Open and standardized – the world of IP Protocols

H.323: The Leading Standard in Voice over IP

The ITU-T H.323 standards enable convergence of information and communications applications as well as enhanced integration and interoperability among IP Communication products. This protocol family therefore forms the technological basis for IP Communication.

Introduction

H.323 is the leading standard in the Voice over IP (VoIP) area. The term VoIP stands for more than only voice transmission in IP networks. It covers an abundance of applications that are now being successively integrated due to the universality and ubiquity of the IP networks. Enhanced performance of IP and Ethernet networks, as well as the improved manageability of the bandwidth, allow traditional switched network applications such as Automatic Call Distribution, Real-time Messaging and Teleworking to be offered in IP networks. In addition to voice applications, H.323 provides mechanisms for video communication and data collaboration, in combination with the ITU-T T.120 series of standards.

The H.323 standard published in 1996 by the ITU-T represents the basis for data, voice and video communication via IP-based LANs and the Internet. The H.323 standard refers to many other standards, all known and referred to as members of the H.323 family of standards, such as H.245, H.225, H.450, etc.

H.323 regulates the technical requirements for visual telephony, which means the transmission of audio and video in packet based networks. Since IP is the prevailing protocol in packet-based networks—with about 90 percent market share—the H.323 standard is interpreted as a standard for multimedia communication in IP networks. Per definition, H.323 focuses on IP packet-based networks that do not provide any guaranteed service quality; e.g. packets can be lost and there is no prioritization of the real-time (voice and video) traffic over non-real-time, and therefore delay-insensitive, data traffic. Although this was a requirement in 1996, recent development in IP networking technology introduced Quality of Service (QoS) mechanisms that led to improved voice/video quality. Nevertheless, with majority of IP networks still not having QoS capabilities today, the mechanisms of H.323 help provide reliable communication. Since IP runs on any existing layer 2 technology, H.323 can be used over Ethernet, Fast Ethernet, Gigabit Ethernet, FDDI and Token-Ring. Recent implementation proved that H.323 can be also used beyond LANs, in multisite configurations over Wide Area Networks based on T1, Frame Relay, and ATM technology. Since February 1998 version 2 of the H.323 standard has been available under the name Packet-Based Multimedia Communications System; version 3 became available at the end of September 1999.

H.323 is often characterized as an "umbrella specification," because it refers to various other ITU standards. Not only the topology, but also its parts as well as the protocols and standards are specified in H.323. The H.323 components are terminal, gateway, gatekeeper, and Multipoint Control Units (MCUs). Terminals represent the end devices of every connection. Gateways establish the connection in other networks, i.e. gateways connect the H.323 network with the switched network of PBXs and Central Office switches. Gatekeepers take over the task of translating between telephone number, e.g. in accordance to the E.164 numbering standard, and IP addresses. They also manage the bandwidth, and provide mechanisms for terminal registration and authentication. MCUs are responsible for establishing multipoint conferences. The H.323 standard makes the distinction between callable and addressable end devices: all components are addressable; gatekeepers are, however, not callable.

The four components communicate by exchanging information flows among each other. These are split into five categories:

- Audio (digitized and coded) voice
- Video (digitized and coded full-motion image communication)
- Data (files such as text documents or images)
- Communication control (exchange of supported functions, controlling logical channels, etc.)
- Controlling connections (connection setup and connection release, etc.)

The information flows are routed according to the standard H.225.0 to the LAN-interface. This is based on the transport layer (layer 4) in accordance with ISO/OSI Reference Model. The interface to the bottom is the H.225.0 layer, the LAN interface is not a subject matter of H.323 and depends on the network used. The prerequisite here is not only the reliable, but also the unreliable transmission which is handled in IP networks via TCP respective to UDP. In the following the existence of a LAN interface for IP networks is assumed. In an upward direction the interfaces consist of audio and video codes, respective to the standards H.225.0 and H.245. The terminal and user interface features do not lie within the range of H.323 either.

