Computer Networks 31 (1999) 225–235

Multimedia conferencing over ISDN and IP networks using ITU-T H-series recommendations: architecture, control and coordination

Mark Reid *

VideoServer Inc., 63 Third Avenue, Burlington, MA 01803, USA

Abstract

This article discusses the architecture, control protocols, and coordination mechanisms used in today's multimedia conferencing standards for ISDN (Integrated Services Data Network) and IP (Internet Protocol) networks. Specifically, the ITU-T's (International Telecommunication Union – Telecommunication Standardization Sector) H.320 and H.323 multimedia conferencing standards are discussed and compared with respect to their architecture and position in the market. © 1999 Elsevier Science B.V. All rights reserved.

Keywords: Multimedia; Conferencing

1. Introduction

This article discusses the architecture, control protocols, and coordination mechanisms used in today's multimedia conferencing standards for ISDN (Integrated Services Data Network) and IP (Internet Protocol) networks. Multimedia conferencing refers to meetings at a distance that involve some combination of audio, video, and data collaboration. When these meetings contain all three media types they are called videoconferences and are conducted using either the ITU-T's (International Telecommunication Union – Telecommunication Standardization Sector) H.320 multimedia conferencing standard for ISDN or H.323, the multimedia conferencing standard for IP Networks [1–23].

2. Multimedia conferencing architecture

All of today's ITU-T standard-based multimedia conferencing protocols are built from a similar base architecture. Each standard shares the same basic components that are organized to deliver control, audio, video, and data collaboration to the user. The differences between the standards are driven from each standards optimal packaging of data for transmission on the network for which they were designed to service.

Multimedia conferencing equipment can be categorized into terminals or end-user devices that are used to participate in a conference and network

* Tel.: +1-781-505-2368; fax: +1-781-505-2101; e-mail: mreid@videoserver.com.

1389-1286/99/$ - see front matter
PII: S0169-7552(98)00264-5

devices that are used to provide services to the terminals. Network devices provide services like multipoint, gatewaying or protocol conversion, transcoding, rate matching, call forwarding and call transfer.

Fig. 1 shows the design of a generic multimedia terminal that aims to deliver control services and collaboration tools for audio, video, and data.

The generic multimedia terminal architecture includes the following basic components:

User interface – The user interface provides the ability to place and terminate calls and configure and monitor the system.

Audio I/O equipment – Audio input output equipment includes devices such as microphones, speakers, headsets, and speakerphones.

Video I/O equipment – Video input output equipment includes devices such as cameras, TVs, computer monitors, and VCRs.

Data applications – Data applications include applications such as electronic white boards, file transfer, and application sharing.

System interface – The system interface includes the program display, keyboard, keypad, or other input output device that is used to control or monitor the system.

Call control – Call control provides call setup and tear down processing, and supplementary services such as call transfer and forward.

Connection control – Connection control provides control for a call once a connection is established. Connection control includes services like capability negotiation, the maintenance of the media channels, and mode selection.

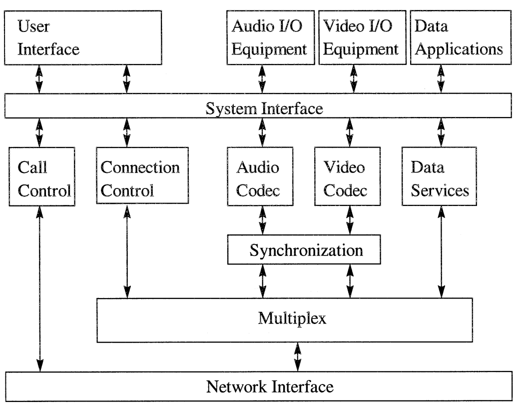

Fig. 1. Generic multimedia terminal architecture.

Audio codec – The audio codec performs the coding (compression) and decoding (decompression) of audio signals.

Video codec – The video codec performs the coding (compression) and decoding (decompression) of video signals.

Data services – Data services include the protocols and services used to transport data generated by applications like white board and file transfer between endpoints in a conference. Data services include services and protocols such as a reliable transport protocol, a one to many data distribution service, or a multipoint (many to many) data distribution service.

Synchronization – In general, video codecs require processing delays that exceed the delays associated with audio codecs. To maintain synchronization, audio must be delayed in the system to match the video.

Multiplex – The multiplex combines control, audio, video, and data signals into a stream of information that can be transmitted as a single bit stream or as a series of packets. The multiplex also parses the received bit stream or received packets and re-constitutes the control, audio, video, and data signals.

