TANDBERG

See: productivity

See: passion

See: performance

Video Conferencing Protocols

Welcome to the TANDBERG University prerequisite Video Conferencing Protocols. Before commencing you are requested to ensure that you have completed the Introduction to the TANDBERG University eLearning Experience Module that is available through this portal.

Lesson: Video Conferencing Protocols

Learning Objectives

On completion of this lesson, you will demonstrate an understanding of some of the protocols which are used in a video conference. You will know how to:

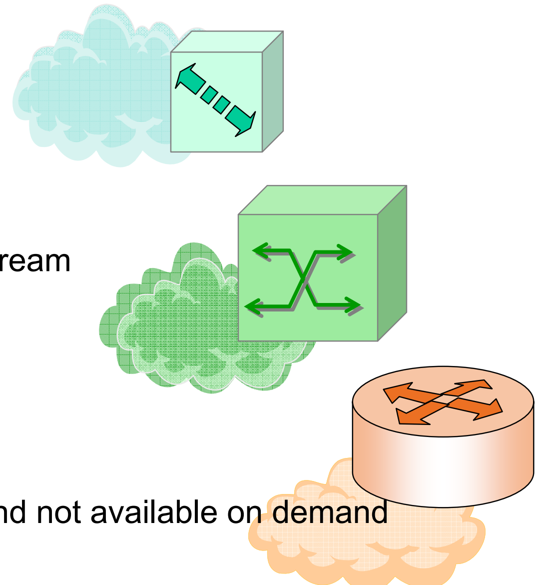
- Explain the difference between H.320, H.323 and SIP
- Explain bandwidth and it's effect on video/audio quality and which standards are used in a call
- Describe the differences between G.711, G.722, G.722.1, G.728, and AAC-LD
- Explain the differences between H.261, H.263, H.264, QCIF, CIF, SIF, 4CIF, 4SIF, iCIF, iSIF, and HD

TANDBERG See: knowledge slide 2 Video Conferencing Protocols

Video Conferencing Protocols

Circuit Switched

- Bandwidth is guaranteed
- Bandwidth is not shared
- On demand
- Information is in a single bit stream



Packet Switched

- Bandwidth is not guaranteed
- Bandwidth is shared
- Circuits are built as needed and not available on demand
- Information is in packets

TANDBERG See: knowledge slide 3 Video Conferencing Protocols

Video conferencing can be run across two types of networks. These are called ‘Circuit Switched’ and Packet Switched’.

Circuit Switched networks have the following characteristics:

•Bandwidth is guaranteed through the network

•Bandwidth is not shared once connection is established

•Just like a phone call – you bring up a circuit through the network on demand

•Information is sent as a single bit stream

Packet Switched networks have the following characteristics:

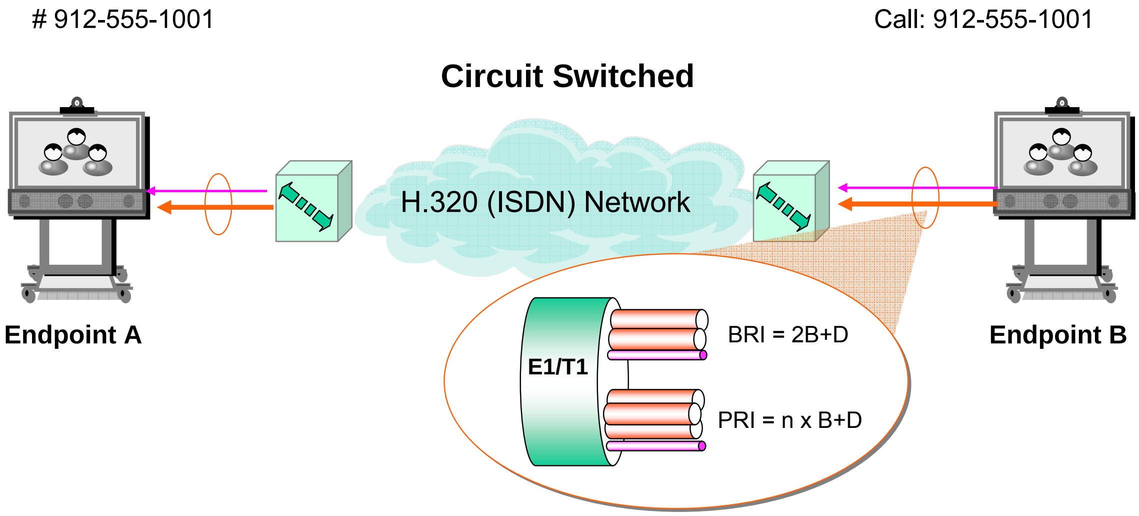
•Bandwidth is NOT guaranteed

•Bandwidth IS shared

•Circuits are built as needed and are not available on demand

•Information is sent after being broken down into packets

Video Conferencing Protocols: H320



TANDBERG See: knowledge slide 4 Video Conferencing Protocols

On most TANDBERG codecs and infrastructure devices, there is a choice of which type of network configuration you are using to place video calls.

