15

H.323 Protocols

The H.323 standard provides a foundation for audio, video, and data communications across IP-based networks, including the Internet. H.323 is an umbrella recommendation from the International Telecommunications Union (ITU) that sets standards for multimedia communications over Local Area Networks (LANs) that do not provide a guaranteed Quality of Service (QoS). These networks dominate today's corporate desktops and include packet-switched TCP/IP and IPX over Ethernet, Fast Ethernet and Token Ring network technologies. Therefore, the H.323 standards are important building blocks for a broad new range of collaborative, LAN-based applications for multimedia communications. They include parts of H.225.0-RAS, Q.931-H.245, RTP/RTCP and audio/video/data codecs, such as the audio codecs (G.711, G.723.1, G.728, etc.), video codecs (H.261, H.263) that compress and decompress media streams and data codecs (T.120).

Media streams are transported on RTP/RTCP. RTP carries the actual media and RTCP carries status and control information. The signalling, with the exception of RAS, is transported reliably over TCP. The following protocols deal with signalling:

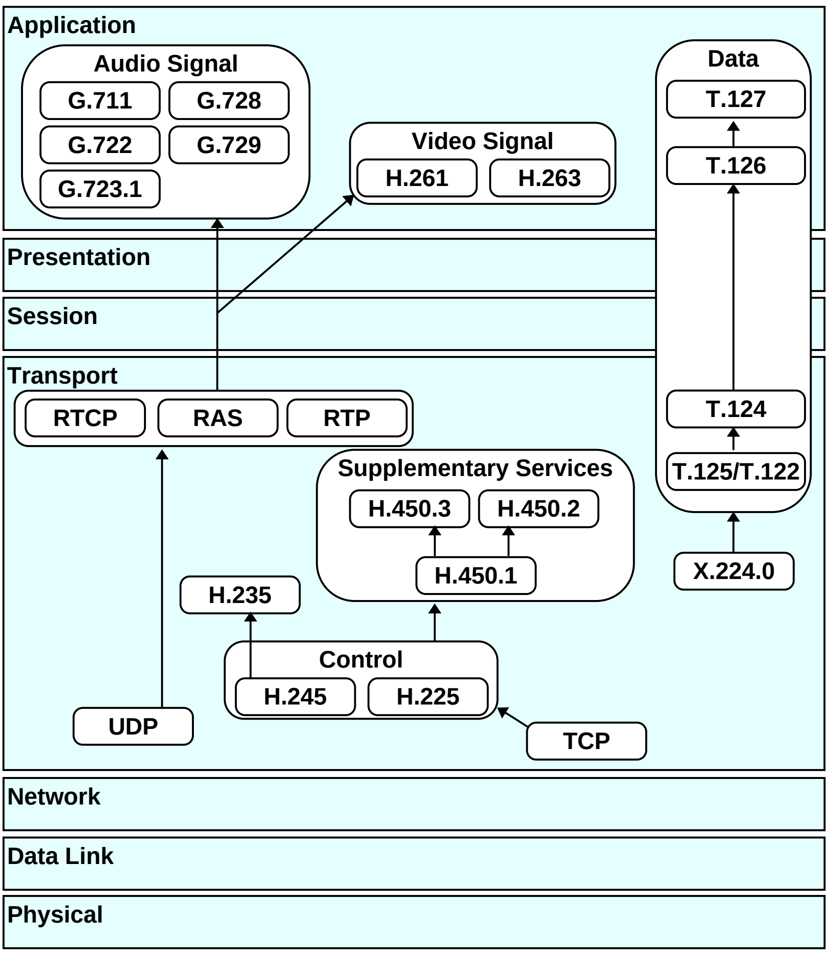
- RAS: manages registration, admission and status.

- Q.931: manages call setup and termination.
- H.245: negotiates channel usage and capabilities.

In addition, the following protocol provide optional features within the H.323 framework:

- H.235: security and authentication.
- H.450.x: supplementary services.

The following diagram illustrates the H.323 protocols in relation to the OSI model:



H.323 protocols in relation to the OSI model

RTP

RFC 1889 http://www.cis.ohio-state.edu/htbin/rfc/rfc1889.html

The Real-time Transport (RTP) Protocol provides end-to-end network transport functions suitable for applications transmitting real-time data such as audio, video or simulation data, over multicast or unicast network services. RTP does not address resource reservation and does not guarantee quality-of-service for real-time services. The data transport is augmented by a control protocol (RTCP) to allow monitoring of the data delivery in a manner scalable to large multicast networks, and to provide minimal control and identification functionality. RTP and RTCP are designed to be independent of the underlying transport and network layers. The protocol supports the use of RTP-level translators and mixers.