Network interface – The network interface makes the necessary adaptation between the actual network and the lowest level components of the multimedia architecture.

3. ISDN based conferencing architecture

The ISDN based multimedia conferencing market has been deployed using the ITU-T's H.320 conferencing standard. H.320 became a standard in 1990 and is used as the default mode of operation by every major manufacturer of ISDN based videoconferencing equipment. The protocol is built on a bit level multiplex that can use multiple ISDN B-channels (64 Kbps) to send control, audio, video, and data information. The multiplex, which formats data in a synchronous stream where fields are distinguished by bit position, does not require packet headers. The lack of packet headers makes the multiplex very efficient in its use of bandwidth but requires a complex implementation and a significant amount of processing power to operate in real-time.

3.1. H.320 components

H.320 consists of terminals and multipoint control units called MCUs. The terminal is the end-user device that is used to participate in a point-to-point or multipoint conference. A MCU is a network device that enables a terminal to participate in a multipoint conference. Fig. 2 shows the relationship of the terminal and the MCU for H.320.

H.320 media streams follow the ISDN circuits that are established during call setup. For a multipoint conference, all terminal connections (e.g. control and media) are made between the terminal and the MCU. It is possible to cascade multiple MCUs together in a single conference, but even in this case, all of the terminals are connected directly to one of the MCUs.

3.1.1. H.320 terminal

Fig. 3 shows how the generic multimedia terminal architecture has been adapted for H.320 and ISDN.

A H.320 terminal includes the following components:

Call control – H.320 uses the ISDN call signalling protocol Q.931 for call setup and tear down. The call setup protocol runs in the D-Channel of the ISDN connection. Q.931 is signalled by the terminal to the network, is not within the scope of the H.320 standard, and does not run through the H.221 multiplex.

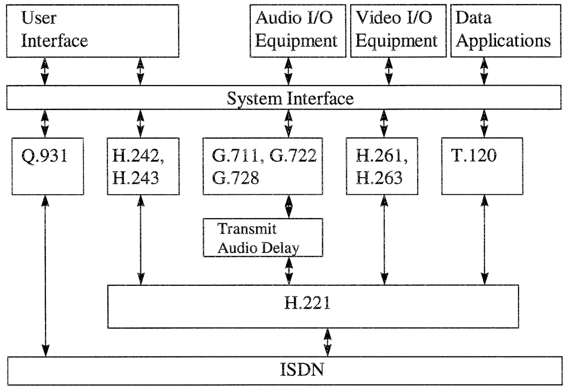

Fig. 3. H.320 terminal architecture.

Connection control – Connection control is carried out by the H.242 and H.243 protocols. H.242 describes the procedures for endpoint capability exchange, restricted network operation, mode switching, and mode selection for the various audio, video, and data modes in a point-to-point call. H.243 describes the procedures required for multipoint conferences in which multiple participants are joined together via a multipoint control unit (MCU). H.242 and H.243 use command and control indications called BAS codes (bit-rate allocation symbol) that are defined in the H.230 specification. A bit-rate allocation symbol is a code word that is a command, indication, or definition of a terminal's capability to structure the capacity of the channel.

Audio codec – Although H.320 specifies G.711 as its base audio codec, most 128 Kbps conferences (Basic Rate ISDN) use G.728 for their audio mode

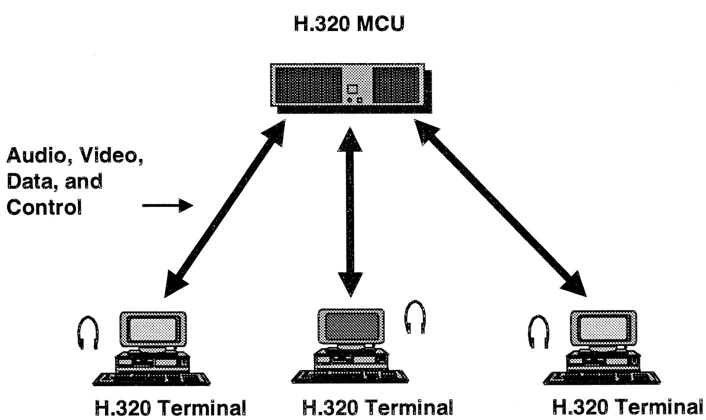

Fig. 2. H.320 components.

and most 384 Kbps (Primary Rate ISDN) conferences use G.722. Table 1 shows the details of the audio algorithms specified for use in H.320.

