

VCON

VIDEOCONFERENCING

WHITE PAPER

H.323 STANDARD

January 1998

<i>International</i>	<i>North America</i>	<i>China</i>	<i>Germany</i>	<i>Japan</i>	<i>Spain</i>	<i>United Kingdom</i>
VCON Ltd. P.O. Box 12747, 22 Maskit Street Herzliya 46733 Israel	VCON Inc. 17440 N. Dallas Pkwy, Suite 200 Dallas TX 75287-7308 USA	VCON China Unit 1703 Hanwei Plaza No. 7 Guang Hua Road Chao Yang District Beijing 100020 China	VCON GmbH Voltastrasse 6 Dietzenbach 63128 Germany	VCON Japan Sumitomo Corporation Kanda Bld. 1F 3-24-4 Kandanshiki-cho Chiyoda-ku Tokyo, 101 Japan	VCON Spain Concha Espina 63, 5 Dcha Madrid 28016 Spain	A2 Westacott Business Center Maidenhead Office Park, Little Maidenhead Berkshire SL6 3RT United Kingdom
Ph: +972-9-959-0059 Fx: +972-9-956-7244	Ph: +1-972-735-9001 Fx: +1-972-735-9099	Ph: +86-10-6561-42198 Fx: +86-10-6561-42197	Ph: +49-6074-35882 Fx: +49-6074-35883	Ph: +81-3-5280-7789 Fx: +81-3-3219-7958	Ph: +34-1-5158900 Fx: +34-1 5158905	Ph: +44-1628-829555 Fx: +44-1628-829777

Introduction

The H.323 standard encompasses audio, video and data communications across packet-based networks – LAN, Intranet, Extranet and Internet. The H.323 standard was developed to allow multimedia products and applications from multiple vendors to interoperate. Compatibility is the key concern for vendors and end users for LAN-based products in the consumer, business, entertainment, and professional markets.

The International Telecommunications Union (ITU) sets standards for multimedia communications over Local Area Networks (LAN) as well as other forms of communication. Today’s corporate networks include Ethernet, Fast Ethernet, Token Ring and ATM that can be interconnected by various private and public Wide Area Networks (WAN) – ISDN, V.35, E1/T1, Frame Relay, ATM and others. The H.323 standard is an important building block for a broad new range of collaborative, LAN-based applications for multimedia communications. In addition, the standard allows direct connection to PPP based networks via ISDN and POTS.

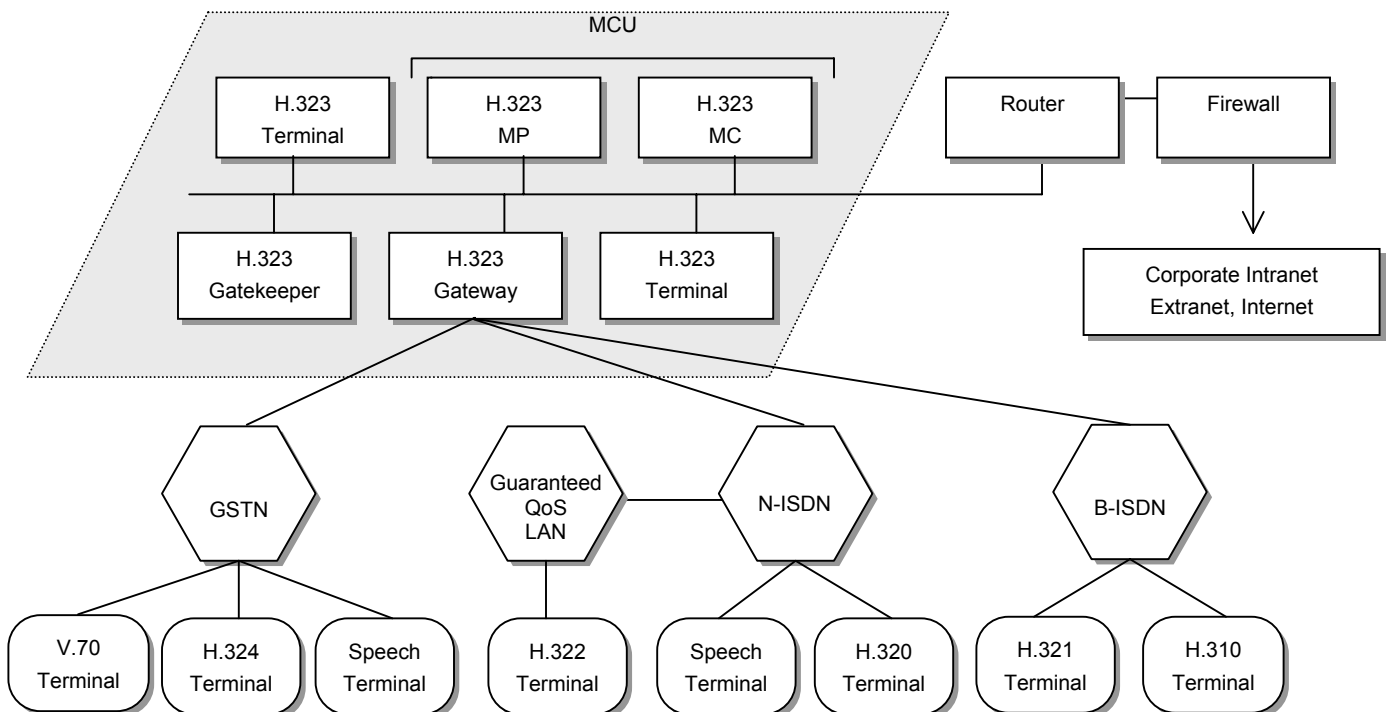
In 1996 the ITU approved the H.323 specification. The standard is broad in scope and includes both stand-alone devices and embedded personal computer technology as well as point-to-point and multipoint conferences.

The standard addresses issues such as call and session control, multimedia and bandwidth management for point-to-point and multipoint conferences.

H.323 is only a part of a larger series of communications standards set by the ITU that enable videoconferencing across a range of networks. Known as H.32x, this series includes H.320 and H.324, which address ISDN and POTS communications respectively.

This paper provides a summary of the H.323 standard, its benefits, architecture, and applications.

Diagram 1: The scope of H.323 standard.



H.323 Applications:

- Desktop and Room System Videoconferencing
- Internet Telephony and Video Telephony
- CTI
- Collaborative Computing
- Intranet and Internet Business Conferencing
- Distance Learning
- Support and Help Desk Applications
- Interactive Shopping
- Banking Kiosks
- Videoconferencing services and Billing Applications
- TeleMedicine
- Security Systems
- Audio/Video Mail
- Audio/Video Broadcasting
- Video-on-demand
- Others

The Importance of H.323

The H.323 Recommendation is comprehensive, yet flexible, and are set to grow into the mainstream market for several reasons:

- H.323 sets multimedia standards for existing corporate infrastructure (i.e. IP-based networks). Designed to compensate for the absence of guaranteed Quality of Service (QoS), H.323 allows customers to use multimedia applications without changing their network infrastructure.
- IP LANs are becoming more powerful. Ethernet bandwidth is expanding from 10 Mbps to 100 Mbps, and Gigabit Ethernet is already on the horizon.
- Many companies have moved to the low cost switching network technology allowing practically unlimited number of the simultaneous videoconferencing sessions without the performance degradation of other networking applications.
- H.323 can provide additional benefits when more network services are available - ATM QoS, RSVP for resources reservation, firewalls, etc.
- By defining guideline for device-to-device, application-to-application, and vendor-to-vendor interoperability, H.323 allows products to interoperate freely with other H.323-compliant products. Alternatively, the standard is very flexible allowing vendor-specific enhancements.
- Network loading can be managed using a centralized or distributed gatekeeper. With H.323, the network manager can restrict the amount of network bandwidth available for conferencing. Multicast support also reduces bandwidth requirements.
- H.323 provides capability of the network security on different levels including centralized call control, firewalls, encryption, etc.
- H.323 has the support of many computing and conferencing companies and organizations, including Intel, Microsoft, and Netscape. The efforts of these organizations will generate a higher level of awareness in the market.