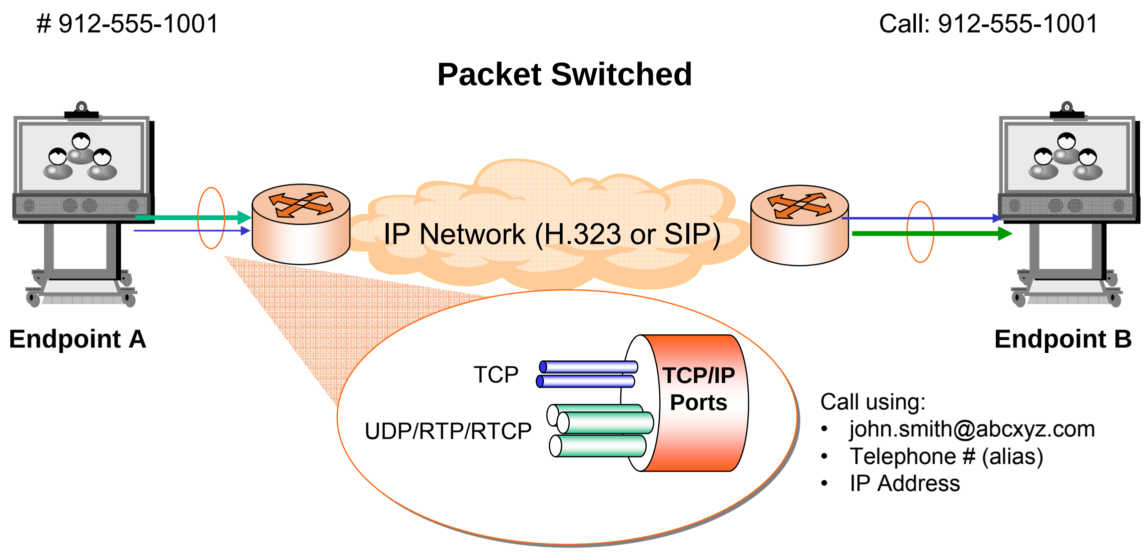
The three choices are:

- H.320
- H.323
- and Session Initiation Protocol (SIP)

Often there is also a choice to configure an “IP” network. This option often refers to H.323.

H.320 defines how circuit switched networks are used in video communications. By far the most common circuit switched network is ISDN or Intergrated Services Digital network. The H.320 standard includes its signaling mechanisms and how voice, video, and other payload are transmitted over the ISDN interface.

Video Conferencing Protocols: H320 vs IP

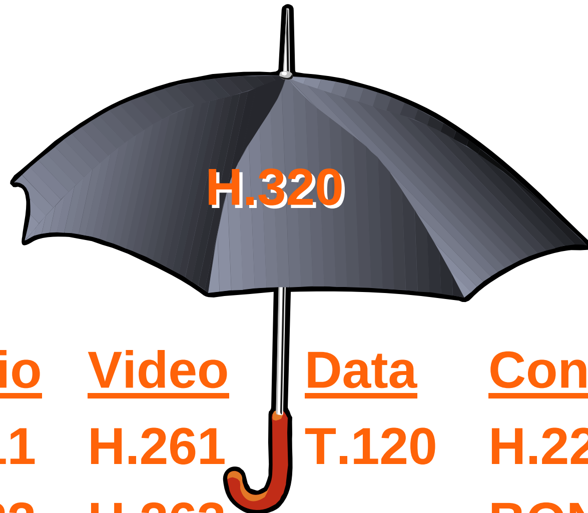


TANDBERG See: knowledge slide 5 Video Conferencing Protocols

IP configuration often refers to H.323, however both H.323 and SIP are signaling protocols that operate on IP based networks. Like the H.320 protocol, the H.323 standard includes its signaling mechanisms and how voice, video are transmitted.

SIP, however, does not handle the voice, video, and other payload; SIP just defines signaling procedures which are used to set up, maintain, and tear down the IP connections that carry the voice, video, and other payload signals.

Standards



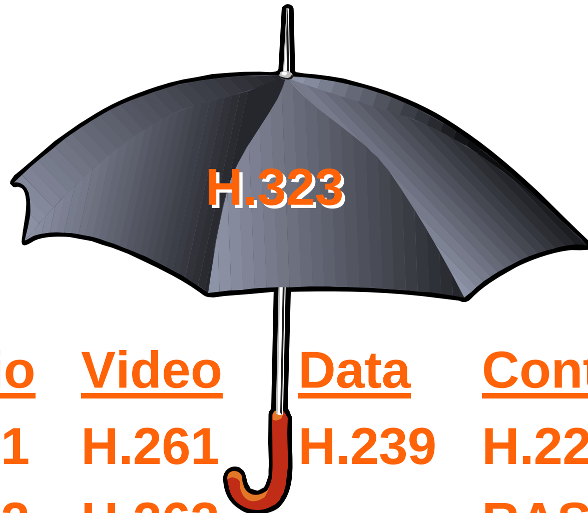
Audio	Video	Data	Control
G.711	H.261	T.120	H.221
G.722	H.263		BONDing
G.728	H.264		

TANDBERG See: knowledge slide 6 Video Conferencing Protocols

In order for it to be a success, video conferencing has to be built around standards so that systems from different manufacturers can communicate with each other. There are a huge range of standards within video conferencing and some of these we need to know about. H.320 is the 'umbrella' standard that dictates how video conference traffic runs over Circuit Switched Networks which include ISDN networks. An umbrella standard is basically a group of standards that when working together achieve the end goal. In this case, the H.320 standard is a group of standards that govern all the parts of our video conference, for example how we compress our video signal or how we display our video picture.

H.320 has a number of standards which are 'required' in order for a system to be 'standards compliant' and a number of standards that are 'optional'. Since most of the 'optional' standards are later developments and improvements to the original 'required' standards they are not really optional if you want to produce a market leading product. However it is essential that the 'required' standards are also present so that new equipment can still communicate with old equipment.

Standards

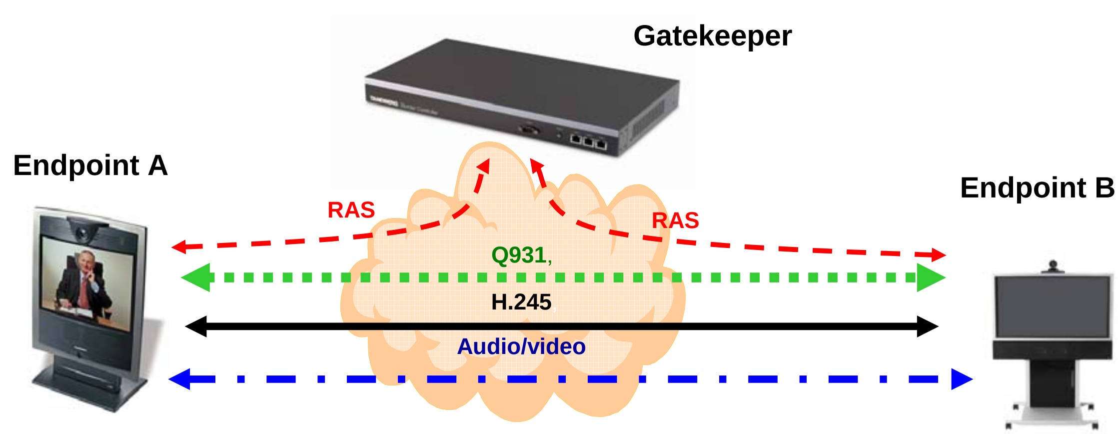


Audio	Video	Data	Control
G.711	H.261	H.239	H.225
G.722	H.263		RAS
G.728	H.264		Q.931
			H.245

