



go green with

21st Century Video

...become environmentally friendly
use green solutions

Video Conferencing Standards and Terminology

Video Conferencing Standards and Terminology.

There is an ever increasing number of standards, terminologies and buzz-words used within the video conferencing industry that can make understanding what is both available and compatible a minefield. We have the H.300's, the G.700's, the T.120's and the H.460's, not to mention ISDN, LAN, WAN, ADSL, VPN and POTS all mixed with NTSC, PAL and CIF. To complicate matters more, we also have to deal with the forthcoming media-enabled 3G mobile phone and how this links in with existing systems. This document explains what these standards, terminologies and buzz-words mean, how they relate to the various communications infrastructures of video conferencing and how they relate to each other.

It is assumed that the reader has a general knowledge of Video Conferencing systems. However, the following technical papers are available to provide more information:

- ***How do I choose a Video Conferencing system?***
- ***H.323 Terminals, Gatekeepers, Gateways & MCUs.***
- ***Global Dialling Scheme (GDS) for Schools Video Conferencing.***
- ***H.323 Dial Plan and Service Codes used by Gatekeepers etc.***
- ***IP Ports and Protocols used by H.323 Devices.***
- ***Cost Efficient ISDN Conferencing, including Multipoint Access.***
- ***H.221 Framing used in ISDN Conferences.***

International Telecommunications Union & The Internet Engineering Task Force.

Telecommunications standards are set by the United Nations agency, International Telecommunications Union (ITU) and the Internet Engineering Task Force (IETF). Products that adhere to these standards allow users to participate in a conference, regardless of their platform. These standards for desktop video conferencing ensure compatibility on a worldwide basis. The ITU has developed the H, G and T Series of standards whilst the IETF has developed Real-Time Protocol (RTP), Real-Time Control Protocol (RTCP) & Resource Reservation Protocol (RSVP).

Transport Protocols.

There are several standards based transport protocols used with conferencing, TCP, UDP & RTP. Generally, each configures the data into packets, with each packet having a 'header' that identifies its contents. The protocol used is usually determined by the need to have reliable or unreliable communications.

TCP is a reliable protocol designed for transmitting alphanumeric data; it can stop and correct itself when data is lost. This protocol is used to guarantee sequenced, error-free transmission, but its very nature can cause delays and reduced throughput. This can be annoying, especially with audio.

User Datagram Protocol (UDP) within the IP stack, is by contrast, an unreliable protocol in which data is lost in preference to maintaining the flow.

Real-Time Protocol (RTP) was developed to handle streaming audio and video and uses IP Multicast. RTP is a derivative of UDP in which a time-stamp and sequence number is added to the packet header. This extra information allows the receiving client to reorder out of sequence packets, discard duplicates and synchronise audio and video after an initial buffering period. Real-Time Control Protocol (RTCP)

is used to control RTP.

Available Transport Media.

ISDN, LAN, WAN, Internet, ADSL (Asynchronous Digital Subscriber Lines), SDSL (Synchronous Digital Subscriber Lines) and VPN, (Virtual Private Networks) are the popular transport media used in desktop video conferencing. They all have strengths and weaknesses that should be considered carefully before deciding upon which one to use. The worldwide availability of the Internet has virtually stopped the use of POTS (Plain Old Telephone Service) as a direct means of connecting video conferencing systems. However, the forthcoming media-enabled 3G mobile phone has caused the creation of a derivative of the H.324 POTS standard in the form of 3G-324M as well as next generation Gateways to transcode the new protocols.

Integrated Digital Services Network (ISDN).

ISDN supports isochronous (regular timed) data transmission and the bandwidth is guaranteed once the connection is established. With it, all information such as audio, data and video is transmitted in digital form at high speed over the public switched telephone network (PSTN). There are two available ISDN connections, Basic Rate Interface (BRI) and Primary Rate Interface (PRI). Essentially, a BRI provides two 64kbps B-channels and one 16kbps D-channel whilst a PRI in Europe provides 30 x 64kbps B-channels and one 64kbps D-channel.

ISDN connections usually aggregate the BRI and share the same number for both B channels. Known as ISDN-2, this provides a line speed of 128kbps is typically used in a desktop conference over ISDN. For increased bandwidth, ISDN-6 provides a line speed of 384kbps and is typically used in room-based conferences over ISDN. With ISDN-6, the sequence in which the lines are aggregated must be known and adhered to! Furthermore, if the connection is going to use some form of 'switch', this must be configured to pass both voice and data!