H.323 Topology

The topology given by H.323 comprises four main components and specifies their mandatory and optional functionality. The main components are terminals, gateways, gatekeepers and multipoint control units (MCU). A H.323 zone consists of the quantity of these components within a LAN segment. Several segments can be linked via routers. Per H.323 zone there is a maximum of one gatekeeper, the other components can exist in a random number.

Terminals

Terminals represent the end points of each H.323 connection and can be realized in hardware or software. The audio transmission via G.711 and support by the control protocols H.245, H.225 and RAS are mandatory. The use of other audio codes and the option to transfer video and data are optional. If these additional services are offered, certain codes have to be used.

If several codecs are available for the same kind of data, the codec to be used is negotiated at the beginning of a connection via H.245. Each communication begins and ends with an H.323 terminal, whereby several audio and video connections are possible simultaneously. Coding and decoding can take place in asymmetric operation even with various codecs.

Within a LAN segment terminals can intercommunicate directly. For connections in other networks or in another LAN an additional component is required: the gateway.

Gateways

Gateways serve to establish connections with the telephone network or via a PBX system and are optional components of H.323 topology. The function of the gateways is to convert the various data formats in transport, process control and audio/video processing. Data communication of the gateways with the terminals is via H.245 and H.225. Some of the gateway functions are not exactly specified in H.323 and are left up to the manufacturer, for example, the maximum number of connected terminals, the maximum number of connections to other networks, the number of simultaneously independent conferences as well as the supported conversion and multipoint functions.

Gatekeepers

Gatekeepers take over various control and management functions within an H.323 zone and also belong to the optional components. If a gatekeeper exists, its services have to be used by the terminals. Per H.323 zone only one gatekeeper is permitted. The two main tasks of the gatekeeper are address conversion and bandwidth management. The address conversion function serves to control the connection; bandwidth management is designed to avoid overload situations. Both functions are realized via the RAS protocol defined in H.225.0. The network administrator is able to allocate a part of the total bandwidth to H.323 connections and release the rest for other applications. If the preset limit has been reached, the gatekeeper rejects further connection requests from terminals or an increase in bandwidth for already existing connections, and prevents network overloads. The criteria to determine whether bandwidth is available is not the subject matter of H.323.

As the gatekeeper also takes over access control of the terminals via RAS, it can also reject connections if individual terminals are not authorized.

Finally the gatekeeper can also play a role by receiving and routing the H.245 channels in connections between two users. If the conference is extended to three or more users the gatekeeper routes the H.245 control channels to a multipoint controller which then takes over the task of controlling the conference.

Multipoint Control Units

Multipoint Control Units (MCU) are used in the case of conferences with more than two users. They ensure that connections are properly setup and released, that audio and video are mixed, and that the data are distributed among the conference. Each of the H.323 terminals sends its data to the MCU. An MCU consists of a Multipoint Controller (MC) and any number of Multipoint Processors (MP). The Multipoint Controller takes care of the H.245 and negotiating the general functions for audio and video processing and controls the resources by determining which data flows are to be transmitted by the MP(s). Multipoint Processors (MP) receive media streams from conference participants, processes them and distributes media streams to the terminals in the conference. Video processing refers to all algorithms and formats, audio processing only to the algorithms, data processing only to the flows. In video processing by MP, switching and mixing is also required. Switching ensures that a certain data flow is sent if several data flows are available (for example with the matching video sequences, if the speaker in a conference changes identified by an audio signal, or if a change is requested via H.245). Mixing allows several data flows to be combined, whereby the image created is split into several segments and re-coded.

MPs also perform audio switching and mixing. Incoming signals are decoded in a standard procedure according to Pulse-Code-Modulation (PCM) or analogously, combined in a suitable way and then coded in the desired audio format. In this combination interference signals and ancillary noises can be diminished.