Video codec – H.320 specifies the H.261 video codec as mandatory and the H.263 video codec as optional. Support for the H.263+ video codec will be added to H.320. H.261, H.263 and H.263+ describe a set of signals that can be transmitted to convey a compressed moving picture. The standards only specify how to transmit a compressed picture, not how to encode it, and not how to take the compressed picture and render it on a display. The generality in the video standards has allowed manufacturers who have better compression technology to distinguish their video quality from their competitors when comparing video quality on links of the same bit rate. When viewing a video image, it is important to remember that for the most part, the video quality is the result of the remote system's encoder and not the result of the manufacturer's system where the image is being viewed. The specifications define video formats at various resolutions with the most common being quarter CIF (QCIF or Quarter Common Intermediate Format) 178 by 144 or full CIF (CIF or Common Intermediate Format) 356 by 288.

Data applications – Standard data applications defined for use with H.320 come from the T.120 protocol suite. These applications include T.126 for still image exchange and white boarding, T.127 for file transfer, and T.128 for application sharing.

Data services – Data services in H.320 are provided by the T.120 protocol suite. The data services include generic conference control (GCC) called T.124 and multipoint communication service (MCC) called T.122. The multipoint communication service layer protocol is defined in T.125 and the transport protocol stack definitions are specified in T.123. T.123 is called the Network Specific Transport Protocols specification and defines the protocols for running over networks like ISDN, the public switched telephone network (PSTN), and the required protocol layers to utilize networks that transport TCP/IP.

Synchronization – (Transmit Audio Delay) In H.320, the audio is delayed in the transmitter to compensate for the video encode path delay to allow the receiver to maintain lip synchronization. Although the addition of audio delay for synchronization is optional, most vendors implement this feature due to customer's preference for synchronized audio and video. In H.324, the ITU-T standard for multimedia conferencing on the public switched telephone network, instead of the audio being delayed in the transmitter; transmitters send a skew indication that contains the average skew by which the video signal trails the audio signal. In this case, the receiver optionally delays the audio to match the inbound video.

Multiplex – H.221 provides the H.320 multiplex capability by dynamically subdividing the overall channel on a bit by bit basis for control, audio, video, and data. The B-channels are divided into octets in which each bit position in the octet is actually a different sub-channel (e.g. in any given octet there maybe bits from each of the control, audio, video, and data streams). The complexity of the bit level multiplex is the reason that H.320 is typically implemented with a hardware add-on card.

3.1.2. H.320 MCU

H.320 defines a Multipoint Control Unit (MCU) that can host multipoint conferences on an ISDN network. A MCU provides rate matching, audio mixing, video switching, continuous presence video mixing processing, multipoint data services, and spe-

Table 1
H.320 audio algorithms

Algorithm	Audio bandwidth	Type of algorithm	Bit rate (Kbps)	Frame duration or sample rate
G.711	Narrowband (3 kHz)	PCM (Pulse Code Modulation)	48, 56, or 64	8 kHz sample rate
G.722	Wideband (7 kHz)	Subband ADPCM (Adaptive Differential PCM)	48, 56, or 64	8 kHz sample rate
G.728	Narrowband (3 kHz)	LD-CELP (Low Delay Code Excited Linear Prediction)	16	2.5 ms frame

cial conferencing services like chair control for the participants of a multipoint call. To conduct a multipoint conference, a MCU performs the capability negotiations between its attached endpoints to achieve a common operating mode and controls conference resources such as who is being heard and who is being seen.

MCU functionality can be broken into three main components: control processing, media processing, and media distribution.

Control processing – A MCU is responsible for the connection control processing associated with each terminal connection in a conference. To manage each connection, a MCU negotiates the audio, video, and data algorithm to be used by each terminal in a conference. A MCU uses its algorithm capabilities and its available resources to determine which mode each conference participate will operate in. A MCU may transcode between the various audio, video, and data algorithms, perform rate matching processing for terminals connected at different transmission rates, or perform video mixing to deliver a continuous presence video image to a terminal. The process of negotiating a capability set that works for each terminal in a conference is called ‘‘scuming’’ by the H.320 development community. The term is shorthand for ‘‘Selected Communication Mode’’ (SCM) which represents the mode selected for a conference.