Main Advantages of H.323

Network Independence

H.323 is designed to run on top of common network architectures. As network technology evolves, and as bandwidth management techniques improve, H.323-based solutions will be able to take advantage of those enhanced capabilities.

Bandwidth Management

Video and audio traffic is bandwidth-intensive and could clog the corporate shared network if not monitored and controlled. H.323 addresses this issue by providing bandwidth management. The number of simultaneous H.323 connections within their network or the amount of bandwidth available to H.323 applications can be limited. On one hand, these limits ensure that critical traffic on the LAN will not be disrupted; on the other hand, they leave enough resources for other network activities. Every H.323 terminal can provide the bandwidth management for the specific session. Such a mechanism (called Automatic Bandwidth Adjustment) increases and decreases video bit rate due to the network behavior - changes in the latency, jitter and packets lost.

Inter-Network Conferencing

Many users want to conference to a remote site. It is possible to do this in a pure H.323 environment using the direct PPP connection or the corporate Intranet, Extranet, or even Internet. H.323 also establishes a means of linking packet-switched based with circuit-switched based (H.320, H.324) or with ISDN-based videoconferencing systems. H.323 uses common codec technology from different videoconferencing standards to minimize or eliminate protocols transcoding and to provide optimum performance.

Platform and Application Independence

H.323 is not tied to any hardware or operating system. H.323-compliant platforms will be available in many sizes and shapes, including video-enabled personal computers, dedicated platforms and turnkey boxes.

Multipoint Support

H.323 can support conferences of three or more endpoints with or without requiring a centralized Multipoint Control Unit (MCU). Multipoint capabilities can be distributed and implemented as a part of H.323 endpoints (terminals).

Multicast Support

H.323 supports multicast transport in multipoint conferences when group management protocol (like IGMP) is used and supported by the network. Multicasting sends a single packet of information to a number of destinations on the network without replication. A multicast transmission uses bandwidth more efficiently since all stations in the multicast group (and only these stations) read a single stream.

Interoperability

Users want to communicate without worrying about compatibility at the receiving point. In addition to ensuring that the receiver can decompress the information, H.323 establishes methods to exchange capabilities between clients and to set common capabilities for the conference. The standard also establishes common call setup and control protocols.

Codec Standards

H.323 establishes standards for compression and decompression of audio and video data streams, ensuring that equipment from different vendors will have some area of common support. It also gives full flexibility for additional features and enhanced performance based on vendor specific hardware and software.

Flexibility

An H.323 conference can include end-points with different capabilities. For example, a terminal with audio-only capabilities can participate in a conference with terminals that have video and/or data capabilities. Furthermore, an H.323 multimedia terminal can share the data portion of a videoconference with a T.120 data-only terminal, while sharing voice, video, and data with other H.323 terminals.

H.323 Architecture

The H.323 Recommendation covers the technical requirements for audio and video communication services over packet-switched networks. H.323 references to the T.120 specifications for data conferencing and enables conferences that include a data capability. The scope of H.323 does not include the network itself or the transport layer that may be used to connect various networks.

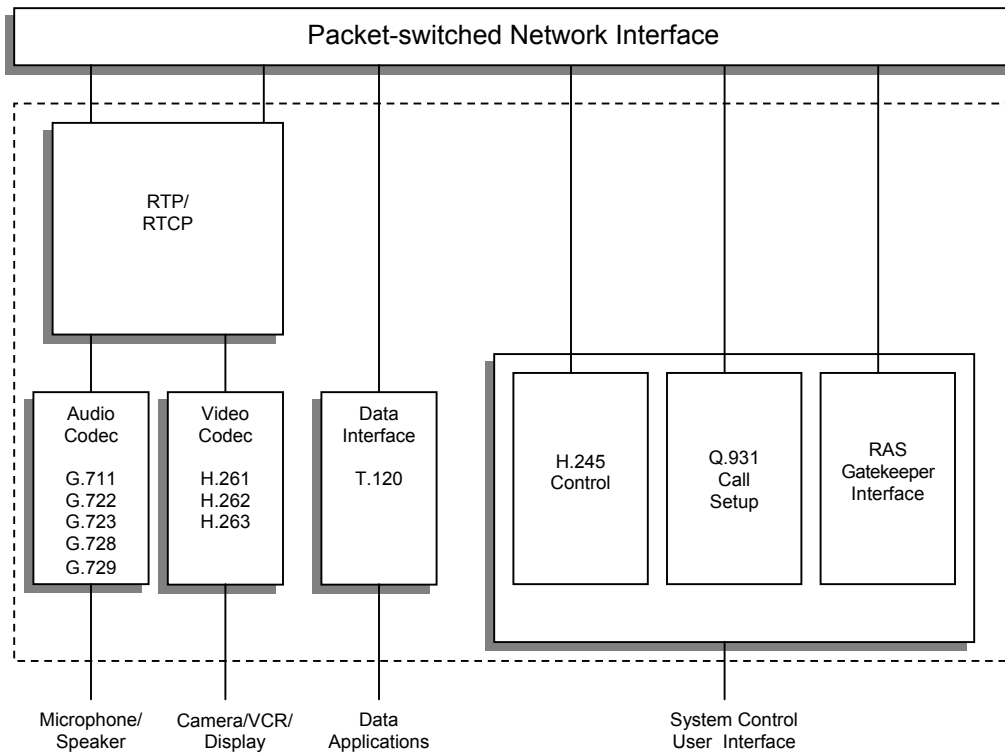
H.323 defines four major components for a network-based communications system: Terminals, Gateways, Gatekeepers, and Multipoint Control Units.

Terminals

Terminals are the client end-points that provide real-time one-way or two-way communications. All terminals must support voice communications; video and data are optional. H.323 specifies the modes of operations required for different audio, video, and/or data terminals to work together. It will be the dominant standard of the next generation of Internet phones, audio conferencing terminals, and video conferencing technologies.

All H.323 terminals must also support H.245, which is used to negotiate channel usage and capabilities. Three additional components are: a scaled-down version of Q.931 for call signaling and call setup; Registration/Admission/Status (RAS), which is a protocol used to communicate with a Gatekeeper; and support for RTP/RTCP for sequencing audio and video packets.

Diagram 2: Terminal Components.



Gateways

The Gateway is optional in an H.323 conference. Gateways provide many services, including a translation function between H.323 conferencing endpoints and other terminals. This function includes translation between transmission formats (i.e. H.225 to H.221) and between communications procedures (i.e. H.245 to H.242). In addition, the Gateway also translates between audio and video codecs and performs call setup and clearing on both the LAN and the WAN side. Diagram 3 shows an H.323/H.320 Gateway.