The format of the RTP Fixed Header Fields is shown in the following illustration:

0	1	2	3	4	5	6	7	Octet
V		P	X	CSRC count				1
M	Payload type							2
Sequence number								3
Timestamp								4
SSRC								5
CSRC								6

RTP structure

V

Version. Identifies the RTP version.

P

Padding. When set, the packet contains one or more additional padding octets at the end which are not part of the payload.

X

Extension bit. When set, the fixed header is followed by exactly one header extension, with a defined format.

CSRC count

Contains the number of CSRC identifiers that follow the fixed header.

M

Marker. The interpretation of the marker is defined by a profile. It is intended to allow significant events such as frame boundaries to be marked in the packet stream.

Payload type

Identifies the format of the RTP payload and determines its interpretation by the application. A profile specifies a default static mapping of payload type codes to payload formats. Additional payload type codes may be defined dynamically through non-RTP means.

Sequence number

Increments by one for each RTP data packet sent, and may be used by the receiver to detect packet loss and to restore packet sequence.

Timestamp

Reflects the sampling instant of the first octet in the RTP data packet. The sampling instant must be derived from a clock that increments monotonically and linearly in time to allow synchronization and jitter calculations. The resolution of the clock must be sufficient for the desired synchronization accuracy and for measuring packet arrival jitter (one tick per video frame is typically not sufficient).

SSRC

Identifies the synchronization source. This identifier is chosen randomly, with the intent that no two synchronization sources within the same RTP session will have the same SSRC identifier.

CSRC

Contributing source identifiers list. Identifies the contributing sources for the payload contained in this packet.

RTCP

RFC 1889 http://www.cis.ohio-state.edu/htbin/rfc/rfc1889.html

The RTP control protocol (RTCP) is based on the periodic transmission of control packets to all participants in the session, using the same distribution mechanism as the data packets. The underlying protocol must provide multiplexing of the data and control packets, for example using separate port numbers with UDP.

The format of the header is shown in the following illustration:

0	1	2	3	4	5	6	7	Octet
Version		P	Reception report count					1
Packet type								2
Length								3-4

RTCP structure

Version

Identifies the RTP version which is the same in RTCP packets as in RTP data packets. The version defined by this specification is two (2).

P

Padding. When set, this RTCP packet contains some additional padding octets at the end which are not part of the control information. The last octet of the padding is a count of how many padding octets should be ignored. Padding may be needed by some encryption algorithms with fixed block sizes. In a compound RTCP packet, padding should only be required on the last individual packet because the compound packet is encrypted as a whole.

Reception report count

The number of reception report blocks contained in this packet. A value of zero is valid.

Packet type

Contains the constant 200 to identify this as an RTCP SR packet.

Length

The length of this RTCP packet in 32-bit words minus one, including the header and any padding. (The offset of one makes zero a valid length and avoids a possible infinite loop in scanning a compound RTCP packet, while counting 32-bit words avoids a validity check for a multiple of 4.)

RAS

H.225: http://www.itu.int/itudoc/itu-t/rec/h/h225-0.html

Registration, Admission and Status (RAS) channel is used to carry messages used in the gatekeeper discovery and endpoint registration processes which associate an endpoint's alias address with its call signalling channel transport address. Since the RAS channel is an unreliable channel, H.225.0 recommends timeouts and retry counts for various messages. An endpoint or gatekeeper which cannot respond to a request within the specified timeout may use the Request in Progress (RIP) message to indicate that it is still processing the request. An endpoint or gatekeeper receiving the RIP, resets its timeout timer and retry counter.

RAS messages are in ASN.1 syntax. They consists of an exchange of messages.

H.225

H.225: http://www.itu.int/itudoc/itu-t/rec/h/h225-0.html

H.225.0 is a standard which covers narrow-band visual telephone services defined in H.200/AV.120-Series Recommendations. It specifically deals with those situations where the transmission path includes one or more packet based networks, each of which is configured and managed to provide a non-guaranteed QoS, which is not equivalent to that of N-ISDN, such that additional protection or recovery mechanisms beyond those mandated by Rec. H.320 are necessary in the terminals. H.225.0 describes how audio, video, data, and control information on a packet based network can be managed to provide conversational services in H.323 equipment.