An individual combination of the incoming audio data can be supplied to each user whereby private communication is enabled within conferences. The audio data transmitted should not be contained in the audio data received. MC and MP can be co-located with other components, e.g. with H.323 terminals

Processing Audio, Video and Data Flows

The key function of H.323 components is to exchange information flows. A distinction is made between audio, video and data flows which are processed with certain codecs. All three information flows are transmitted via logic channels in accordance with H.225.0.

Audio Processing with G.711 and G.723.1

Audio transmission has to be supported by the H.323 terminals via G.711 codec. G.711 was originally designed for ISDN networks with fixed transmission rates, and has an output of 64 kbit/s. Although feasible in most LAN environments, G.711 cannot be used on low bandwidth links; therefore, ITU-T specified G.723 is a preferred codec due to its exceptional compression of voice to 5.3-6.3 kbit/s. Further optional audio codecs are G.722, G.728, G.729 and MPEG1, all of them offering benefits for certain environments and applications. The H.323 endpoints can support any of these codecs, and can advertise and negotiate the usage of these codecs in communications to other endpoints.

Video Processing with H.261 and H.263

Video transmission is an optional function of H.323 terminals. If it is supported it has to be handled via the ITU-T standards H.261 and optionally via H.263. The H.261 standard uses transmission rates of n x 64 kBit/s (n = 1, 2, ... 30) and can therefore for example use several ISDN channels. H.261 uses intra- and inter-frame coding similar to MPEG. Motion compensation is an optional function.

The more recent H.263 standard is compatible with H.261, but features by far better image quality as a result of 1/2 Pixel Motion Estimation, Predicted Frames and is also suitable for lower transmission rates. H.263 defines five image formats. The interaction with H.261 takes place via the QCIF format supported by both.

Data Conferences with T.120

For a transmission of data between endpoints, the H.323 standard refers to the ITU-T T.120 standard that can be used for various applications in the field of Collaborative Work, such as White-boarding, Application Sharing, and joint document management. T.120 is independent of the operating system and transport protocol and is supported by more than 100 companies. The characteristics of T.120 comprise:

- Multipoint data conferences,
- Transmission with error correction and acknowledgement of receipt, control of certain package sequences at the receiving station – also from different transmitting stations,
- Independent of the underlying transmission layer (LAN, modem ...) and of the network (POTS, ISDN, CSDN, LAN)
- Interoperability and platform independence,
- Support of heterogeneous topologies (star, cascading, series connection ...),
- Scalability (PC to multiprocessor), standard compatibility (e.g. to H.320) and future reliability (ATM, Frame-Relay, security aspects).

T.120 utilizes layer architecture similar to the ISO/OSI layer model: top layers (T.126, T.127) are based on the services of lower layers (T.121 to T.125) and contain protocols for special conference applications such as common notebook (White-board) or multipoint file transmission.

Real-Time Transport Protocol (RTP)

H.323 is directed at networks without special service quality. For the transmission of real-time data, such as audio or video, additional Mechanisms are introduced to guarantee successful communication. H.225.0 protocol therefore refers to the Real-time Transport Protocol (RTP) from the Internet Engineering Task Force (IETF). RTP is specified in RFC 1889 and 1890, and enables ascertain real-time compatibility. If used under the TCP/IP protocol family RTP is based on UDP, and marks the UDP/IP packets with a timestamp and a sequence number. The receiving station is therefore able to sort incoming packets and play them in the correct sequences. In case a packet gets lost during transmission, RTP can play the previous packet instead of re-transmitting. Since voice and video are time critical applications, re-transmitting packets would take too long and be of no use. RTP also identifies duplicated packets, and plays only one of the copies.

To distinguish between different RTP connections the contents of the package can be described via the field Payload Type.
An optimal supplement to RTP is the Real-time Transport Control Protocol (RTCP) which contains all control functions of RTP.