The purpose of the Gateway is to reflect the characteristics of a H.323 endpoint to a non-H.323 endpoint and vice versa. The primary applications of Gateways are likely to be:

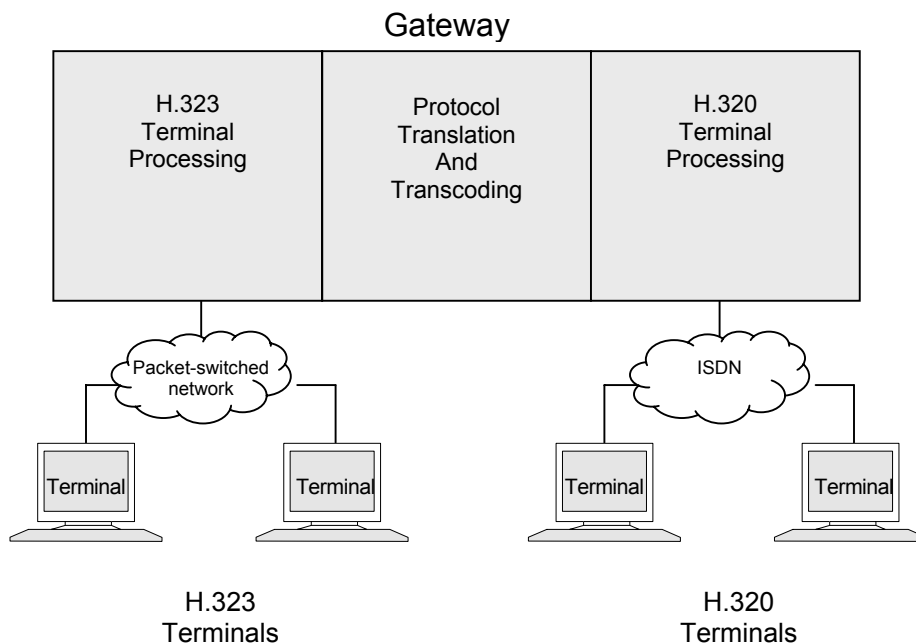
- Establishing links with analog and digital speech terminals.
- Establishing links with non-H.323 terminals.
- Gatekeeper functionality (mandatory).

Gateways are not required if connections to other standards are not needed, since endpoint may directly communicate with other H.323 endpoints over the same LAN, Intranet, Extranet or even Internet.

Terminals communicate with Gateways using the H.245 and Q.931 protocols.

The actual number of H.323 terminals that can communicate through the Gateways is not subject to standardization. Similarly, the number of WAN connections and the number of simultaneous independent conferences supported are left to the manufacturer. By incorporating Gateway technology into the H.323 specification, the ITU has positioned H.323 as the glue that holds the world of standards-based conferencing endpoints together.

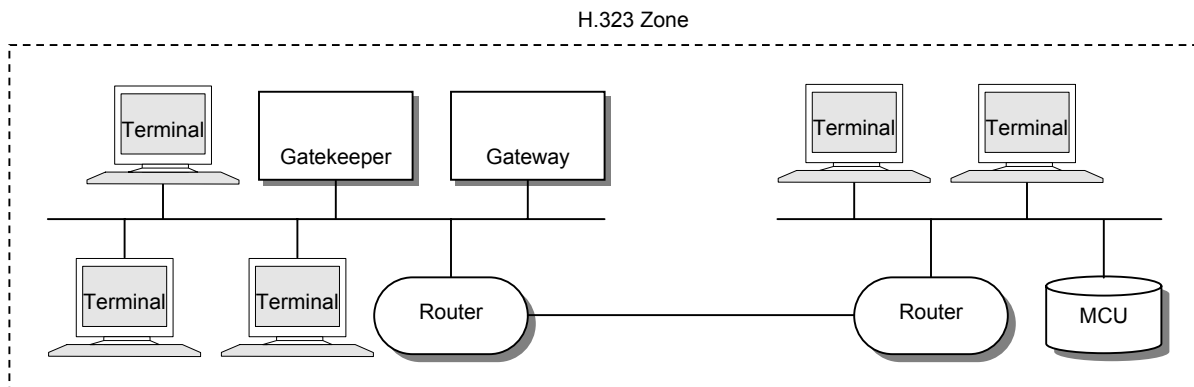
Diagram 3: H.323/H.320 Gateway



Gatekeepers

Gatekeepers perform a number of important functions that help preserve the integrity of the corporate data network. The first one is address translation from H.323 aliases for terminals and gateways to network addresses, as defined in the RAS specification. The second function is access control, preventing non-authorized videoconferencing sessions. The third function is bandwidth management, which is also designated within RAS. For instance, if a network manager has specified a threshold for the number of simultaneous conferences on the LAN, the Gatekeeper can refuse to make any more connections once the threshold is reached. The effect is to limit the total conferencing bandwidth to some fraction of the total available; the remaining capacity is left for email, file transfers, and other LAN activities. The fourth function is to manage a number of terminals, gateways, and MCUs as a single logical group known as the H.323 zone.

Diagram 4: H.323 Zone



A Gatekeeper is not required in an H.323 system. However, if a Gatekeeper is present, it is mandatory that terminals make use of its services.

Gatekeepers can also play a role in multipoint connections. To support multipoint conferences, users would employ a Gatekeeper to receive H.245 Control Channels from two terminals in a point-to-point conference. When the conference switches to multipoint, the Gatekeeper can redirect the H.245 Control Channel to a Multipoint Control Unit (MCU). The Gatekeeper need not process the H.245 signaling; it only needs to pass it between the terminals and the MCU.

LANs that contain Gateways should also contain a Gatekeeper to translate incoming E.164 addresses into Transport Addresses.

Gatekeeper Functions

Address Translation	Translation of Alias to Transport Address using a table that is updated with Registration messages. Other messages of updating the translation table are also allowed.
Admissions Control	Authorization of LAN access using Admission Request, Confirm and Reject (ARQ/ARC/ARJ) messages. LAN access may be based on call authorization, bandwidth, or some other criteria. Admissions Control may also be a null function which admits all requests.
Bandwidth Control	Support for Bandwidth Request, Confirm and Reject (BRQ/BCF/BRJ) messages. This may be based on bandwidth management. Bandwidth Control may also be a null function which accepts all requests for bandwidth changes.
Zone Management	The Gatekeeper provides the above functions for terminals, MCUs, and Gateways which have registered within its Zone of control.
Call Control Signaling	In a point-to-point conference, the Gatekeeper may process Q.931 call control signals. Alternatively, the Gatekeeper may send the endpoints Q.931 signals directly to each other.
Call Authorization	The Gatekeeper may reject a call from a terminal based on the Q.931 specification. The reasons for rejection may include, but are not limited to restricted access to/from particular terminals or Gateways or restricted access during certain periods of time. The criteria for determining if authorization passes or fails is outside the scope of H.323.
Bandwidth Management	The Gatekeeper may reject calls from a terminal if it determines that sufficient bandwidth is not available. This function also operated during an active call if a terminal requests additional bandwidth. The criteria for determining if bandwidth is available is outside the scope of H.323.
Call Management	The Gatekeeper may maintain a list of ongoing H.323 calls in order to indicate that a called terminal is busy or to provide information for the Bandwidth Management function.

Multipoint Control Units (MCU)

A Multipoint Control Unit (MCU) supports conferences between three or more endpoints. Under H.323, an MCU consists of a Multipoint Controller (MC), which is required, and zero or more Multipoint Processors (MP). The MC handles H.245 negotiations between all terminals to determine common capabilities for audio and video processing. The MC also controls conference resources by determining which, if any, of the audio and video streams will be multicast.

The MC does not deal directly with any of the media streams. This is left to the MP, which mixes, switches, and processes audio, video and/or data bits. MC and MP capabilities can exist in a dedicated component or be a part of other H.323 components.

Multipoint Conferences

Multipoint conference capabilities are handled in a variety of methods and configurations under H.323. The Recommendation uses the concepts of centralized and decentralized conferences as described in Diagrams 5 and 6 below.

