

The H.323 protocol

The H.323 protocol is the common name for the International Telecommunications Union (ITU) Recommendation that defines packet-based multimedia communications systems. The most commonly deployed packet-based networks are, of course, those based on the TCP/IP suite of communications, which inter-connect to form the Internet. The H.323 recommendation is a widely adopted umbrella protocol that defines standard behaviour for setting up and proceeding with audio and video calls. It is known as an umbrella protocol because it depends on, and references, other protocols for call signalling, media transport and media encoding. Full details of the protocol can be found in the Recommendation itself, which is available from the ITU, and in the VTAS document “An Introduction to H.323 Videoconferencing”.

The H.323 protocol works very well where it is used within the same organisational IP network. However, problems are likely to occur when an organisation wishes to make H.323 calls to other organisations. This will usually mean traversing firewalls and NAT boundaries. The issues that have to be solved in order for H.323 to work across NAT boundaries and/or firewalls can be summarised as „network border traversal problems“. Before considering these problems, and possible solutions, it is worth examining how the H.323 protocol works in more detail.

Whether the call is mediated by a gatekeeper or not, the communication between the endpoints uses ITU Recommendation H.225.0 for call signalling. The call setup procedure can be paraphrased as shown in Table 1:

Endpoint	Protocol	Message	Port
A	H.225.0	Can we set up a call?	1720
B	H.225.0	OK, call proceeding	1720
B	H.225.0	Alerting user (ringing)	1720
A	H.225.0	What port shall we use for the next bit?	1720

B	H.225.0	Let's do H.245 on these ports	1720
A	H.245	I can do this and that (these are the speeds/encodings/decodings/etc I am capable of)...	Ports between
B	H.245	I can do this, and that...	As ab
A	H.245	Shall we use this encoding, that speed etc? On these ports?	As ab
B	H.245	Yes, OK	As ab
A	H.245	Right, let's go...	As ab
A + B	RTP	Media content – two ports (content and control) in each direction, per media	A group of contiguous ports, the last 1024

The acceptable play-out of real-time media is dependent on the media data being delivered in a timely manner. There is no point in resending media packets as they are continuously being decoded and passed to the application for display at the other end. By the time the re-sent snippet arrived it would be too late for inclusion in this rolling process. TCP/IP has built-in checks to ensure that data packets are delivered. In the face of a congested network, TCP will control the transmission of packets by 'backing-off' and actually slowing transmission rates. TCP will also check on packet delivery and ask for the re-transmission of any packets that have not arrived at their destination. This is obviously unsuitable for videoconferencing and telephony applications, and for this reason (amongst others) TCP and IP datagrams are not used for media transmission; instead they use the User Datagram Protocol (UDP), which does not have the control element of TCP. This means that media does not get delayed but is churned out by the transmitter regardless of what is happening on the network from moment to moment. UDP does lack some of the functionality required for a real-time exchange, and so there is another protocol layer between the UDP packet and the actual data payload of encoded media. This layer is provided by the Real Time Protocol (RTP) which provides features like time stamping. The Real Time Control Protocol (RTCP) allows participating endpoints in a call to report back on QoS (Quality of Service) parameters. This creates the potential for the endpoint to make adjustments to the encoding and transmission during the call in order to improve QoS.

Source URL: <https://community-stg.jisc.ac.uk/library/janet-services-documentation/h323-protocol>