Media processing – A MCU processes the audio, video, and data received from each terminal in a conference based on the configuration defined for that conference. For audio, a MCU may mix any number of audio streams to allow for seamless interaction between the conference participants and transcode between different audio algorithms to account for different terminal capabilities. To prevent echo at a terminal where someone is speaking, a MCU may create specialized audio mixes that do not contain the speaker's voice for each speaker in a conference. For video, a MCU is typically ‘‘switching video’’ to conference participants based on voice activation. In this mode, the video associated with the loudest speaker is sent to the terminals in a conference. As with audio, a MCU may be transcoding between video algorithms to account for different terminal capabilities or transmission rates. The MCU may also mix video to deliver a continuous presence image that contains the video from more than once source. In this mode, a MCU takes a number of video streams and mixes them together into a single picture. For data, a MCU performs the function of a T.120 top provider.

Media distribution – H.320 media is restricted to follow the path and bandwidth restrictions setup for the ISDN call. A MCU distributes individual audio, video, data, and control streams to each terminal in a conference. Fig. 2 shows the media flow for a H.320 multipoint conference.

4. H.323 based conferencing architecture

IP based videoconferencing has grown steadily over the past four to five years as multiple vendors have been piloting proprietary systems on corporate Intranets and the Internet. In 1995, a set of those vendors realized that a standard was necessary for the market to expand and started work on H.323. H.323 was ratified by the ITU-T in May of 1996 and is now the basis for IP conferencing applications ranging from corporate networks to the Internet.

The standard was designed with two goals in mind, first, to provide a mechanism for interactive multimedia communications on IP (packet) based networks and second, to define interoperability mechanisms between packet based H.323 endpoints and other H-series standards and traditional telephony. H.323 endpoints can operate over any network that supports the transmission of IP traffic.

H.323 includes all of the audio and video algorithms defined for H.320 and H.324 but only mandates G.711 audio and H.261 video. H.323 is built on top of the Internet Engineering Task Force's (IETF) Real-time Transport Protocol (RTP) [24] and Real-time Transport Control Protocol (RTCP) specifications that define audio and video transport and control. H.323 call setup is based on an extended Q.931, and its capability exchange mechanisms on H.245. H.323 includes the T.120 series standards for data collaboration and has provisions to allow T.120 terminals to participate in the data portion of a H.323 conference.

H.323 includes point to point and multipoint conferences where multipoint conferences have centralized control and either centralized or distributed audio and video. Distributed audio or video typically implies the use of IP multicast to distribute the media rather than a single unicast audio or video stream. Data in a multipoint conference is specified using the T.120 series standards.

In addition to a terminal, H.323 defines a Gatekeeper for call management, a Gateway for interoperability with other multimedia standards, a Multipoint Controller (MC) representing the control portion of a Multipoint Control Unit (MCU), and a Multipoint Processor (MP) representing the media processing portion of a MCU.

The gatekeeper controls access to the network for terminals, gateways, and MCUs and provides address translation for all H.323 components. A H.323 Gateway provides real-time, two-way communications between H.323 terminals on the packet-based network and other ITU-T terminals or telephones on the PSTN or ISDN. A Multipoint Controller (MC) provides the control services for three or more endpoints participating in a multipoint conference. A Multipoint Processor (MP) provides the media services (e.g. mixing, switching) for three or more endpoints participating in a multipoint conference. Together, the MC and the MP make up a Multipoint Conferencing Unit or MCU.

Fig. 4 shows the H.323 components in a typical network deployment. In this figure, switched circuit network is abbreviated SCN.

4.1. H.323 components

4.1.1. H.323 terminal

Fig. 5 shows how the generic multimedia terminal architecture has been adapted for H.323 and IP Networks.

A H.323 terminal includes the following components:

Call control – H.323 uses an extended Q.931 for its call setup and tear down protocol that was defined specifically for IP network operation. The extended protocol is defined in H.225.0 and uses Q.931 user-to-user information elements to transfer IP specific information like the H.323 dialling alias. H.323 also contains a registration, admission and status (RAS) protocol that defines the procedures and commands for endpoint to Gatekeeper communications.

Connection control – Connection control in H.323 is done by the H.245 protocol. H.245 specifies the commands and procedures for endpoint capability exchange, master slave negotiation, mode switching and mode selection for the various audio, video, and data modes. H.245 uses logical channels to specify what information is being transmitted and contains commands equivalent to those defined in H.242 and H.243.