Diagram 5: Decentralized Conference with Multicast

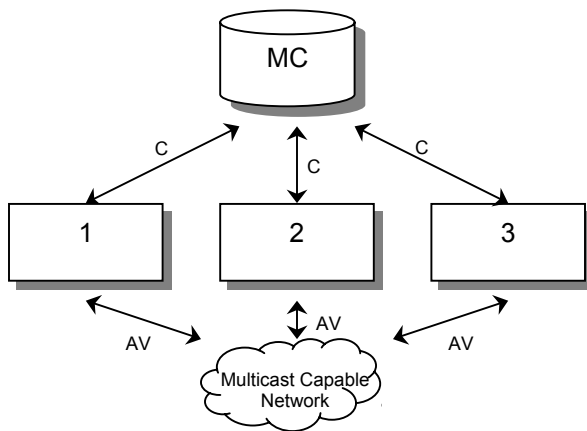
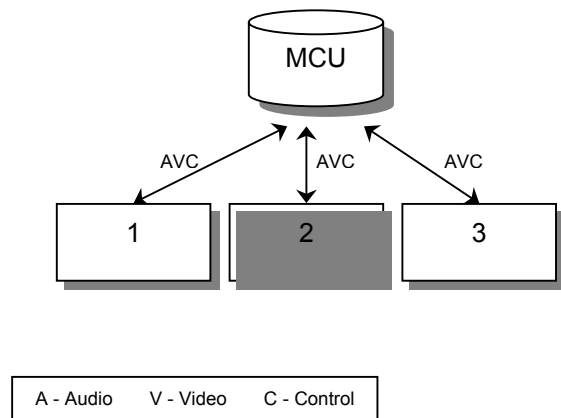


Diagram 6: Centralized Conference with Unicast



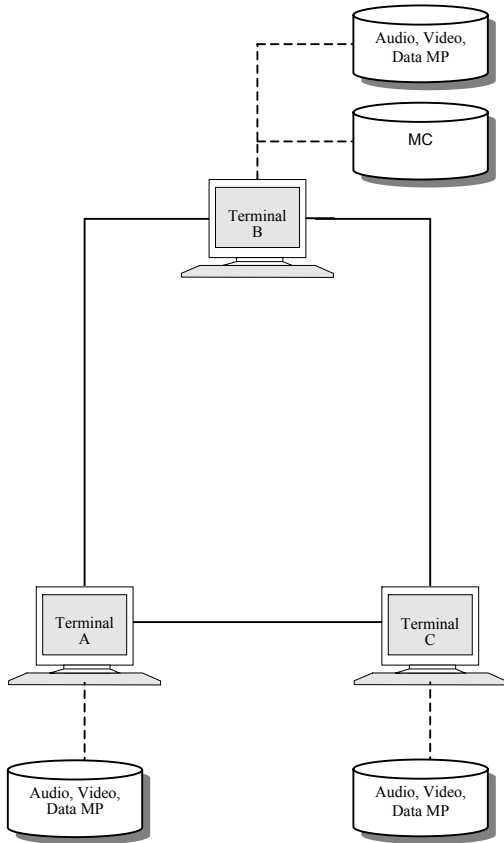
Centralized multipoint conferences require the existence of an MCU to facilitate a multipoint conference. All terminals send audio, video, data, and control streams to the MCU in a point-to-point fashion. The MC centrally manages the conference using H.245 control functions that also define the capabilities for each terminal. The MP does the audio mixing, data distribution, and video switching/mixing functions typically performed in multipoint conferences and sends the resulting streams back to the participating terminals. The MP may also provide conversion between different codecs and bit rates and may use multicast to distribute processed video. A typical MCU that supports centralized multipoint conferences consists of an MC and an audio, video and data MP.

Decentralized Multipoint Conferences make use of multicast technology. Participating H.323 terminals multicast audio and video to other participating terminals without sending the data to an MCU. Note that control of multipoint data is still centrally processed by the MC, and H.245 Control Channel information is still transmitted in a point-to-point mode to an MC (excluding loosely-coupled conferences described later).

Receiving terminals are responsible for processing the multiple incoming audio and video streams. Terminals use H.245 Control Channels to indicate to an MC how many simultaneous video and audio streams they can decode. The number of simultaneous capabilities of one terminal does not limit the number of video or audio streams which are multicast in a conference. The MP can also provide audio and video selection and /or mixing in a decentralized multipoint conference (Diagram 7).

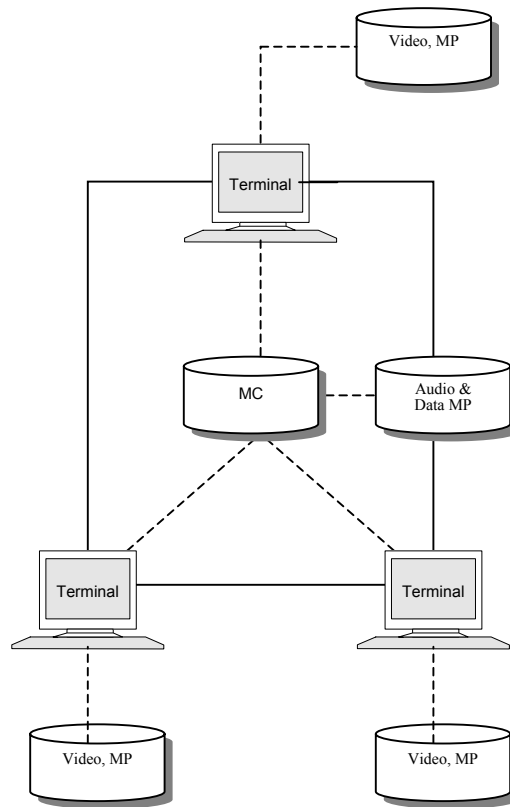
Multipoint Conference

Diagram 7



Decentralized
Audio, video, and data

Diagram 8



Hybrid
Centralized audio, data, and control
Decentralized video

Hybrid multipoint conferences use a combination of centralized and decentralized features. H.245 signals either an audio or video stream is processed through point-to-point messages to the MC. The remaining signal (audio or video) is transmitted to participating H.323 terminals through multicasting (Diagram 8).

H.323 also supports mixed multipoint conferences in which some terminals are in a centralized conference, others are in a decentralized conference, and an MCU provides the bridge between the two types. The terminal is not aware of the mixed nature of the conference, only of the mode of conference in which it sends and receives.

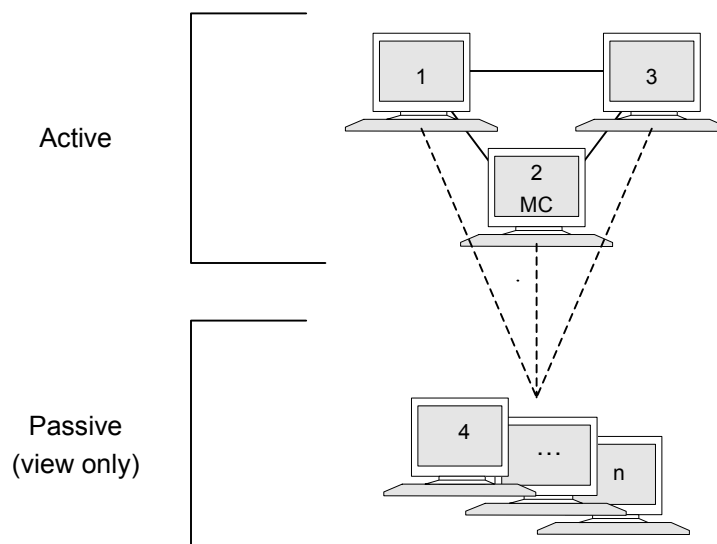
By supporting both multicast and unicast approaches, H.323 spans current generation and future networking technologies. Multicast makes more efficient use of network bandwidth, but places higher computational loads on the terminals, which have to mix and switch their own audio/video receiving streams in case of many-to-many multipoint videoconference. Additionally, multicast support is required in network routers and switches.