TANDBERG See: knowledge slide 7 Video Conferencing Protocols

Just as the H.320 standard is an umbrella standard, so too is the H.323 standard. You may notice that a number of the standards shown above under the umbrella are in fact the same as those within the H.320 standard.

Video Conferencing Protocols: H.323



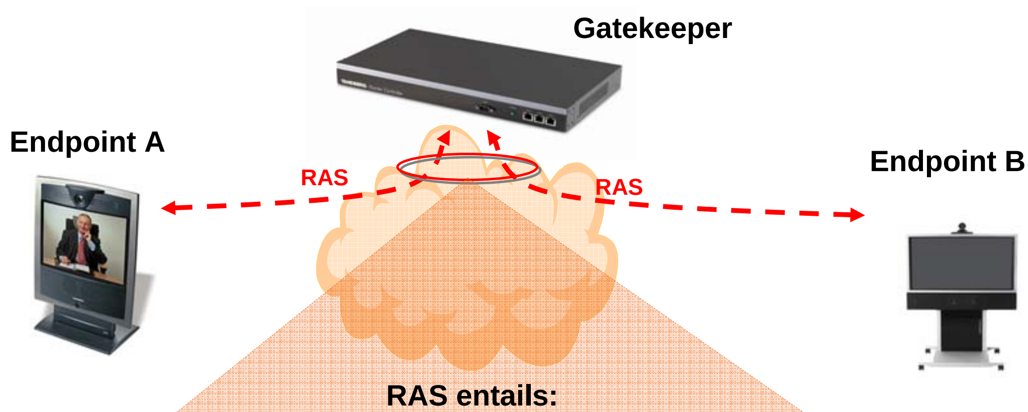
TANDBERG See: knowledge slide 8 Video Conferencing Protocols

H.323 is not a single protocol. It is actually a series of protocols that together create the “umbrella” of H.323. Although H.323 does enable endpoints to establish point to point connections without one; it relies on the centralized intelligence of a gatekeeper to control the communications in a multi-point session. Two of the signaling protocols under the H.323 umbrella are H.225 and H.245:

- H.225 admits endpoints onto the video network and establishes connections between endpoints. Two main components to H.225 are Q.931 and RAS.
- Registration Admission Status (RAS)- RAS allows control of the video network as a whole one zone at a time. It gives gatekeepers the ability to manage the devices in a zone. Used between endpoints and gatekeepers.
- Q.931 is responsible for establishing and tearing down connections between H.323 endpoints.

- Once the connections have been established, H.245 negotiates with the communicating devices a Capabilities Exchange (CapEx). H.245 establishes which protocols an endpoint supports, which protocol combinations that endpoints support, which protocols an endpoint prefers, the best possible bit rate available to all endpoints, the master-slave relationships between endpoints, etc.

Call Signaling with H.323 - RAS



RAS entails:

- **RRQ (E.164 alias, H.323 ID and IP address)**
- **RCF**
- **RRJ**
- **ARQ (Called units E.164 alias/H.323 ID, Source dial digits and source IP address and requested bandwidth)**
- **ACF**
- **ARJ**
- **ARQ**

TANDBERG See: knowledge slide 9 Video Conferencing Protocols

The RAS Process follows the sequence below:

A Registration Request (RRQ) is sent to identify the endpoint that wishes to be admitted to the network. An RRQ contains the following addresses that can be used to contact the endpoint:

- E.164 alias
- H323id
- IP address

The gatekeeper sends a Registration Confirm (RCF) to confirm that endpoint's entry into the video network.

If for any reason the endpoint is not admitted, it will send a Registration Reject (RRJ) to the endpoint.

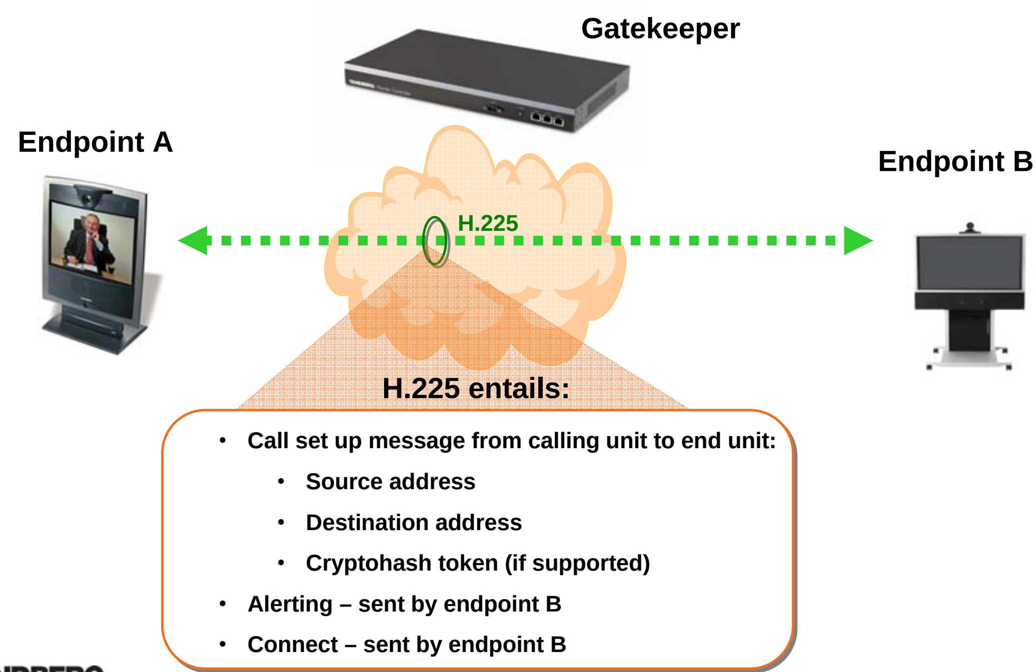
In order to set up a video connection, an endpoint will send an Admission Request (ARQ). An ARQ contains:

- Called units E.164 alias/H323id
- Requested call bandwidth

The result of an ARQ could be an Admission Confirm (ACF) to confirm the connection, provide the remote IP address and the call's bandwidth to the calling endpoint; or an Admission Reject (ARJ) to deny the connection.