In the past, most conferences would have been between just two participants as ISDN is essentially a point-to-point connection. However, multipoint technology now makes it possible for groups of people to participate in a conference and share information. To hold a multipoint conference over ISDN, participants use a Multipoint Control Unit (MCU), that connects and manages all the ISDN lines. This can be either a separate MCU or an endpoint with an embedded H.320 multipoint capability.

H.320 is the ITU standard for ISDN conferencing and includes:

Audio: G.711, G.722, G.722.1, G.728, AAC-LC, AAC-LD

Video: H.264, H.263, H.261

Data: H.239

Control: H.221, H.231, H.242, H.243

Local Area Network (LAN) or Intranet and Wide Area Network (WAN).

100 Mbps LANs with switches and routers are used in most companies today and these have enough bandwidth to support desktop conferences. With a LAN offering significantly more bandwidth than ISDN, the video quality within a conference is much higher and can approach that of television. Technology as also helped, we now have communications advancements such as Gigabit Ethernet (1000 Mbps), Faster Switches, Asynchronous Digital Subscriber Lines (ADSL), Synchronous Digital Subscriber Lines (SDSL) and Virtual Private Networks (VPN) that have increased and/or secured bandwidth, whilst IP Multicasting has reduced network loading in conferences involving more than two participants.

Unlike ISDN networks, LANs and WANs use TCP/IP protocol and the H.323 standard defines how to assemble the audio, video, data and control (AVDC) information into an IP packet. Most companies use DHCP and allocate dynamic IP addresses to PC's. Therefore, in order to correctly identify a user, the H.323 endpoints are usually registered with a Gatekeeper and 'called' into a conference by their H.323 alias. The Gatekeeper translates the alias into the corresponding IP address. Another method of identifying H.323 users is for them to register their presence using Light Directory Access Protocol (LDAP) with a Directory Service such as Microsoft's Site Server ILS or Windows Active Directory.

To hold a multipoint conference over IP, H.323 systems require some form of Multipoint Conference Server (MCS). This is also referred to as an H.323 Multipoint Control Unit (H.323 MCU), which is not the same as an H.320 MCU; hence it is important to be clear about what you mean when using the term MCU. To hold a large scale multipoint conference over IP, participants must use a separate dedicated MCU connected to the IP network. For small scale multipoint conferences, there are now endpoints with an embedded H.323 multipoint

capability that support up to 6 endpoints in a single conference.

H.323 is the ITU standard for LAN conferencing and includes:

Audio: G.711, G.722, G.722.1, G.723.1, G.728, G.729, AAC-LC, AAC-LD

Video: H.264, H.263, H.261

Data: H.239

Control: H.225, H.245, H.460

Cellular Networks.

The cellular phone network is a readily available form of wireless multimedia delivery and with the forthcoming media-enabled 3G mobile phone or Personal Digital Assistants, PDAs, that support the CDMA2000 or WCDMA Air Interface, there is sufficient bandwidth to enable IP-based multipoint audio and video conferencing to existing desktop video conferencing systems when used in-conjunction with next generation Gateways and MCU's that also support these new protocols.

3G-324M is an extension by the 3rd Generation Partner Project (3GPP) and 3rd Generation Partner Project2 (3GPP2) to the ITU H.324M standard for 3G mobile phone conferencing and includes:

Audio: G.722.2 (AMR-WB), G.723.1

Video: MPEG-4, but not H.264

Control: H.223 A/B, H.245

Internet, ADSL, SDSL & VPN.

With its ever increasing popularity, people have sought to use the Internet in more ways than just a means of sending email or browsing interesting sites.

Like LANs, the Internet, ADSL, SDSL and VPNs are other forms of TCP/IP based networks and hence can be used as a transport media in desktop conferencing systems. Not to be confused with POTS, the Internet uses a modem as a TCP/IP dial-up adapter in order to gain access to the network. What the users must do is to get their Internet Service Provider (ISP) to provide them with a fixed IP address. Alternatively, users can register their presence using LDAP with a Directory Service such as Microsoft's Site Server ILS or Windows Active Directory. This is how you determine the address of the machine that you want to conference with. Obviously, speed is limited to that of the slowest link, but most ISPs support ISDN Dial-up at 128kbps as well as V.92 modems at 56kbps.

For a more secure and faster connection, ISP and telecoms companies are now offering VPN over ADSL and SDSL links. A VPN provides a secure tunnel over the providers network by applying encryption between sites. With most Firewalls supporting VPN pass-thru, there is no need to open lots of ports. However, be wary of applying too much encryption as this can cause an unacceptable delay in the transmission between sites.