While the theoretical limit on the number of participants is high, in many circumstances users will find that more than 10-20 participants per conference are insufficient.

There is a requirement for applications with a very high number of conferences (i.e. distance learning, audio and video broadcasting). H.323 defines such applications as so called "loosely-coupled" videoconferencing (regular conferencing is called "tightly-coupled"). Two different groups may be a part of such a conference: Panel (active participants) and Passive participants. The Panel members behave exactly as in tightly-coupled H.323 multipoint videoconferencing. Passive participants can only monitor the conference using RTP/RTCP rules. The Panel may be divided into two groups: permanent and temporary members. Any temporary members may be exchanged with Passive participants by means of their role in the session. The roster of the conference may be built automatically using RTCP identification fields that all users have to send periodically. It gives the permanent members the capability of inviting Passive participants to the Panel as temporary members. On the other hand, a Passive participant is able to join the Panel if he/she knows the MC address. The joining-in of either Active or Passive participants can be enabled using announcement procedures. For such announcements, the standard defines that IETF Session Description Protocol (SDP) be used. The file containing SDP script can be sent to one or a group of participants with specific security requirements using e-mail (SMTP), the Web (HTTP), or IETF Session Protocol (SAP). This SDP script defines the role the participant plays in the announced conference (permanent active, temporary active, passive, passive capable of joining the panel and the streams available for him (audio, video, data)).

An example of such loosely-coupled decentralized videoconferencing is shown in Diagram 9 below.

Diagram 9: Active and Passive Participants in Decentralized Multipoint



An MC may be located within a Gatekeeper, Gateway, Terminal, or MCU.

Consider a simple example where a multipoint conference is set up between three clients (Figure 7). One client terminal (Client B) performs the MC function. All the terminals could use multicast to participate in a decentralized conference. An MP function on each node would mix and present the incoming audio and video signals to the user. This approach minimizes the need for specialized network resources and significantly reduces the total system cost. However, the network must be configured to support multicasting.

One advantage of centralized multipoint conferencing is that all H.323 terminals and gateways support point-to-point communications. In this case, equipment from different vendors may participate in such conferencing. The MCU may output multiple unicasts to the conference participants and no special network capabilities are required. Alternatively, the MCU may receive multiple unicast, mix audio and switch video, and output a multicast stream conserving network bandwidth.

Another benefit of centralized MCU may be when it is implemented as a single unit together with H.320/H.324 gateway or network equipment - router, switch, or firewall.

In a hybrid environment, a part of the participants save network resources by using multicasting and distributed MP. One of the end-points will play the role of a centralized MP for stations without such a capability. The hybrid environment may be built for different media, for example, audio using a centralized MCU, and video using a decentralized one.

Communications Under H.323

Communications under H.323 can be considered as a mix of audio, video, data and control packets. Audio capabilities, Q.931 call setup, RAS control, and H.245 signaling are required. All other capabilities, including video and data conferencing are optional. When multiple algorithms are possible, the algorithms used by the encoder are derived from information passed by the decoder during the H.245 capability exchange. H.323 terminals are also capable of asymmetric operation (different encode and decode algorithms) and can send/receive multiple video and audio channels.

Control

The call control functions are the heart of the H.323 terminal. Overall system control is provided by three separate signaling channels: the H.245 Control Channel, the Q.931 Call Signaling Channel, and the RAS Channel. Control functions include signaling for call setup, capability exchange, signaling of commands and indications and messages to open and describe the content of logical channels. All audio, video and control signal pass through a control layer that formats the data streams into messages for output to the network interface. The reverse process takes place for incoming streams.

The H.245 Control Channel is a reliable channel that carries control messages governing operation of the H.323 entity, including capabilities exchange, opening and closing of logical channels, preference requests, flow control messages, and general commands and indications. Capabilities exchange is one of the fundamental capabilities in the ITU recommendation; H.245 provides for separate receive and transmit capabilities as well as for methods to describe these details to other H.323 terminals. There is only one H.245 Control Channel between any two terminals.

The Call Signaling Channel uses Q.931 to establish a connection between two terminals. The RAS signaling function performs registration, admission, bandwidth changes, status, and disengage procedures between endpoints and Gatekeepers. RAS is not used if a Gatekeeper is not present.

Audio

Audio signals contain digitized and compressed speech. The compression algorithms supported by H.323 are all proven ITU standards. H.323 terminals must support the G.711 voice standard for speech compression. Support for other ITU voice standards is optional.

The different ITU recommendations for digitizing and compressing speech signals reflect different tradeoffs between speech quality, bit rate, computer power, and signal delay. G.711 generally transmits voice at 56 or 64 kbps, well within the bandwidth limits likely on a LAN. As G.723 operates at very low bit rates, it is a popular audio codec in H.323 applications. This codec also eliminates transcoding modules in H.323-H.324 Gateways. The codec also supports built-in silence detection to save even more bandwidth and to support half-duplex audio systems. The most powerful H.323 terminals will also implement high quality 16 kbps G.728 codec also allowing to save transcoding modules in H.323-H.320 Gateways. Due to the high processing power required, it is implemented in hardware based codecs only. H.323 room systems also support G.722 codec for a superior audio quality.

Video

While video capabilities are optional, any video-enabled H.323 terminal must support the H.261 codec (support for H.263 is optional). Video information is transmitted at a rate no greater than that selected during the capability exchange. H.261, which provides a measure of compatibility across many of the different ITU recommendations is used with communication channels that are multiples of 64 kbps.

Flexible H.261 implementations give the capability of generating any bit rate, even if it is not a multiple of 64 kbps. This type of codec increases video quality in many cases. For example, there is a 128 kbps session and G.728 audio is used, more than 100 kbps for video is available (depending on data rates and overhead).

H.263 is a backwards-compatible update to H.261. H.263 picture quality is greatly improved by using a motion-estimation technique, predicted frames, and a Huffman coding table optimized for low bit rate transmissions. H.263 defines five standardized picture formats. Communications between H.261 systems and H.263 systems is facilitated because both must support QCIF.

If the H.323 terminals support both H.261 and H.263 codecs, it is recommended to use H.263 for low bit rate conversations and H.261 for high bit rate sessions. The use of exact bit rate depends on vendor specific implementations and application requirements. The advantage of hardware based codecs is their capability to support high bit and frame rates in large picture formats: CIF, 4CIF and 16CIF.