The called endpoint also sends an Admission Request (ARQ) to accept an incoming call.

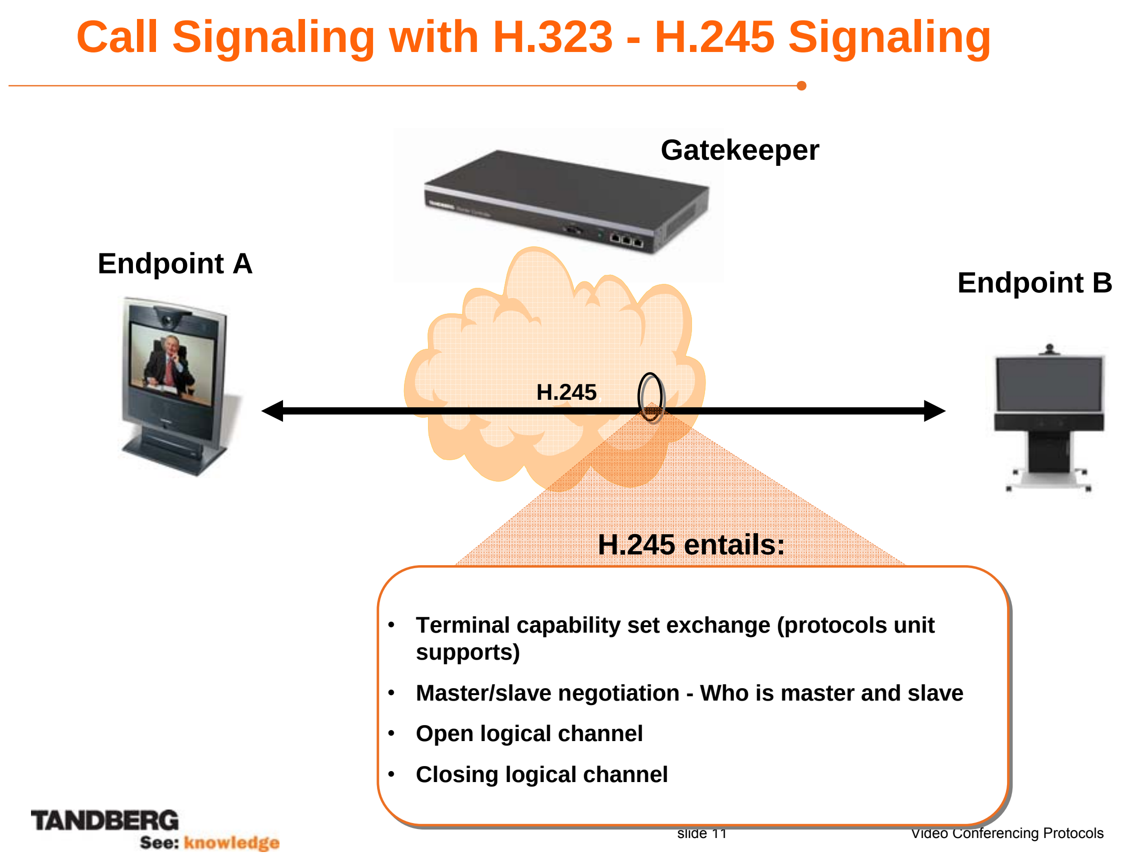
Call Signaling with H.323 – H.225 (Call Setup)



TANDBERG See: knowledge slide 10 Video Conferencing Protocols

During call setup, messages are sent from the calling unit to the called unit. Some of the important information that these messages contain are:

- Source address (E.164 alias/H323id)
- Destination address (E.164 alias/H323id)
- Cryptohash token if encryption is supported (AES/DES)



Part of the connection process is to identify the capabilities of each device participating in the video conferencing session. Also know sometimes as CapEx, for Capabilities Exchange; terminal capability set exchange exchanges information about the type of protocols the H.323 units supports (G711 and H263 etc), the combination of protocols that the unit supports, and the protocol combinations that are preferred by the endpoint.

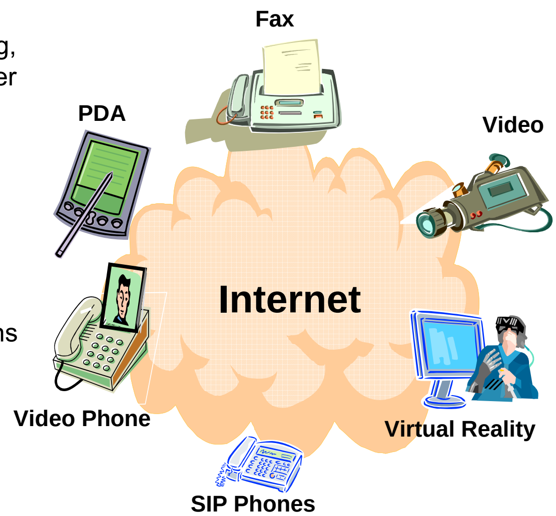
In videoconferencing, it is necessary to identify a "Main" endpoint so that the endpoints all know the role that they must play to successfully complete a call. H.245 negotiates this master-slave relationship between the devices. If there is an MCU/MPS involved in the connection, then the MCU will always be the master device.