RTP was designed as open and versatile protocol and therefore functions not only with IP, but also with other protocols, such as IPX, CLNP or ATM (AAL5). RTP supports not only Unicast, but also Multicast, e.g. it can be employed in multicast enabled networks.

It is important to underline the fact that RTP neither guarantees certain transmission rates nor voice quality or an error-free transmission. The receiving station is enabled to identify faulty or incomplete transmissions and reacts to them with suitable methods. These are, for example:

- Omitting faulty data
- Balancing package errors by duplicating the previous packets
- Renewed request if too many missing packets are received via RTCP

Supplementary Services in H.323 networks (H.450.x series)

H.323 defines basic call setup and tear-down but does not provide in its original form any mechanisms for supplementary services (features), e.g. call transfer, call forward, message waiting indication, etc. Beyond the requirement to provide this type of supplementary services within the IP network, there are several requirements with regard to seamless interworking with existing PBXs:

- Standardized protocol (vendor-independent)
- Independence of the network configuration and topology (packetized versus switched networks),
- Simple interworking with existing QSIG based PBX networks
- Simple interworking with the public ISDN network (DSSI) in Europe, National-ISDN in USA).
- Possibility of introducing manufacturer-specific extensions or of realizing manufacturer-specific features

(standardized termination mechanisms) as well as simple development of features (via open APIs).

The protocol which approximates requirements in the switched networks is "QSIG". QSIG is available as a world-wide standard "PSSI" (Private Signaling System I) from ISO/IEC JTCI for networking ISDN private branch exchanges (PBX). ITU-T Study Group 16 therefore selected and further developed QSIG as the basis for H.323 service characteristics (H.450 series) to meet the requirements of an H.323 environment.

The first H.450 standards were passed in February 1998 by the ITU-T Study Group 16 and included::

- H.450.1 "Generic Functions for the Control of Supplementary Services in H.323". This standard lays down basic functions which are used by all further H.450.x standards.
- H.450.2 "Call Transfers": Enable a call between users A and B to be transferred to a user C (with or without inquiry connection).
- H.450.3 "Call Diversion". This standard enables call diversion to a user C if "busy"(CF-B), "no reply" (CF-NR), "unconditional"(CF-U) or with "user request" (Call Deflect).

Further supplementary services of the H.450 series were passed in May 1999:

- H.450.4 "Call Hold": Enables a user A to hold a call of user B temporarily (call hold) (e.g. before an inquiry connection to C is established)
- H.450.5 "Call Park and Pickup": This standard describes the procedures to park a call (e.g. at a server) anl later pick it up from the park position (e.g. server).
- H.450.6 "Call Waiting": Enables a busy user B to present an incoming call, whereby user A is informed of the call status of B.
- H.450.7: "Message Waiting Indication": A "Message Center" (e.g. a Messaging System) informs a user of incoming messages, such as voicemail, video, fax or e-mail).

The following H.450.x standard were recently approved by ITU-T:

- H.450.8: "Names Identification Service": Transmission of user names of user A and user B within the course of connecting the call.
- H.450.9: "Call Completion Services for H.323 networks" A user A who calls a busy or not answering user B, can apply for an "automatic return call" by means of this standard. That means that the connection between user A and user B is automatically established as soon as user B is available again (without user A having to repeat his attempt to contact user B).

Multi-point conferencing has been an integral part of H.323 from the beginning and is covered by the standards H.225.0 and H.245. Therefore, there is no specific H.450.x standard for multi-point conferencing.

H.450.x supplementary services can be used as "Stand-a-lone" features (e.g. user initiates call transfer via an appropriate user interface), but can also be used as module or further applications. e.g. Call Park or Call Transfer can be applied within the scope of Automatic Call Distribution (ACD).

Recommendations for interworking between H.450 supplementary services with the appropriate PBX features (QSIG) are currently being configured in ECMA TC32 TG17. Recommended H.450 implementation profiles have already been passed by ETSI.