ITU Image Formats for Videoconferencing

Videoconferencing Picture 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	702 x 576	N/A	Optional
16CIF	1408 x 1152	N/A	Optional

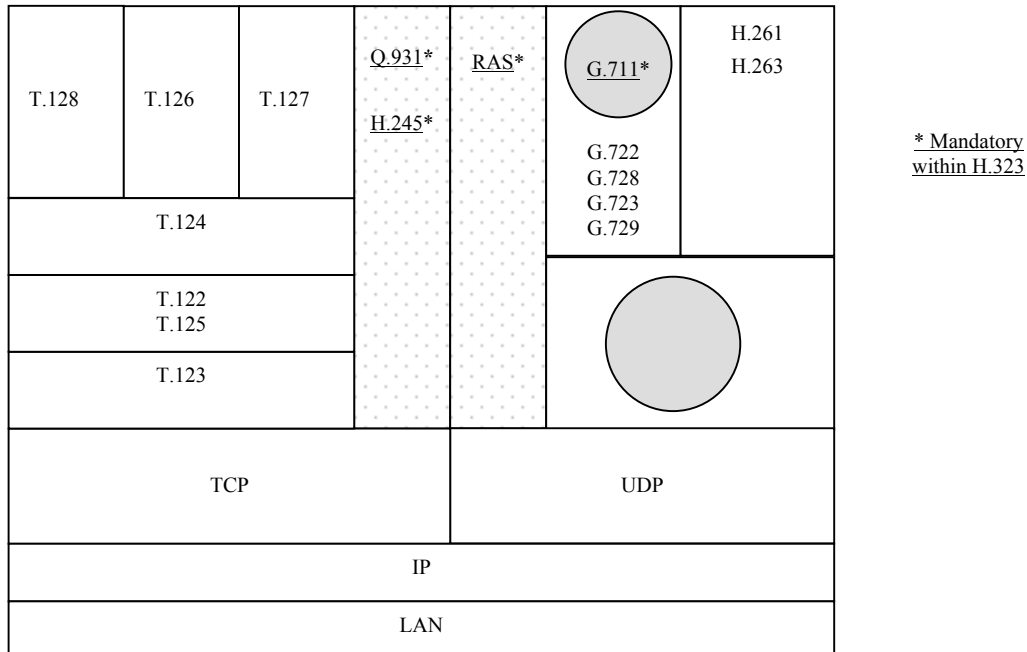
Data

Data conferencing is an optional capability. When supported, data conferencing enables collaboration through applications such as shared whiteboards, application sharing, and file transfer.

H.323 supports data conferencing through the T.120 specification. An ITU standard, T.120, addresses point-to-point and multipoint data conferences. It provides interoperability at the application, network, and transport level.

An H.323 system can support data by incorporating T.120 capabilities into clients and Multipoint Control Units. The MCU controls and mixes data conferencing information. Presently, T.120 supports tree based unicast architecture only, to be updated for multicast support.

Diagram 10: H.323 Control together with audio-video and data protocols



IP Networking and Multimedia Conferencing

H.323 uses both reliable and unreliable communications. Control signals and data require reliable transport because the signals must be received in the order in which they were sent and cannot be lost. Audio and video streams lose their value with time. If a packet is delayed, it may not have relevance to the end user. Audio and video signals use the more efficient unreliable transport. Most existing H.323 products use IP for the information exchange.

Reliable transmission of messages uses a connection-oriented mode for data transmission. In the IP stack, this type of transmission is accomplished with TCP. Reliable transmission guarantees sequenced, error-free, flow-controlled transmission of packets, but can delay transmission and reduce throughput. H.323 uses reliable (TCP) end-to-end services for the H.245 Control Channel, the T.120 Data Channels, and the Call Signaling Channel.

Within the IP stack, unreliable services are provided by User Datagram Protocol (UDP). Unreliable transmission is a mode that promises nothing more than “best effort” delivery. UDP offers minimal control information. H.323 uses UDP for the audio, video, and the RAS Channel.

The Real-Time Protocol (RTP) is developed by the Internet Engineering Task Force (IETF) to handle streaming audio and video for multimedia information transport over the Internet’s Multicast Backbone (MBONE). It was slightly changed by ITU to fit H.323 requirements.

An RTP header containing a timestamp and sequence number envelops each audio and video packet. With appropriate buffering at the receiving station, timing and sequence information allows the application to eliminate duplicate packets; reorder out-of-sequence packets; synchronize sound, video and data; and achieve continuous playback in spite of varying latencies.

The Real-Time Control Protocol (RTCP) is used for the control of RTP and monitors the quality of service, conveys information about the session participants, and periodically distributes control packets containing quality and identification information to all session participants.

In Multipoint videoconferencing, H.323 terminals may use Internet Group Management Protocol (IGMP) for audio and video multicasting. IGMP is supported by major switches and routers for more efficient transport in Intranet, Extranet and Internet.

Having sufficient bandwidth for a multimedia call is critical and difficult to ensure in large packet networks like the Internet or corporate Intranet. Another IETF protocol, the Resource Reservation Protocol (RSVP) allows a terminal to request a specific bandwidth, priority, maximum latency and processor time for a particular data stream and receive a reply indicating whether the request has been granted.

Although RSVP is not an official part of the H.323 standard, many products will support it, because QoS may be critical to the success of videoconferencing. However, RSVP needs to be supported by any intermediate networking equipment, especially routers.

ATM environment

H.323 defines procedure for extended functionality when the terminal is connected to an ATM network. Usage of this mode, including required ATM parameters, has to be negotiated between all participants. The difference is that audio and video packets are sent using native ATM addressing with optional Quality of Service (QoS) type setting. All other packets are sent as in regular H.323 implementation (i.e. using IP). Currently, standard defines using only AAL5 cells. Other adaptation layers usage, and native ATM multicasting capabilities are for further study.

Interoperability

In the past few years, interoperability testing has come to the forefront of the conferencing industry. Sponsored by the IMTC and dozens of individual hardware and software companies, interoperability testing enables developers to test their H.32x- and T.120 compliant products with others.

While the ITU's role is that of a standards-setting body, the IMTC focuses on the practical validation and promotion of standards. The IMTC's emphasis is on multimedia teleconferencing, including still-image graphics, full-motion video, and data teleconferencing. The IMTC is focused on ensuring the adoption of the required standards and education of the market.

IMTC-organized events are intended to facilitate the development and delivery of standards-based conferencing products and services and to continue promoting the importance of industry wide interoperability as a base for building consumer confidence. Testing is likely to extend over a protracted period of time as multiple vendors cooperate to test a multi-dimensional matrix of equipment, networks, codecs and protocols.

The ITU Umbrella Recommendations for Transmission of Non-Telephone Signals

	H.320	H.321	H.322	H.323	H.324
Approved Date	1990	1995	1995	1996	1996
Network	Narrowband switched digital ISDN	Broadband ISDN, ATM	Guaranteed bandwidth packet-switched networks	Packet-switched networks (LAN/ WAN, ATM)	PSTN or POTS - the analog phone system
Video	H.261 H.263	H.261 H.263	H.261 H.263	H.261 H.263	H.261 H.263
Audio	G.711 G.722 G.728	G.711 G.722 G.728	G.711 G.722 G.728	G.711 G.722 G.723 G.728 G.729	G.723
Multiplexing	H.221	H.221	H.221	H.225.0	H.223
Control	H.230 H.242	H.242	H.242 H.230	H.245	H.245
Multipoint	H.231 H.243	H.231 H.243	H.231 H.243		
Data	T.120	T.120	T.120	T.120	T.120
Comm. Interface	I.400	AAL I.363 AJM I.361 PHY I.400	I.400 & TCP/IP	TCP/IP AAL	V.34 modem

Implementing H.323

With H.323 standard beginning to take root in the market, equipment vendors and software providers face the challenge of implementing the complex H.323 standard.

Hundreds of engineers from VCON, 8x8, RadVision, VideoServer and other companies are involved in the developing of high quality H.323 products with proven interoperability – from H.323 terminals through Gateways and Gatekeepers to MCUs.

Continued maintenance and updates to the technology will be available from these companies. They are active participants in standards bodies and continuously track changes to the H.323 specification.