The structure of H.225 follows the Q.931 standard as shown in the following illustration:

8	7	6	5	4	3	2	1	Octet
Protocol discriminator								1
0	0	0	0	Length of call ref				2
Call reference value								3 (-4)
0	Message type							
Information elements								

H.225 structure

Protocol discriminator

Distinguishes messages for user-network call control from other messages.

Length of call ref

The length of the call reference value.

Call reference value

Identifies the call or facility registration/cancellation request at the local user-network interface to which the particular message applies. May be up to 2 octets in length.

Message type

Identifies the function of the message sent. The following message type are used:

000	xxxxx	Call establishment message:
	00001	ALERTING
	00010	CALL PROCEEDING
	00111	CONNECT
	01111	CONNECT KNOWLEDGE
	00011	PROGRESS
	00101	SETUP
	01101	SETUP ACKNOWLEDGE
001	xxxxx	Call information phase message:
	00110	RESUME
	01110	RESUME ACKNOWLEDGE
	00010	RESUME REJECT
	00101	SUSPEND
	01101	SUSPEND ACKNOWLEDGE
	00001	SUSPEND REJECT
	00000	USER INFORMATION
010	xxxxx	Call clearing messages:
	00101	DISCONNECT
	01101	RELEASE
	11010	RELEASE COMPLETE
	00110	RESTART
	01110	RESTART ACKNOWLEDGE
011	xxxxx	Miscellaneous messages:
	00000	SEGMENT
	11001	CONGESTION CONTROL
	11011	INFORMATION
	01110	NOTIFY
	11101	STATUS
	10101	STATUS ENQUIRY

Information elements

Two categories of information elements are defined: single octet information elements and variable length information elements, as shown in the following illustrations.

8	7	6	5	4	3	2	1	Octet
1	IEI			Contents of IE				1

Single octet information element format (type 1)

8	7	6	5	4	3	2	1	Octet
1	IE Identifier							1

Single octet information element format (type 2)

8	7	6	5	4	3	2	1	Octet
1	IEI							1
Length of contents of IE								2
Contents of IE								3-n

Variable length information element format

H.245

H.245: http://www.itu.int/itudoc/itu-t/rec/h/h245.html

H.245 is line transmission of non-telephone signals. It includes receiving and transmitting capabilities as well as mode preference from the receiving end, logical channel signalling, and control and indication. Acknowledged signalling procedures are specified to ensure reliable audiovisual and data communication.

H.245 messages are in ASN.1 syntax. They consists of an exchange of messages.
MultimediaSystemControlMessage message types can be defined as request, response, command and indication messages. The following additional message sets are available:

- Master Slave Determination messages.
- Terminal Capability messages.
- Logical Channel signalling messages.
- Multiplex Table signalling messages.
- Request Multiplex Table signalling messages.
- Request Mode messages.
- Round Trip Delay messages.
- Maintenance Loop messages.
- Communication Mode messages.
- Conference Request and Response messages.
- Terminal ID.
- Commands and Indications.

H.261

H.261: http://www.cis.ohio-state.edu/htbin/rfc/rfc2032.html

The H.261 describes a video stream for transport using the real-time transport protocol, RTP, with any of the underlying protocols that carry RTP.

The format of the header is shown in the following illustration:

0	1	2	3	4	5	6	7	Octet
SBIT			EBIT			I	V	1
GOBN				MBAP				2
MBAP	QUANT					HMVD		3
HMVD			VMVD					4

H.261 header structure

SBIT
Start bit. Number of most significant bits that are to be ignored in the first data octet.

EBIT
End bit. Number of least significant bits that are to be ignored in the last data octet.

I
INTRA-frame encoded data flag. Set to 1 if this stream contains only INTRA-frame coded blocks. Set to 0 if this stream may or may not contain INTRA-frame coded blocks. The sense of this bit may not change during the course of the RTP session.

V
Motion Vector flag. Set to 0 if motion vectors are not used in this stream. This is set to 1 if motion vectors may or may not be used in this stream. The sense of this bit may not change during the course of the session.

GOBN
GOB number. Encodes the GOB number in effect at the start of the packet. This is set to 0 if the packet begins with a GOB header.

MBAP

Macroblock Address Predictor. Encodes the macroblock address predictor (i.e., the last MBA encoded in the previous packet). This predictor ranges from 0-32 (to predict the valid MBAs 1-33), but because the bit stream cannot be fragmented between a GOB header and MB 1, the predictor at the start of the packet can never be 0. Therefore, the range is 1-32, which is biased by -1 to fit in 5 bits. This is set to 0 if the packet begins with a GOB header.