H.245 is more flexible than H.242 or H.243 in that it has advanced capability structures and proce-

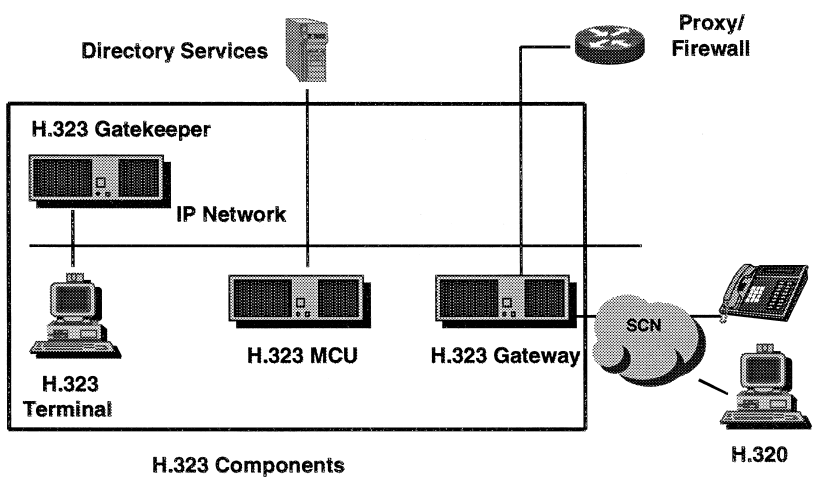

Fig. 4. H.323 components.

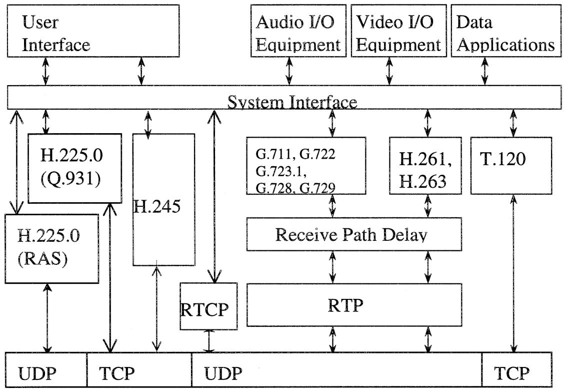

Fig. 5. H.323 endpoint architecture.

dures for signalling multiple media transmission modes. H.245 defines the following services:

- capability definitions with dependencies,
- the ability to request the transmission of a particular audio, video or data mode,
- the ability to manage logical channels,
- the ability to establish a master terminal for the purpose of managing logical channels,
- the ability to control the bit rate of individual logical channels,
- the ability to measure the round trip delay between endpoints.

Audio codec – H.323 includes the same audio algorithms as H.320 and specifies G.711 as mandatory. H.323 includes the low bandwidth audio codecs G.723.1 and G.729 for Internet Telephony applications and dial-up links. H.323 specifies how to packetize audio streams based on a whole number of samples or frames depending on the algorithm and allows terminals to stop sending audio when no information (e.g. silence) is detected at the microphone input. Table 2 shows the H.323 audio algorithms that are supported in addition to those listed in Table 1:

Video codec – H.323 uses the same video algorithms as H.320 with the additional specification of how to packetize the video streams for transmission on a packet-based network. These video algorithms, which were designed to be transmitted in a continuous stream, do not have an elegant method of handling a large piece of missing data (e.g. a dropped packet). To handle this, H.323 specifies a mechanism to packetize a video stream whereby each packet contains either a GOB (Group of Blocks) or a MB (Macroblock) with the resultant packet never reaching a size greater than the MTU (Maximum Transmittable Unit) of the network. The packet size limit prevents a video packet from being segmented in the network and enables protocols such as RSVP (Resource Reservation Protocol) that use the UDP port to identify packets to operate on all packets in the video stream.

Data applications – The same data applications defined in the T.120 protocol suite for use with H.320 are also defined and used for H.323.

Data services – Data services in H.323 are provided by the T.120 protocol suite and are run on an IP network specific stack defined in T.123.

Synchronization – (Receive Path Delay) In H.323, the transmitter sends timestamps on each audio and video packet which represent the capture time of the data. In this case, the receiver optionally delays the audio to match the video.

H.323 terminals differ from H.320 terminals in that they must handle inter-packet arrival time jitter. Terminals are typically designed with a packet arrival jitter queue that exceeds the average video to audio codec delay skew by a significant amount. Therefore, as long as the packet arrival jitter queue remains significantly larger than the video to audio codec delay skew, there is no need for additional audio delay to cover lip synchronization. As an example, if the receive queue for audio is 150 ms and the video is rendered to synchronize to the audio queue output, then an audio to video codec skew of 20 ms would effectively reduce the jitter resilience by 20 ms, but not require additional processing in the receiver.