H.323 Glossary

Application Sharing. This is a feature that allows two or more people to work together when one the individuals does not have the same application, or same version of the application. In application sharing, one user launches the application and it runs simultaneously. All users can impute information and otherwise control the application using the keyboard and mouse. Files associated with the application can be easily transferred, so the results of the collaboration are available to all users immediately. The person who launched the application can lock out the other person from making changes, so the locked-out person sees the application running but cannot control it.

Application Viewing. In personal conferencing, the users sharing the application can see every keystroke or mouse movement made by the one user who is running the application. The other users have no control over the application.

ATM. Asynchronous Transfer Mode. High speed low-delay transport technology, integrating multiple data types (voice, video, and data). ITU has selected ATM as the basis for the future broadband network because of its flexibility and suitability for both transmission and switching. May be used in the phone and computer networks of the future.

Audio. Signals that carry sounds.

Audio Bridge. Equipment that mixes multiple audio inputs and feeds back composite audio to each station after removing the individual station's input.

Automatic Bandwidth Adjustment. Algorithm in H.323 endpoint for automatic increasing and decreasing video bit rate due to the network behavior.

B channel. The ISDN circuit-switched bearer channels, capable of transmitting 64kps of digitized information.

B-ISDN. Broadband ISDN. The ITU-T is developing the B-ISDN standard, incorporating the existing ISDN switching, signaling, multiplexing and transmission standards into a higher-speed specification that will support the need to move different types of information around the public switched network.

Bandwidth. A term that defines the information carrying capacity of a channel - its throughput. In analog systems, it is the difference between the highest frequency that a channel can carry and the lowest, measured in hertz. In digital systems the unit of measure of bandwidth is bits per second.

Bit. Binary Digit. The basic signaling unit in all digital transmission systems.

Bit rate. The number of bits of information transmitted over a channel in a given second. Typically expressed bps.

Bps. Bits per second, a unit of measurement of the speed of data transmission and thus of bandwidth.

BRI. Basic Rate Interface. In ISDN there are two interfaces, the BRI and the PRI or Primary Rate Interface. The BRI offers two circuit-switched B (bearer) channels of 64 kbps each and one packet-switched 16 kbps D (delta) channel that is used for exchanging signals with the network.

Bridge. In videoconferencing vernacular, a bridge connects three or more conference sites so that they can simultaneously communicate. Bridges are often called MCUs - Multiple Conferencing Units. A bridge is also considered a device that interconnects LAN segments at the data-link layer of the OSI model to extend the LAN environment physically. They work with frames of data, forwarding them between networks. They learn station addresses and they resolve problems with loops in the topology by participating in the spanning tree algorithm. Finally, the term bridge can be used in audio conferencing to refer to a device that connects multiple (more than two) voice calls so that all participants can hear and be heard.

Broadcasting. In packet-switched networks, this means sending a packet to all users connected to the specific network.

Call. Multimedia communication between two or more H.323 endpoints.

Call Signaling Channel. Reliable channel used to convey call setup messages following Q.931

Centralized Multipoint Conference. A call in which all participating terminals communicate in a point-to-point fashion with an MCU.

Caller ID. An identification (number, name) of the party being called. This identification is of interest when you transfer or forward a call. For example, when an unanswered call is forwarded to a voice messaging system, the called-ID of the original call is used to locate the mailbox of the called party.

CCITT. Consultative Committee for International Telegraphy and Telephony. As of 1994 known as the International Telecommunications Union. See ITU.

CIF. Common Intermediate Format, an optional part of the ITU-T's H.261 and H.263 standards. CIF specifies 288 non-interlaced luminance lines, that contain 176 pixels. CIF is to be sent at frame rates of 7.5, 10, 15, or 30 per second. When operating with CIF, the number of bits that result can not exceed 256 K bits (where K equals 1024).

Circuit-switched. An ISDN bearer service that provides a 64 kbps (sometimes 56 kbps) path between two users for the duration of the call. The term is also used for the networks with behavior similar to ISDN.

CODEC. A sophisticated digital signal-processing unit that takes an analog input and converts it to digital on the sending end. At the receiving end, another codec reverses this by reconvertng the digital signal back to analog. Codec is a contraction of code/decode (some experts in the video industry assert it also stands for compress/decompress). A codec takes the form of a set of hardware or software components, or a combination of both.

Compression. Reducing the representation of the information, but not the information itself. Reducing the bandwidth or number of bits needed to encode information or encode a signal, typically by eliminating long strings of identical bits or bits that do not change in successive sampling intervals (e.g., video frames). Compression saves transmission time or capacity. It also saves storage space on storage devices such as hard disks, tape drives, and floppy disks.

Decentralized Multipoint Conference. Conference in which the participating terminals multicast to all other participating terminals without an MCU.

Document Sharing. See Whiteboard

E.164. Address format for ISDN networks. See ITU Recommendation E. 164 (1991). Added as alias for H.323 terminals.

Endpoint. A Terminal, Gateway, or MCU.

Ethernet. A LAN running on coaxial or twisted pair wiring, at 10 or 100 mbps. In Ethernet, all terminals are connected to a single common highway or bus.

Ethernet switch. A device than connects local area networks (LAN). Ethernet switching is viewed as one solution to deliver 10 Base-T or 100 Base-T networks that are bandwidth-constrained because of a new requirement to carry multimedia messages and interactive videoconferencing communications. To qualify as an Ethernet Switch, a device must be capable of switching packets from one Ethernet segment to another "on the fly" and exhibit very low port-to-port latency.

Full-duplex. A communication protocol in which the communications channel can send and receive data at the same time. Compare to half-duplex, where information can only be sent or received in one direction at a time.

G.711. An ITU-T Recommendation entitled, "Pulse Code Modulation (PCM) of Voice Frequencies". G.711 defines how a 3.1 kHz audio signal is encoded at 64 kbps using Pulse Code Modulation (PCM) and either mu-law (US and Japan) or A-law (Europe).

G.721. An ITU-T Recommendation that defines how a 3.1 kHz audio signal is encoded at 32 kbps using Adaptive Differential Pulse Code Modulation (ADPCM).

G.722. An ITU-T Recommendation that defines how a 7.5 kHz audio signal is encoded at a data rate of 64 kbps.

G.723. An ITU-T Recommendation entitled, "Dual Rate Speech Coder for Multimedia Communication Transmitting at 5.3 and 6.4 kbps".

G.728. An ITU-T Recommendation for audio encoding using Low Delay Code Excited Linear Prediction (CELP). The bandwidth of the analog audio signal is 3.4 kHz whereas after coding and compression the digitized signal requires a bandwidth of 16 kbps.

Gateway. The gateway allows H.323 systems to interoperated with other H.32x products. For instance, the gateway could link the H.323 session with an H.320 (ISDN-based) system; an H.321 (ATM-based) system; an H.322 (iso Ethernet-based) system; or an H.324 (POTS-based) system. At the present, most H.323 gateway implementations are concerned with linking H.323 and H.320/H.324 systems across a LAN/WAN connection.

GateKeeper. A gatekeeper is a utility that controls H.323 videoconference access on a packet-switched network. It requires that multimedia terminals register "at the gate", which is accomplished when the terminal provides its address. The gatekeeper translates network addresses and aliases to make connections. It can also deny access or limit the number of simultaneous connections to prevent congestion.

H.221. A framing portion of the ITU-T's H.320 Recommendation that is formally known as "Frame Structure for a 64 to 1920 kbps Channel in Audiovisual Teleservices". The Recommendations specifies synchronous operation in which the coder and decoder synchronize timing.

H.222. ITU-T Recommendation specifies generic coding of moving pictures and associated audio information.