H.245 is also responsible for opening and closing the logical channels that are used to carry the video and audio information. These logical channels will usually contain Real-time Transport Protocol (RTP) over User Datagram Protocol (UDP) and Internet Protocol (IP), as discussed in the *TANDBERG and IP* prerequisite learning session.

Session Initiation Protocol (SIP) Overview

SIP:

- Protocol developed for initiating, modifying, and terminating interactive multimedia user sessions such as the following:
  - voice
  - video
  - instant messaging
  - online games
  - virtual reality
- Provides all of the call processing functions and features of PSTN.



SIP is used strictly for signaling. It does not carry any voice, video, or streaming information; it sets up and enables that information transfer. SIP relies less on centralized intelligence then does H.323 due to its simplicity, however; it is commonly deployed and integrated into a network using the same architecture as H.323 with a SIP registrar in place of a gatekeeper.

SIP was defined by the Internet Engineering Task Force and is therefore not a standard; unlike H.323 which is a standard.

SIP is a simplified signaling protocol as compared to H.323; there are only six types of SIP Requests (i.e. signaling messages):

•Invite

•Ack(knowledge)

•Options

•Bye

•Cancel

•Register

SIP also offers a more flexible signaling platform than H.323 by including manufacturer, user defined, and experimental fields to the signaling messages for greater capabilities and easier customization of features. This flexibility is also a major drawback. It creates compatibility issues between vendors and has even been the catalyst for the evolution of multiple SIP "standards".

SIP Registrar

- Provides Registration for SIP endpoints
- Requires SIP Domain(s) to be configured on TANDBERG VCS
- Proxy registrations to SIP registrar: The proxy can forward the registration to your ”home domain” registrar
- Will not act as a SIP registrar if domains are not configured

TANDBERG
See: knowledge

slide 13 Video Conferencing Protocols

Using a SIP Registrar

The registrar provides registration for SIP endpoints. In order for a SIP endpoint to be contactable via its registered alias, it must register its location with a SIP Registrar.

- SIP URI’s always take the form X@Y where X indicates the unit and Y the domain. Both values can be a name or an IP address
- To a SIP Registrar, you must configure it with the SIP Domain(s) for which it will be authoritative. Registrars will accept registration requests for any endpoints attempting to register with an alias that includes that domain.
- Proxy registrations to a SIP registrar will forward the registration to your ”home domain” registrar

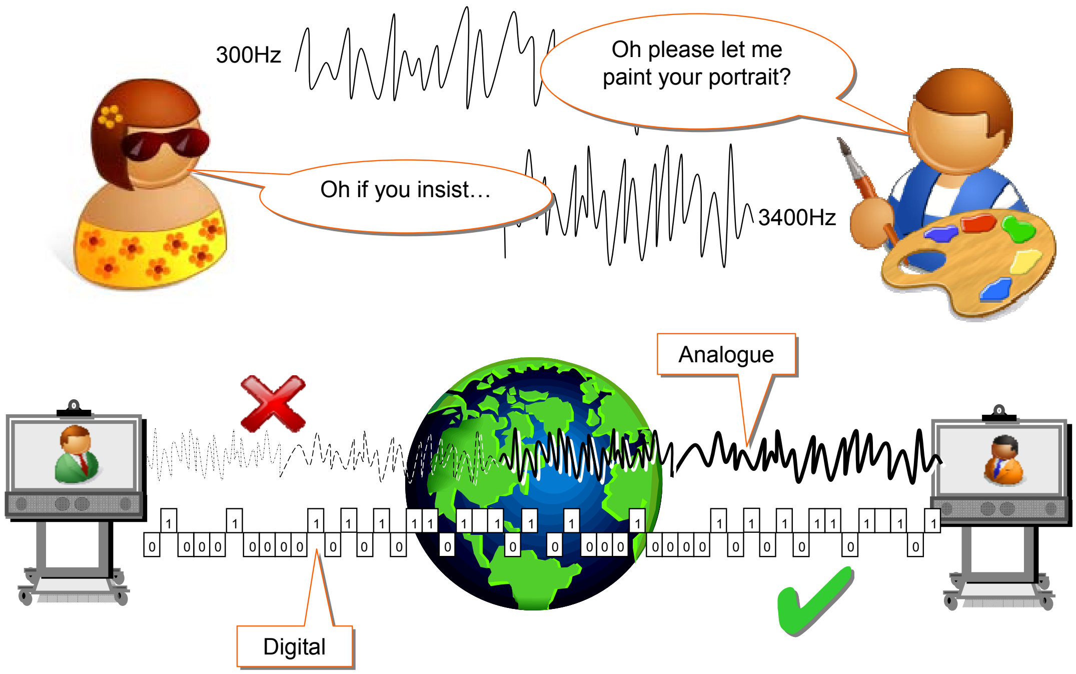
SIP Proxy

- Forward requests (such as REGISTER and INVITE) from endpoints or other Proxy Servers.
- Requests forwarded on to other Proxy Servers or to the destination endpoint.
- Behavior determined by the SIP Registration Proxy Mode setting

Using a SIP Proxy Server

The role of a Proxy Server is to forward requests (such as REGISTER and INVITE) from endpoints or other Proxy Servers.

Video Conferencing Protocols: Video and Audio



slide 15 Video Conferencing Protocols

Video and audio communications are primarily about humans communicating with humans. Humans both speak with and listen to analog signals. Most of this human to human communication takes place between 300 Hz and 3400 Hz, although other frequencies can be and are used as well.

Video and audio signals are most accurately recreated and transmitted over very short distances as an analog signal. However, most of the time audio and video signals are transmitted; they are transmitted over long distances, possibly thousands of miles when the signal reaches its destination. Over long distances, analog signals lose power and gain large amounts of noise making the signal received at the destination useless. For long distance transmissions; digital media are far superior. Digital bit streams can easily be recreated periodically over the entire transmission without gaining a significant amount noise.

Some bandwidth limitations can be overcome by using the right coding protocol. A protocol's variables, such as, compression technique, error correction, coding scheme, and ability to extrapolate (predict) can improve signal quality by reducing the need for bandwidth, correcting transmission errors, and/or reducing the amount of information that needs to be transmitted.