Table 3 summarizes the different methods used for media synchronization.

Table 2
Additional H.323 audio algorithms

Algorithm	Audio bandwidth	Type of algorithm	Bit rate (Kbps)	Frame size
G.723.1	Narrowband (3 kHz)	MPE/ ACELP	5.3, 6.4	30 ms frame
G.729	Narrowband (3 kHz)	ACELP	8	10 ms frame

Table 3
Media synchronization methods

	H.323 (IP networks)	H.320 (ISDN)	H.324 (PSTN)
Location where audio delay is inserted	Receiver	Transmitter	Receiver
Mechanism for lip synchronization	Transmitter sends timestamps on each audio and video packets	Transmitter delays audio to match delay in video stream	Transmitter indicates to receiver the skew between audio and video streams

Multiplex – H.323 uses the Real-time Transport Protocol (RTP) for media stream packetization and synchronization. In the case of an IP based network, the UDP port number serves as the multiplex for streams moving across the network. The Real-time Transport Control Protocol (RTCP) provides both transmit and receive statistics between conference endpoints based on the transmission and reception of RTP packets.

4.1.2. Gatekeeper

The gatekeeper controls access to the network for terminals, gateways, and MCUs and provides address translation for all H.323 components. The gatekeeper provides the best opportunity for manufacturers to add standards plus features into their solution. A smartly designed gatekeeper can allow the deployment of multimedia conferencing to be tailored to a specific network configuration.

The gatekeeper's key functions include:

Admissions control – Authorizes network access for all endpoints based on a programmed criteria.

Registration and address resolution – The gatekeeper manages terminals, gateways, and MCUs. Gateways and MCUs can be added to a network without the need to configure their location in individual terminals. The gatekeeper acts as an address translation device so users may use aliases such as their email address as their name instead of a network address (e.g. 121.8.4.63) for dialling purposes.

Call management – In addition to maintaining lists of ongoing calls, rejected calls, call accounting on the use of WAN links, and the configuration of the other H.323 components, a gatekeeper can provide a call routing function. The call routing function permits endpoints to dial packet or switched based terminals without actually knowing where those terminals are or what network they are connected to. For a typical call, a user will click a name in a phone book and then be connected. The details of whether or not the call uses a gateway should be completely invisible to the user of the system.

4.1.3. Gateways

A H.323 Gateway provides real-time, two-way communications between H.323 terminals on the packet-based network and other ITU-T terminals or telephones on the public switched telephone network (PSTN) or ISDN. A gateway has the characteristics of a H.323 terminal on the packet network and of another ITU-T terminal or telephone on the switched network and provides the necessary protocol conversion between the different terminal types.

4.1.4. Multipoint controller, multipoint processor, and MCU

A Multipoint Controller (MC) provides the control services for three or more endpoints participating in a multipoint conference. A Multipoint Processor (MP) provides the media services (e.g. mixing, switching) for three or more endpoints participating in a multipoint conference. Together, the MC and the MP make up a Multipoint Conferencing Unit or MCU.

The MC provides for capability negotiation with all endpoints to achieve a common operating mode and may control conference resources such as who is multicasting video. The MC may exist in an endpoint, gatekeeper, gateway, or as the core of a H.323 MCU.

Even though three to four person conferences without an MCU using IP multicast are possible, most conferences will require an MCU to handle the audio mixing and processing requirements of hosting a conference. On the Internet, where access is typically via a low bit rate pipe, a MCU will be required to link multiple users via audio mixing and video switching as the pipe itself will not be able to handle all of the individual media streams.

Although a H.323 MCU is logically split between the MC and the MP, its functionality can be broken into the same components as a H.320 MCU: control processing, media processing, and media distribution.

Control processing – Control processing for H.323 is functionally the same as H.320 with the exception that in addition to negotiating the modes of a conference, a MCU also determines how the media in the conference will be distributed. A H.323 MCU can unicast, multicast, or use a combination of the two transport mechanisms to distribute media in a conference.

Media processing – A H.323 MCU performs the same media processing functions of rate matching, transcoding, mixing, and switching as performed by a H.320 MCU.

Media distribution – In H.323, a MCU has multiple options for media distribution. A MCU can conduct a conference using unicast, multicast, or a combination of both transport mechanisms. The MCU can mix transport mechanisms between endpoints in a conference, or mix the transports for each media stream it is sending to a single endpoint. As shown in Fig. 6, a MCU can be multicasting a mixed audio and switched video stream while at the same time unicasting individual audio mixes to specific terminals in a conference.