ADSL and SDSL, whilst being faster than ISDN, are only as fast as the slowest uplink when used for Video Conferencing. Again, users should get their Service Provider to provide them with a fixed IP address for their xDSL Modem/Router/Firewall. Most xDSL Modems now incorporate a Router and Firewall. Depending upon whether the Video Conferencing system is PC or non-PC based, it can either be located behind an **H.323 Intelligent Firewall or Proxy** (PC-based) or outside in the DMZ (non-PC based). Otherwise, too many Firewall ports may have to be opened in order to provide access, which defeats the objectives of having a Firewall. Alternatively, some newer xDSL Modem/Router/Firewalls now support Universal Plug and Play (UPnP). This feature when used with UPnP enabled endpoints negotiates opening just the required ports.

H.323 is the ITU standard used for Internet conferencing and includes:

Audio: G.723.1, G.722.1, G.728

Video: H.264, H.263, H.261

Data: H.239

Control: H.225, H.245, H.460

Video standards:

H.261 - video codec for audiovisual services at p x 64Kbps.

H.263 - video codec for narrow telecommunications channels at < 64 Kbps.

Notable elements of the standard are image size. QCIF is Quarter Common Intermediate Format and represents a 176x144 pixel image. This is the minimum size that must be supported to be H.320 compliant. CIF is the optional full- screen H.320 video image of 352x288 pixels and requires considerably more computing capability.

Note: whilst this is termed full-screen, it is nowhere near the size of a typical PC screen (1024x768) pixels or that of a UNIX workstation (1280x1024) pixels.

H.264/AVC - a new video codec standard offering major improvements image quality.

Ratified in late 2003, this new codec standard was a development between the ITU and ISO/IEC Joint Video Team, (JVT) and is known as H.264 (ITU name) or ISO/IEC 14496-10/MPEG-4 AVC (ISO/IEC name).

This new standard surpasses H.261 and H.263 in terms of video quality, effective compression and resilience to transmission losses, giving it the potential to halve the required bandwidth for digital video services over the Internet or 3G Wireless networks. H.264 is likely to be used in applications such as Video Conferencing, Video Streaming, Mobile devices, Tele-Medicine etc. Current 3G mobiles use a derivate of MPEG-4, but not H.264.

Audio standards:

G.711 - Pulse Code Modulation of voice frequencies (PCM), where 3.1 kHz analogue audio is encoded into a 48, 56 or 64 kbps stream. Used when no other standard is equally supported.

G.722 - 7 kHz audio encoded into a 48, 56 or 64 kbps stream. Provides high quality, but takes bandwidth.

G.722.1 - 7 kHz audio encoded at 24 and 32 kbps for hands-free operation in systems with low frame loss.

G.722.1 Annex C - The ITU's adoption of Polycom's Siren 14 - a 14 kHz audio codec.

G.722.2 - Coding of speech at around 16 kbps using Adaptive Multi-Rate Wideband, AMR-WB. Five mandatory modes, 6.60, 8.85, 12.65, 15.85 and 23.85 kbps.

G.723.1 - 3.4 kHz dual rate speech codec for telecommunications at 5.3 kbps & 6.4 kbps.

G.728 - 3.4 kHz Low Delay Code Excited Linear Prediction (LD-CELP) where 3.4 kHz analogue audio is encoded into a 16 kbps stream. This standard provides good quality results at low bitrates.

G.729 A/B - 3.4 kHz speech codec that provides near toll quality audio encoded into an 8 kbps stream using the AS-CELP method. Annex A is a reduced complexity codec and Annex B supports silence suppression and comfort-noise generation.

Data and Control standards:

H.221 - defines the transmission frame structure for audiovisual teleservices in channels of 64 to 1920 Kbps; used in H.320

H.223 - specifies a packet-orientated multiplexing protocol for low bit rate multimedia communications; Annex A & B handles light and medium error prone channels of the mobile extension as used in 3G-324M.

H.224 - defines real-time control protocol for simplex applications using the H.221 LSD, HSD and HLP channels.

H.225 - defines the multiplexing transmission formats for media stream packetisation & synchronisation on a non-guaranteed QoS LAN.

H.231 - specifies multipoint control units used to bridge three or more H.320 systems together in a conference.

H.233 - Confidentiality systems for audiovisual services, used by H.320 devices.

H.234 - Encryption key management and authentication system for audiovisual services, used by H.320 devices.

H.235 - Security and encryption for H.323 and other H.245 based multimedia terminals.

H.239 - defines role management and additional media channels for H.300-Series multimedia terminals. How data and web-enabled collaboration work in parallel with video in a conference, allowing endpoints that support H.239 to receive and transit multiple, separate media streams - typically voice, video and data collaboration.

H.241 - defines extended video procedures and control signals for H.300-Series multimedia terminal.