Video Conferencing Protocols: Audio Coding

Protocol	Bit Rate (Kbs) 8Kbs ↔ 128Kbs	Audio Limits	Delay	Quality
G.711	64	4 kHz	40 ms	Fair
G.722	64 56 48	7 kHz	40 ms	Good
G.722.1	48 32 24	7 kHz	40 ms	Good
G.728	16	4 kHz	625 µs	Fair
AAC-LD	48 to 128	20 kHz	20 ms	Excellent-Stereo

G.711 digitizes analog signals using a process called Pulse Code Modulation (PCM). The output of this process is an excellent quality 64000 bit per second digital representation of the original signal with low delay. This is the minimum requirements on all codecs.

G.722 digitizes analog signals using a variation of the PCM process. This process is called Sub-Band Adaptive Differential Pulse Code Modulation (SB-ADPCM). This modulation method allows G.722 to represent up to 7000 Hz audio with a digital 64000 bps, 56000 bps, or 48000 bps depending on the mode.

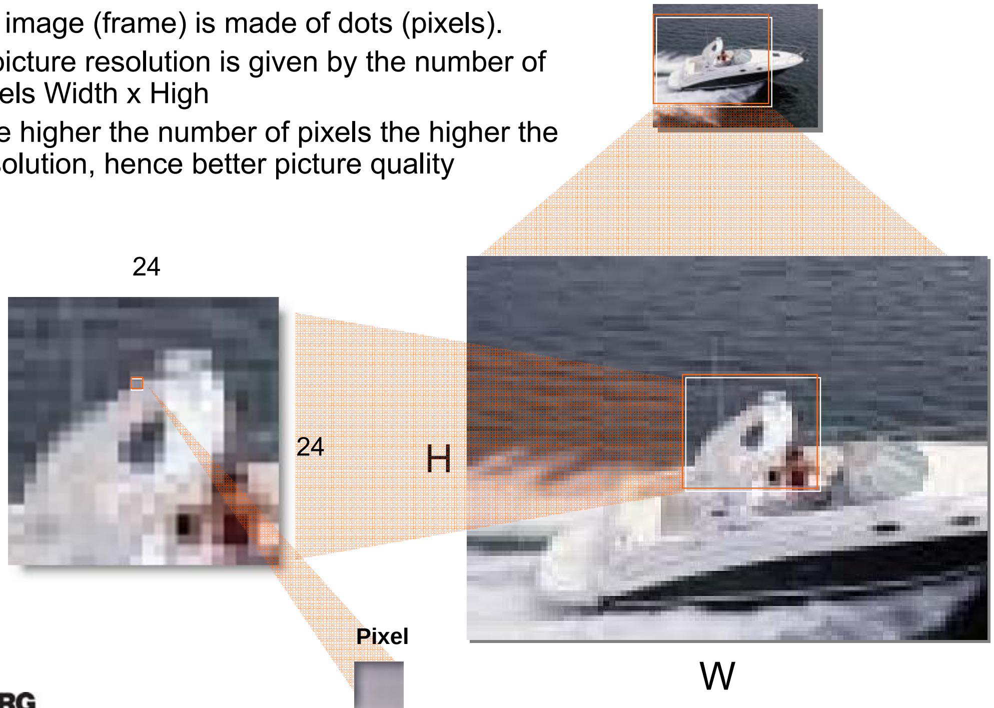
G.722.1 describes a low-complexity extension mode to G.722, which permits 14 kHz audio to be transmitted at 24, 32, and 48 Kbps. This mode provides the same 40 ms delay as the 7 kHz mode.

G.728- There have been a number of patterns identified by audio engineers that can be used to predict human voice sequences. The process that is used to make this prediction is called Code Excited Linear Prediction (CELP). G.728 uses Low-Delay Code Excited Linear Prediction (LD-CELP) to encode analog signals. G.728 can transmit voice using only a 16000 bps signal. The delay for G.728 is 625 µs.

AAC-LD- Advanced Audio Coding- Low Delay (AAC-LD) is a method that encodes stereo audio. It is contained within Motion Picture Experts Group 4 (MPEG4) standard. AAC-LD provides high quality and low delay 20 kHz audio.

Frame

- An image (frame) is made of dots (pixels).
- A picture resolution is given by the number of pixels Width x High
- The higher the number of pixels the higher the resolution, hence better picture quality



TANDBERG See: knowledge

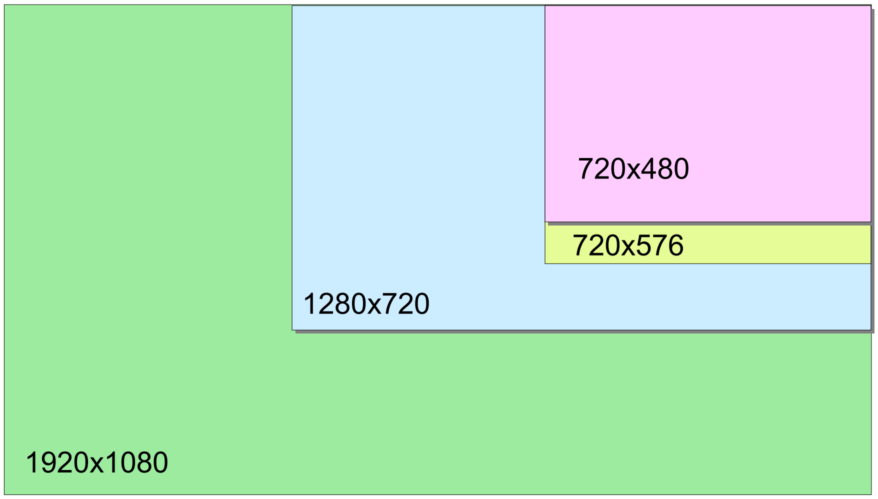
slide 17 Video Conferencing Protocols

•An image (frame) is made of dots (pixels)