5. Multimedia conference coordination and inter-working

Coordination between ISDN and IP based multimedia conferencing is accomplished by gateways that translate between the H.320 and H.323 protocol domains.

The ITU-T standard H.246 defines the gateway requirements for H.323 to H.320 Gateways and in the next ITU-T study period, will be expanded to include H.323 to telephony gateways that are sometimes call Internet Telephony Gateways.

When H.246 is completed, it will describe gateways that provide inter-working between H-series multimedia terminals and other H-series multimedia terminals, voice/voiceband terminals on PSTN or ISDN, V.70 terminals on the PSTN, and multi-call applications on the PSTN.

Fig. 7 shows the H.323 to H.320 call control, link control, and media flow inter-working which is specified in H.246.

5.1. H.323–H.320 inter-working

To properly gateway between a H.323 and H.320 endpoint, H.246 specifies that a gateway must sup-

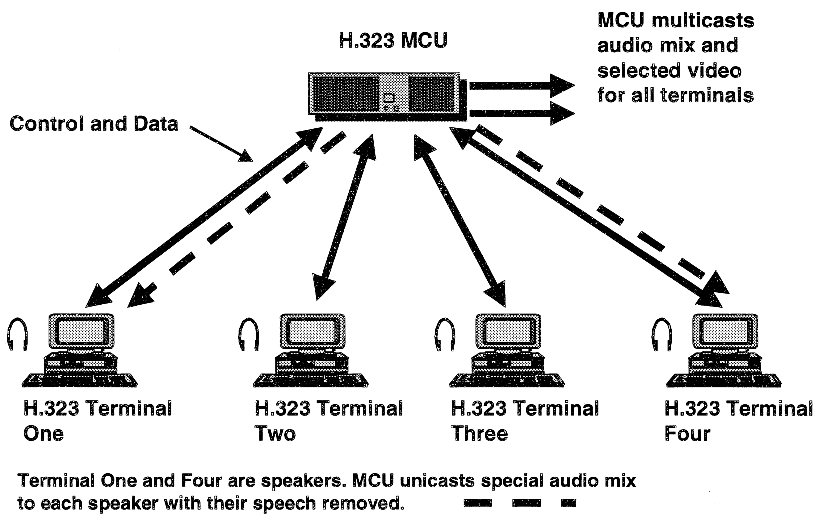

Fig. 6. H.323 multipoint using a mix of unicast and multicast media distribution.

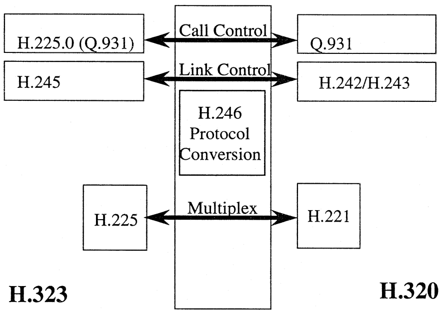

Fig. 7. H.323 to H.320 protocol conversion defined in H.246.

port all of the mandatory functionality specified in the H.320 system specification on its switched circuit network (ISDN) and all of the mandatory functionality specified in the H.323 system specification on its IP network. In requiring this functionality, H.246 guarantees that two standard compliant endpoints attempting to communicate with each other through a standard compliant gateway will succeed!

The following example shows the importance of proper coordination between the protocols:

H.246 specifies that a gateway that receives an audio mute from a H.323 endpoint on its H.245 channel (e.g. logicalChannelInactive) should pass this indication to the H.320 side of its link as an audio mute message (e.g. AIM). The failure of a gateway to implement this coordination can have a disastrous effect on a conference. If you consider a H.323 endpoint dialled into a H.320 multipoint bridge (MCU) through a gateway where its audio is inadvertently muted, there is no way for anyone in the H.320 based conference to determine why there is no audio coming from the muted endpoint. If the conference is set up for voice activated switching, the muted endpoint may never be viewed by anyone in the conference. Coordination between the protocols is key for a successful conference that spans the two networks. Of course for this to work, the endpoints must also send and display audio mute indications.

6. Multimedia conferencing future

Over the next few years, networks supporting IP traffic will be upgraded with high bandwidth technologies and quality of service protocols to improve the latency, packet loss and bandwidth availability problems usually associated with IP network operation. These improvements will evolve slowly with protocols like RSVP being rolled out on corporate campuses, Intranets, and the Internet. In parallel, dial up links will be enabled by higher bandwidth technologies like xDSL.