H.223. Part of the ITU-T's H.324 standard specifying a control/multiplexing protocol, which is formally called "Multiplexing protocol for low bit rate multimedia communication".

H.230. A multiplexing Recommendation that is part of the ITU-T family of video interoperability Recommendations. The Recommendation specifies how individual frames of audiovisual information are to be multiplexed onto a digital channel.

H.231. A Recommendation added to the ITU-T's H.320 family specifying multipoint control unit used to bridge three or more H.320 compliant codecs together in a multipoint conference.

H.242. Part of the ITU-T's H.320 family of video interoperability Recommendations. H.242 specifying the protocol for establishing an audio session and taking it down after the communication has terminated.

H.245. Part of the ITU-T's H.323 and H.324 families defining control of communications between multimedia terminals.

H.261. The ITU-T's Recommendations that allows dissimilar video codecs to interpret how a signal has been encoded and compressed, and to decode and decompress that signal. It also defines two picture formats: CIF and QCIF.

H.320. An ITU-T standard including a number of individual recommendations for coding, framing, signaling and establishing connections (H.221, H.230, H.321, H.242, and H.261). It applies to point-to-point and multipoint videoconferencing sessions and includes three audio algorithms, G.711, G.722 and G.728.

H.323. The H.323 extends the H.320 to Intranet, Extranet or Internet over packet-switched networks: Ethernet, Token-Ring, and others that may not guarantee QoS. It also specifies procedures for videoconferencing over ATM including ATM QoS. It supports both point-to-point and multipoint operations.

H.323 Alias. User logical name used for remote party calling. Translated by Gatekeeper to the network address.

H.324. An ITU-T standard that provides point-to-point data, video, and audio conferencing over analog telephone lines (POTS). It can incorporate H.261 video encoding, but most implementations will probably use H.263, a scalable version of H.261 that adds a 128-by-96 Sub-QCIF (SQCIF) format. Because of H.263's efficient design, it may produce frame rates much like those of today's ISDN H.320 systems through inexpensive hardware-assisted modems. The H.324 family includes H.223, a multiplexing protocol. H.245, a control protocol, T.120, a suite of audiographics protocols and V.34, a modem specification.

IP. Internet Protocol. The most popular network protocol in corporate and public networks. May be used by H.323 endpoints for audio, video, and data packets transfer.

Interoperability. The ability of electronic components produced by different manufacturers to communicate across product line. The trend toward embracing standards has greatly furthered the interoperability process.

ISDN. Integrated Services Digital Network. ISDN is an entirely digital telephone service that can be installed by the local telephone company to replace the old analog local loop (the connection to the telephone company's nearest central switching office) with a digital line. As long-distance lines are usually digital already, replacing the local loop with an ISDN line provides "end-to-end" digital service. Two types of ISDN are: BRI and PRI. ISDN BRI is the interface to connect the desktop to the digital long distance network. ISDN BRI provides two 64 kbps B ("bearer") channel to carry information content, the voice, video, and data substance of a transmission. A separate 16 kbps D ("data") channel is used for call setup and signaling. ISDN BRI is often called "2B+D" ISDN, for its combination of two B and one D channel.

ITU. International Telecommunications Union. One of the specialized agencies of the United Nations that is composed of the telecommunications administrations of 113 participating nations. Founded in 1865 before telephone were invented as a telegraphy standards body. It now develops international standards for interconnecting telecommunications equipment across networks.

Kbps. Kilo-bytes per second - one thousand bits per second.

LAN. Local Area Network. A network of computers and other devices for communication within a restricted geographic area, such as a building or a campus.

Mbps. Megabits per second or approximately one million bits per second.

Multicasting. Sending a packet that can be received by multiple recipients, all of whom are listening on a single multicast address.

Multipoint. Communication configuration in which several terminals or stations are connected. Compare to point-to-point where communication is between two stations only.

Multipoint Controller (MC). An entity which provides for the control of three or more terminals in a multipoint conference.

Multipoint Control Unit (MCU). A device that bridges together multiple inputs so that three parties or more can participate in a video conference.

Multipoint Processor (MP). An entity which provides for the processing of audio, video, and/or data streams in a multipoint conference. The MP provides for the mixing, switching, transcoding, or other processing of media streams under the control of the MC.

Network. A group of stations (computers, telephones, or other devices) connected by communications facilities for exchanging information. Connection can be permanent, via cable, or temporary, through telephone or other communications links. The transmission medium can be physical (copper, wire, fiber optic cable etc.) or wireless, for example via satellite.

POTS. Plain Old Telephone Service. Conventional analog narrowband telephone line using twisted-pair copper wire for transmitting voice calls.

Q.931. Call signaling protocol for setup and termination of calls.

Quality of Service (QoS). Guarantees network resources for specific application requirements.

RAS Channel. An unreliable channel used to convey the Registration, Admissions and Status messages and bandwidth changes between two H.323 entities through a Gatekeeper.

Reliable Transmissions. Connection-oriented data transmission which guarantees sequenced error-free, flow-controlled transmission of messages to the receiver.

Resource Reservation Protocol (RSVP). IETF specification. Allows applications to request dedicated resources.

Real-Time Protocol/Real-Time Control Protocol (RTP/RTCP). IETF specification for audio and video signal management. Allows applications to synchronize and spoil audio and video information.

Router. Equipment that facilitates the exchange of packets between autonomous networks (LANs and WANs) of similar architecture. Routers move packets over a specific path or paths based on the packet's destination, network congestion and the protocols implemented on the network.

Switch. A device that establishes, monitors, and terminates a connection between devices connected to a network.

Switching. The process of setting up a connection between an input and an output. It allows a subscriber to establish communications with multiple parties by sending their address to the switch, which will then attempt to make a connection.

Switched Circuit Network (SCN). A public or private switched telecommunications network such as GSTN or ISDN.

Switch Type. The type of ISDN network you are connected to. This information is available from the ISDN provider and provided to the buyer when purchasing an ISDN line.

T.120. The ITU-T's "Transmission Protocols for Multimedia Data", a data sharing/data conferencing specification that lets users share documents during any H.32x videoconference. Like H.32x specifications, T.120 is an umbrella Recommendation that includes a number of other Recommendations. Data-only T.120 session can be held when no video communications are required, and the standard also allows multipoint meetings that include participants using different transmission media. The mandatory components of T.120 include recommendations for multipoint file transfer and shared-whiteboard implementation.

TCP. Transmission control protocol. A reliable transport layer on top of IP.

Teleconferencing. The use of telecommunications links to provide audio, video and graphics capabilities. These systems allow distant workgroups or individuals to meet. An endpoint which provides for real-time, two-way communications with another Terminal, Gateway, or

UDP. User Datagram Protocol. An unreliable transport layer on top of IP.

Unicast. Application of conferencing, usually over packet-switched networks, where only one user receives data. In contrast to this, multicast application, where data is received by more than one user.

Unreliable Transmission. Connection-less transmission which provides best-effort delivery of data packets. Messages transmitted by the sender may be lost, duplicated, or received out of sequence.

Videoconferencing. A collection of technologies that integrate video with audio, data, or both to convey in real-time over distance for meeting between dispersed sites.

WAN. Wide Area Network. A communications network that services a geographic area larger than that served by a local area network or metropolitan area network.

Whiteboarding. A term used to describe the placement of shared documents on an on-screen "shared notebook" or "whiteboard". Multiple users can simultaneously view and annotate a document.

Zone. In H.323 specifications, a collection of all Terminals, Gateways and MCUs managed by a single Gatekeeper. A zone must include at least one Terminal and may include LAN segments connected using routers.