•A picture resolution is given by the number of pixels Width x High

•The higher the number of pixels the higher the resolution, hence better picture quality

Video Conferencing Protocols: Video Resolution



TANDBERG See: knowledge slide 18 Video Conferencing Protocols

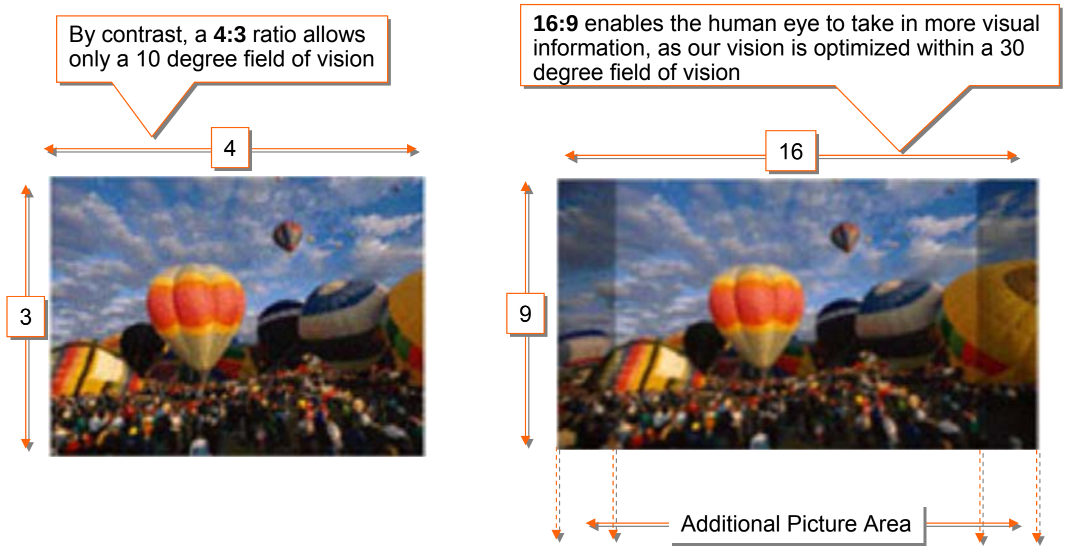
One definition of resolution is, "The fineness of detail that can be distinguished in an image, as on a video display terminal."

This fineness is influenced by many variables, among them are:

- the number of pixels per line
- the number of lines
- the number of image frames per second
- the scan mode
- and the format.

Video Conferencing Protocols: Video Format

The new large digital screens are approximately 1/3 larger than standard NTSC sets, and can display a wider aspect ratio - 16:9



slide 19 Video Conferencing Protocols

While building an image, a video monitor will create the image using one of two different formats; progressive and interlaced.

Progressive- each horizontal line of the image is created in successive order.

Interlaced- every other horizontal line of an image is populated, once these lines are completed, the alternating lines are created.

4:3 and 16:9 formats:

- 4:3; 4 units wide by 3 units high.
- 16:9; 16 units wide by 9 units high

1080i is an HDTV standard referring to a signal with a resolution of up to 1080 by 1920 pixels. It is a 16:9 signal using interlaced build-up of the signal's lines

720p is also an HDTV display standard referring to a signal with a resolution of up to 720 by 1280 pixels. It is a 16:9 signal using progressive build-up of the signal's lines

480p is a SDTV display standard referring to a signal with a resolution of up to 480 by 720 pixels. It is a 4:3 signal using progressive build-up of the signal's lines

Video Conferencing Protocols: Video Standards

System	PAL B,G,H	PAL I	PAL D	PAL N	PAL M	NTSC M
Lines/Field	625/50	625/50	625/50	625/50	525/60	525/60
Horizontal Frequency (Khz)	15.625	15.625	15.625	15.625	15.750	15.734
Vertical Frequency (Hz)	50	50	50	50	60	60
Color Sub Carrier Frequency (MHz)	4.433618	4.433618	4.433618	3.582056	3.575611	3.579545
Video Bandwidth (MHz)	5.0	5.5	6.0	4.2	4.2	4.2
Sound Carrier (MHz)	5.5	6.0	6.5	4.5	4.5	4.5

TANDBERG See: knowledge slide 20 Video Conferencing Protocols

Short for Phase Alternating Line, PAL is the dominant television standard across Europe, delivering 625 lines at 50 half-frames per second.

Short for National Television System Committee, NTSC is the United States committee responsible for creating technological television and video standards.

These standards bodies define two main formats for video that are supported today; CIF and SIF:

Phase Alternating Line (PAL)

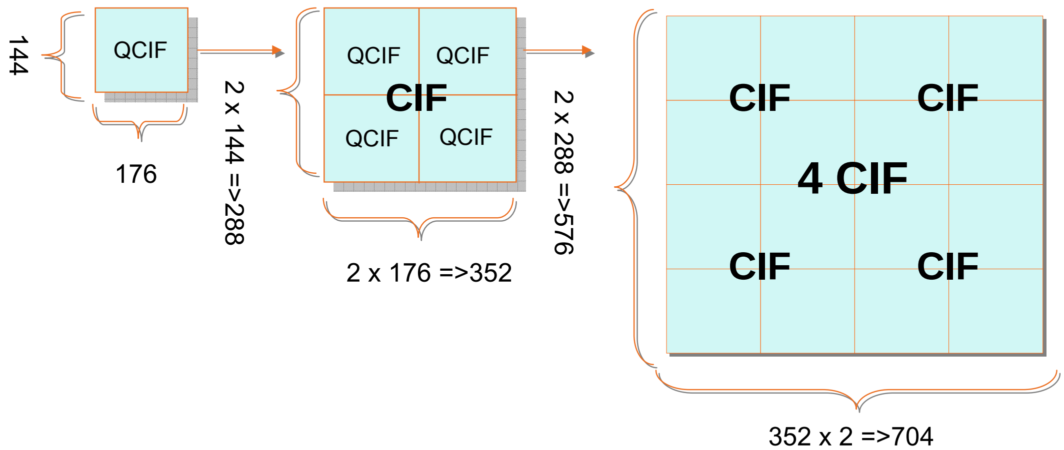
- CIF-With the advent of videoconferencing standards in the early 1990's a 'compromise' video standard between NTSC and PAL was created. Called CIF, for Common Intermediate Format, the idea was to have all videoconferencing systems translate PAL or NTSC to CIF, and by doing so, have each system take half of the step of translating from PAL to NTSC or NTSC to PAL: CIF is based on the resolution of PAL and the frame rate of NTSC. (352x288 at 30 fps)
- QCIF-Quarter CIF
- iCIF (2CIF)- interlaced CIF
- 4CIF- 4xCIF (704x576)