QUANT

Quantizer. Shows the Quantizer value (MQUANT or GQUANT) in effect prior to the start of this packet. Set to 0 if the packet begins with a GOB header.

HMVD

Horizontal Motion Vector Data. Represents the reference horizontal Motion Vector Data (MVD). Set to 0, if V flag is 0 or if the packet begins with a GOB header, or when the MTYPE of the last MB encoded in the previous packet was not MC. HMVD is encoded as a 2's complement number and 10000 corresponding to the value -16 is forbidden (motion vector fields range from +/-15).

VMVD

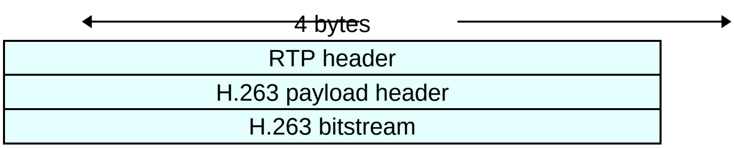
Vertical Motion Vector Data. Represents the reference vertical Motion Vector Data (MVD). Set to 0 if V flag is 0, or if the packet begins with a GOB header, or when the MTYPE of the last MB encoded in the previous packet was not MC. VMVD is encoded as a 2’s complement number and 10000 corresponding to the value -16 is forbidden (motion vector fields range from +/-15).

H.263

RFC 2190 (RTP): http://www.cis.ohio-state.edu/htbin/rfc/rfc2190.html
H.263: http://www.itu.int/itudoc/itu-t/rec/h/h263.html

This protocol specifies the payload format for encapsulating an H.263 bitstream in the Real-time Transport Protocol (RTP). Three modes are defined for the H.263 payload header. An RTP packet can use one of the three modes for H.263 video streams depending on the desired network packet size and H.263 encoding options employed. The shortest H.263 payload header (mode A) supports fragmentation at Group of Block (GOB) boundaries. The long H.263 payload headers (modes B and C) support fragmentation at Macroblock (MB) boundaries.

For each RTP packet, the RTP fixed header is followed by the H.263 payload header, which is followed by the standard H.263 compressed bitstream. The size of the H.263 payload header is variable depending on the modes. The layout of an RTP H.263 video packet is as shown in the following illustration:



RTP H.263 video packet

In mode A, an H.263 payload header of four bytes is present before an actual compressed H.263 video bitstream in a packet. It allows fragmentation at GOB boundaries. In mode B, an eight byte H.263 payload header is used and each packet starts at MB boundaries, without the PB-frames option. Finally, a twelve byte H.263 payload header is defined in mode C to support fragmentation at MB boundaries for frames that are coded with the PB-frames option.

The mode of each H.263 payload header is indicated by the F and P fields in the header. Packets of different modes can be intermixed. The format of the header for mode A is shown in the following illustration:

1	2	3	4	5	6	7	8	Octet
F	P	SBIT			EBIT			1
SRC			I	U	S	A	R	2
R (cont.)		DBQ		TRB				3
TR								4

H.263 mode A payload header structure

F
Flag bit, indicates the mode of the payload header.

P
Optional PB-frames mode as defined by H.263.

0 Normal I or P frame.
1 PB-frames.

When F=1, P also indicates modes:

0 Mode B.
1 Mode C.

SBIT
Start bit, specifies the number of most significant bits that should be ignored in the first data byte.

EBIT
End bit, specifies the number of least significant bits that should be ignored in the last data byte.

SRC
Source format (bit 6, 7 and 8 in TYPE defined by H.263), specifies the resolution of the current picture.

I
Picture coding type (bit 9 in PTYPE defined by H.263):

0 Intra-coded.
1 Inter-coded.

U
Set to 1 if the Unrestricted Motion Vector option (bit 10 in PTYPE defined by H.263) was set to 1 in the current picture header, otherwise 0.

S
Set to 1 if the Syntax-based Arithmetic Coding option (bit 11 in PTYPE defined by H.263) was set to 1 for current picture header, otherwise 0.

A
Set to 1 if the Advanced Prediction option (bit 12 in PTYPE defined by H.263) was set to 1 for current picture header, otherwise 0.

R
Reserved, must be set to zero.

DBQ
Differential quantization parameter used to calculate the quantizer for the B frame based on the quantizer for the P frame, when the PB-frames option is used. The value should be the same as DBQUANT defined by H.263. Set to zero if the PB-frames option is not used.

TRB
Temporal reference for the B frame as defined by H.263. Set to zero if the PB-frames option is not used.