H.242 - defines the control procedures and protocol for establishing communications between audiovisual terminals on digital channels up to 2 Mbps; used by H.320.

H.243 - defines the control procedures and protocol for establishing communications between three or more audiovisual terminals - H.320 multipoint conferences.

H.245 - defines the control procedures and protocol for H.323 & H.324 multimedia communications.

H.246 - Interworking of H-Series multimedia terminal.

H.248 - Gateway Control Protocol.

H.281 - defines the procedures and protocol for far end camera control (FECC) in H.320 calls.

H.282 - Remote device control protocol for multimedia applications.

H.283 - Remote device control logical channel transport.

H.350 - Storing and retrieving video and voice over IP information from enterprise directories.

ANNEX Q - defines the procedures and protocol for far end camera control (FECC) in H.323 calls.

Supplementary Services:

H.450.1 - defines the generic functional protocol for support of supplementary services in H.323.

H.450.2 - defines the Call Transfer supplementary services for H.323.

H.450.3 - defines the Call Diversion supplementary services for H.323.

H.450.4 - defines the Call Hold supplementary services for H.323.

H.450.5 - defines the Call Park and Call Pickup supplementary services for H.323.

H.450.6 - defines the Call Waiting supplementary services for H.323.

H.450.7 - defines the Message Waiting Indication supplementary services for H.323.

H.450.8 - defines the Name Identification supplementary services for H.323.

H.450.9 - defines the Call Completion supplementary services for H.323.

H.450.10 - defines the Call Offer supplementary services for H.323.

H.450.11 - defines the Call intrusion supplementary services for H.323.

H.450.12 - defines the Common Information Additional Network Feature for H.323.

Mobility:

H.501 - Protocol for mobility management in multimedia systems.

H.510 - Mobility for H.323 multimedia systems and services.

H.530 - Symmetric security procedures for H.323 mobility in H.510.

Connectivity:

BONDING - Bandwidth ON Demand Interoperability Group, synchronises the B-channels to transmit as one stream and attain higher data rates.

DID - Direct Inward Dialling is a method of routing H.320 incoming calls directly to H.323 endpoints without operator intervention.

DTMF - Dual Tone Multi-Frequency signals are the type of audio signals used in telephony for tone dialling.

E.164 Number - (User Number). A numeric string given to an H.323 endpoint. If this endpoint registers with a Gatekeeper, then the Gatekeeper can translate the E.164 Number into the endpoints IP address.

H.323 Alias - A logical name given to an H.323 endpoint. If this endpoint registers with a Gatekeeper, then the Gatekeeper can translate the H.323 Alias into the endpoints IP address.

IVR - Interactive Voice Response is a two-stage DID method of routing H.320 calls that is supported by the Gateway. It enables an H.320 endpoint to directly contact an H.323 endpoint using DTMF tones to control the connection.

LDAP - Light Directory Access Protocol. Use by H.323 endpoints to register their presence with Directory Services.

MSN - Multiple Subscriber Numbering. When the PSTN Company assigns a group of telephone numbers to one line.

Q.931 - Signalling protocol for establishing and terminating calls.

RAS - Registration/Admission/Status. A communications protocol used between H.323 endpoints and the Gatekeeper for registration, admission and status messages.

RTP/RTCP - Real-Time Protocol/Real-Time Control Protocol. An IETF specification for audio and video signal management. Allows applications to synchronize audio and video packets.

SIP - Session Initiation Protocol.

TCS-4 - Terminal Control Strings are another DID method of routing H.320 calls that is supported by the Gateway. The TCS-4 string contains information that is used to identify the H.323 endpoint, such as its E.164 number.

Video and PC Window Sizes:

NTSC - National Television Standards Committee, used in USA, Canada & Japan. 640 x 480 pixels.

PAL - Phase Alternation by Line, used in Europe (except France), Africa & Middle East. 768 x 576 pixels.

SECAM - Sequentielle Couleur Avec Memoire, used in France & Russia.

CIF - Common Intermediate Format; optional for both H.261 & H.263, 352 x 288 pixels.

QCIF - Quarter Common Intermediate Format; required by both H.261 & H.263, 176 x 144 pixels.

SQCIF - Sub Quarter Common Intermediate Format; used by 3G mobiles MPEG4 video and H.263, 88 x 72 pixels.

SXGA - 1280 x 1024 pixels - used by high end graphics workstations.

XGA - 1024 x 768 pixels - typical PC or laptop resolution.

SVGA - 800 x 600 pixels.

VGA - 640 x 480 pixels.

Misc:

RTFM - A diagnostic instruction generally given to eager installers of hardware and software when things don't quite work **Read the FFFF.... Manual.!**