National Television System Committee (NTSC)

- SIF- Source Input Format (352x240)
- 4SIF- 4xSIF
- iSIF (2SIF)- interlaced SIF

QCIF – 4 CIF

- QCIF is a standard video resolution with dimensions 176x144 pixels (W x H)
- If we double the number of pixels on both sides we get an other standard video resolution known as CIF 352x288 pixels
- 4CIF = 704 x 576 pixels



TANDBERG
See: knowledge
slide 21
Video Conferencing Protocols

QCIF is a standard video resolution with dimensions 176x144 pixels (W x H).

If we double the number of pixels on both sides we get an other standard video resolution known as CIF 352x288 pixels.

4CIF 704x576

Interlaced Video

- NTSC & PAL are interlaced video
- In a video image, the dots (pixels) are not updated at the same time. The screen is drawn twice – odd fields first and than even fields.
- The frequency depends on the system
  - NTSC = 30 frames/sec = 60 fields/sec
  - PAL = 25 frames/sec = 50 fields/sec
- 1 frame = 2 Fields

TANDBERG See: knowledge   slide 22   Video Conferencing Protocols

NTSC & PAL are interlaced video

In a video image, the dots (pixels) are not updated at the same time. The screen is drawn twice – odd fields first and than even fields.

The frequency depends on the system

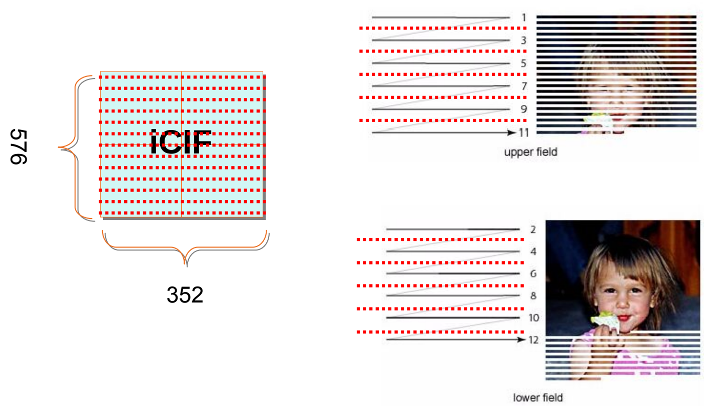
NTSC = 30 frames/sec = 60 fields/sec

PAL = 25 frames/sec = 50 fields/sec

1 frame = 2 Fields

Interlaced CIF - iCIF

An extra field is added to a CIF image increasing the resolution to iCIF 352x576



slide 23 Video Conferencing Protocols

An extra field is added to a CIF image increasing the resolution to iCIF 352x576

Video Conferencing Protocols: H.26x

H.261 (64Kbs – 2Mbs)

Modem

H.263 (64Kbs – 2Mbs)

H.264 (MPEG-4)

TANDBERG See: knowledge slide 24 Video Conferencing Protocols

H.261 is an ITU standard for videoconferencing. H.261 operates in the 64kbps to 2mbps range. All H.323 compliant videoconferencing system are required to support this codec. QCIF and CIF formats are found in H.261. H.261 offers full-pel motion compensation

H.263 is also an ITU standard for video coding. H.263 offers better compression than H.261, particularly in the low bitrate range. The H.263 standard contains a mechanism to define and use customized formats. H.263 supports the following standard source formats:

- SQCIF 128x96@29.97 fps
- 4CIF 704x576@29.97 fps
- 16CIF 1408x1152@29.97 fps

H.264 is the latest ITU standard for video compression. It is based on MPEG-4 and renders roughly equal video quality with H.263, but at half the bit rate (e.g. 256 Kbps instead of 512 Kbps for an H.263 stream.)

Video Interfaces

VGA Socket/female

VGA Plug/male

DVI Socket/female

DVI Plug/male

RCA Plugs and sockets

S-Video Socket/female

S-Video Plug/male

XLR Socket/female and Plug/male

TANDBERG See: knowledge slide 25 Video Conferencing Protocols

VGA (640x480), SVGA (800x600), XGA (1024x768), and UVGA (1280x1024) data display standards.

DVI- (**D**igital **V**isual **I**nterface) A standard interface to a digital display system. DVI sockets are found on flat panel monitors and TVs, DVD players, data projectors and cable TV set-top boxes. The DVI-Integrated (DVI-I) socket on the monitor accommodates both analog and digital signals. It can accept an analog VGA signal from the computer using a DVD-A plug or a digital DVI signal using a DVD-D plug.

RCA- An RCA connector is a plug and a jack designed for use with coaxial cable for frequencies ranging from the very lowest up to several megahertz. An RCA connector is sometimes known as a phono plug and jack.

S-Video- (Super-video) A video color format that combines the three YUV video signals into two channels. Brightness/luma (Y) is in one channel, and color/chroma (U and V) are in another.

XLR- (e**X**ternal **L**eft **R**ight or e**X**ternal **L**ive **R**eturn connector) An audio plug and socket used in professional and high-end audio equipment. It uses a balanced connection and typically locks into the socket. XLR connectors are twice the size of the standard RCA plug and socket.

TANDBERG

See: productivity

See: passion

See: performance

Video Conferencing Protocols

This concludes the Video Conferencing Protocols lesson. You are now advised to proceed to the next pre-requisite lesson that is available as part of your remote learning syllabus accessible through the TANDBERG University portal. Thank you.