TR
Temporal reference for the P frame as defined by H.263. Set to zero if the PB-frames option is not used.

The format of the header for mode B is shown in the following illustration:

<table>
<tr><th>1</th><th>2</th><th>3</th><th>4</th><th>5</th><th>6</th><th>7</th><th>8</th><th>Octet</th></tr>
<tr><td>F</td><td>P</td><td colspan="3">SBIT</td><td colspan="3">EBIT</td><td>1</td></tr>
<tr><td colspan="3">SRC</td><td colspan="5">QUANT</td><td>2</td></tr>
<tr><td colspan="5">GOBN</td><td colspan="3">MBA</td><td>3</td></tr>
<tr><td colspan="6">MBA (cont.)</td><td colspan="2">R</td><td>4</td></tr>
<tr><td>I</td><td>U</td><td>S</td><td>A</td><td colspan="4">HMV1</td><td>5</td></tr>
<tr><td colspan="3">HMV1 (cont.)</td><td colspan="5">VMV1</td><td>6</td></tr>
<tr><td colspan="2">VMV1 (cont.)</td><td colspan="6">HMV2</td><td>7</td></tr>
<tr><td>HMV2</td><td colspan="7">VMV2</td><td>8</td></tr>
</table>

H.263 mode B payload header structure

F, P, SBIT, EBIT, SRC, I, U, S and A are defined the same as in mode A.

QUANT

Quantization value for the first MB coded at the start of the packet. Set to 0 if the packet begins with a GOB header.

GOBN

GOB number in effect at the start of the packet. GOB number is specified differently for different resolutions.

MBA

The address within the GOB of the first MB in the packet, counting from zero in scan order. For example, the third MB in any GOB is given MBA=2.

R

Reserved, set to zero.

HMV1, VMV1

Horizontal and vertical motion vector predictors for the first MB in this packet. When four motion vectors are used for the current MB with advanced prediction option, they are the motion vector predictors for block number 1 in the MB. Each 7-bit field encodes a motion vector predictor in half pixel resolution as a 2's complement number.

HMV2, VMV2

Horizontal and vertical motion vector predictors for block number 3 in the first MB in this packet when four motion vectors are used with the advanced prediction option. This is needed because block number 3 in the MB needs different motion vector predictors from other blocks in the MB. These two fields are not used when the MB only has one motion vector. Each 7-bit field encodes a motion vector predictor in half pixel resolution as a 2's complement number.

The format of the header for mode C is shown in the following illustration:

<table>
<tr><th>1</th><th>2</th><th>3</th><th>4</th><th>5</th><th>6</th><th>7</th><th>8</th><th>Octet</th></tr>
<tr><td>F</td><td>P</td><td colspan="3">SBIT</td><td colspan="3">EBIT</td><td>1</td></tr>
<tr><td colspan="3">SRC</td><td colspan="5">QUANT</td><td>2</td></tr>
<tr><td colspan="5">GOBN</td><td colspan="3">MBA</td><td>3</td></tr>
<tr><td colspan="6">MBA (cont.)</td><td colspan="2">R</td><td>4</td></tr>
<tr><td>I</td><td>U</td><td>S</td><td>A</td><td colspan="4">HMV1</td><td>5</td></tr>
<tr><td colspan="3">HMV1 (cont.)</td><td colspan="5">VMV1</td><td>6</td></tr>
<tr><td colspan="2">VMV1 (cont.)</td><td colspan="6">HMV2</td><td>7</td></tr>
<tr><td>HMV2</td><td colspan="7">VMV2</td><td>8</td></tr>
<tr><td colspan="8">RR</td><td>9</td></tr>
<tr><td colspan="3">RR (cont.)</td><td colspan="2">DBQ</td><td colspan="3">TRB</td><td>10</td></tr>
<tr><td colspan="8">TR</td><td>11</td></tr>
</table>

H.263 mode B payload header structure

F, P, SBIT, EBIT, SRC, I, U, S, A, DBQ, TRB and TR are defined the same as in mode A. QUANT, GOBN, MBA, HMV1, VMV1, HMV2, VMV2 are defined the same as in mode B.

RR
Reserved, set to zero.

H.235

H.235: http://www.itu.int/itudoc/itu-t/rec/h/h235.html

H.235 provides enhancements within the framework of the H.3xx-Series Recommendations to incorporate security services such as Authentication and Privacy (data encryption). H.235 should work with other H series protocols that utilize H.245 as their control protocol.

All H.235 messages are encrypted as in ASN.1.