IP based videoconferencing will grow with high quality end points leveraging server products and applications with a network infrastructure that has been upgraded with bandwidth and service options to support real time audio, video and data. While the IP network is upgraded, a mix of switched and packet based networks will coexist and companies will be deploying end-to-end solutions which leverage packet networks where possible and gateways to the existing switched networks (PSTN and ISDN) when it is appropriate.

Looking forward, IP networks will reign as the transport of choice for multimedia conferencing due to the networks ability to run multiple services and applications while simultaneously providing the transport for a multimedia conference. As consumers become accustomed to combining these technologies on a single network connection, the growth of conferencing on a conferencing dedicated network connection will be drastically reduced. For IP based conferencing to be successful, the network, the endpoints, and the server products and applications will need to provide a solution that is both as reliable and as available as today's public switched network. Consumers will accept nothing less than the reliability and accessibility of today's telephone network for their everyday communications.

References

[1] ITU-T Recommendation H.320, Narrow-band Visual Telephone Systems and Terminal Equipment, 1996.
[2] ITU-T Recommendation H.323, Packet Based Multimedia Communication Systems, 1998.
[3] ITU-T Recommendation E.164, Numbering Plan for the ISDN ERA, 1991.
[4] ITU-T Recommendation G.711, Pulse Code Modulation (PCM) of Voice Frequencies, 1988.
[5] ITU-T Recommendation G.722, 7 kHz Audio-coding within 64 kbit/s, 1988.
[6] ITU-T Recommendation G.723.1, Dual Rate Speech Coder

for Multimedia Communication Transmitting at 5.3 and 6.3 kbit/s, 1996.
[7] ITU-T Recommendation G.728, Coding of Speech at 16 kbit/s using Low-delay Code Excited Linear Prediction, 1992.
[8] ITU-T Recommendation H.221, Frame Structure for a 64 to 1920 kbit/s Channel in Audiovisual Teleservices, 1993.
[9] ITU-T Recommendation H.223, Multiplexing Protocol for Low Bitrate Multimedia Communication, 1996.
[10] ITU-T Recommendation H.225.0, Media Stream Packetization and Synchronization for Visual Telephone Systems on Non-Guaranteed Quality of Service LANs, 1996.
[11] ITU-T Recommendation H.230, Frame-synchronous Control and Indication Signals for Audiovisual Systems, 1995.
[12] ITU-T Recommendation H.243, Procedures for Establishing Communication Between Three or More Audio Visual Terminals using Digital Channels up to 2 Mbit/s, 1993.
[13] ITU-T Recommendation H.261, Video Codec for Audiovisual Services at px64 kbit/s, 1993.
[14] ITU-T Recommendation H.262, Generic Coding of Moving Pictures and Associated Audio: Video, ISO/IEC 13818-2, 1996.
[15] ITU-T Recommendation H.263, Video Coding for Low Bitrate Communication, 1996.
[16] ITU–T Recommendation H.320, Narrow-band ISDN Visual Telephone Systems and Terminal Equipment, 1993.
[17] ITU-T Recommendation H.323, Visual Telephone Systems and Equipment for Local Area Networks Which Provide a Non-guaranteed Quality of Service, 1996.
[18] ITU-T Recommendation H.324, Terminal for Low Bitrate Multimedia Communication, 1996.
[19] ITU-T Recommendation T.120, Data Protocols for Multimedia Conferencing, 1997.
[20] ITU-T Recommendation T.123, Protocol Stacks for Audiographic and Audiovisual Teleconference Applications, 1996.
[21] ITU-T Recommendation H.230, Frame Synchronous Control and Indication Signals for Audiovisual Systems, 1995.
[22] ITU-T Recommendation H.245, Control Protocol for Multimedia Communication, 1998.
[23] ITU-T Recommendation H.246, Interworking of H.Series Multimedia Terminals with H.Series Multimedia Terminals and Voice/voiceband Terminals on GSTN and ISDN, 1998.
[24] H. Schulzrinne, S. Casner, R. Frederick, V. Jacobson, RTP: A Transport Protocol for Real-Time Applications, RFC 1889, Internet Engineering Task Force, 1996.

Mark Reid is Vice President of Engineering at VideoServer Inc. Prior to joining VideoServer, Mark led H.323 development for PictureTel's Personal Systems Division. Mark actively participates in ITU-T Study Group 16 and is the editor of H.246, the standard for multimedia gateways. Mark received a B.S.C.S from the Ohio State University and M.S.E.E from Northeastern University.