia conferences. SIP is HTTP-like, transport protocol neutral, and proxy friendly. It is integrated with native Internet protocols and the Web. It provides explicit extensibility. Despite its tenure, it has not yet been widely adopted to date. But significant interest in SIP is building.

H.323 is real today and is the most widely deployed standard for IP Telephony and multipoint conferencing. It provides the most robust environment for connecting heterogeneous, media-rich terminals over IP. To keep pace with the needs and interests that are now driving the development of IP networks and network resident applications, its evolution must continue at Internet speed.

---

[38] MGCP is a protocol drafted by the Media Gateway Control (megaco) Working Group within the IETF Transport Area see http://www.ietf.org/html.charters/megaco-charter.html .

[39] SIP is part of a suite of protocols developed by the Multiparty Multimedia Session Control (mmusic) Working Group within the IETF Transport Area, see http://www.ietf.org/html.charters/mmusic-charter.html .

[40] http://www.itu.int/plweb-cgi/fastweb?getdoc+view1+www+43218+0++H.gcp

B: Resources

These resources provide additional information on multipoint conferencing, its markets, technologies, standards, products, suppliers, applications and best practices.

Table 13. Multipoint Conferencing Resources

Type	Title / Name	Source / Reference
Conference	MultiMediaCom	Http://www.bcr.com/confer/mmcom99f/Default.htm
Conference	TeleCon	Http://www.teleconexpos.com/
Market Research	Forward Concepts	Http://www.fwdconcepts.com/
Market Research	Frost & Sullivan	Http://www.frost.com/
Market Research	Telespan Publishing Corporation	Http://www.telespan.com/
Presentation	Fundamentals of Videoconferencing and Multimedia Communications	International Multimedia Teleconferencing Consortium ftp://ftp.imtc-files.org/imtc-site/Multimediacom/imtcu.zip
Standards	Communications Standards Review (Standards News and Reports)	Http://www.csrstds.com/ http://www.csrstds.com/http.html
Standards	European Telecommunications Standards Institute	Http://www.etsi.org/
Standards	International Telecommunication Union	Http://www.itu.int/
Standards	Internet Engineering Task Force	Http://ietf.org/
Technical Guide	Microsoft NetMeeting Resource Kit	Http://www.microsoft.com/windows/NetMeeting/Authors/
Web Site	Internet Conferencing	Http://netconference.miningco.com/
Web Site	Meeting by Wire	Http://www.meetingbywire.com/
Web Site	The Windows NetMeeting Zone	Http://netmeet.net/
Web Site	Video Conferencing for Business - H.323 Overview	Intel Corporation http://www.intel.com/proshare/conferencing/h323/index.htm
White Paper	Demystifying Multimedia Conferencing Over the Internet Using the H.323 Set of Standards	James Toga and Hani ElGebaly, Intel Corporation http://developer.intel.com/technology/itj/q21998/articles/art_4.htm
White Paper	Internet Bandwidth Management: iBAND2^{SM} Edition	Stardust Forums, Bob Quinn http://www.stardust.com/iband2/whitepaper.htm
White Paper	Meetings in America	MCI WorldCom http://nmc.mci.com/MeetAmerica/mtg_usa_wp.shtml
White Paper	Meetings in the UK	MCI WorldCom http://nmc.mci.com/meetuk/ukwhitepaper.html