



INTERNATIONAL TELECOMMUNICATION UNION

# ITU-T

TELECOMMUNICATION  
STANDARDIZATION SECTOR  
OF ITU

# H.323

(02/98)

SERIES H: AUDIOVISUAL AND MULTIMEDIA SYSTEMS

Infrastructure of audiovisual services – Systems and  
terminal equipment for audiovisual services

---

## **Packet-based multimedia communications systems**

ITU-T Recommendation H.323

(Previously CCITT Recommendation)

ITU-T H-SERIES RECOMMENDATIONS  
**AUDIOVISUAL AND MULTIMEDIA SYSTEMS**

|                                                                                                            |                    |
|------------------------------------------------------------------------------------------------------------|--------------------|
| Characteristics of transmission channels used for other than telephone purposes                            | H.10–H.19          |
| Use of telephone-type circuits for voice-frequency telegraphy                                              | H.20–H.29          |
| Telephone circuits or cables used for various types of telegraph transmission or simultaneous transmission | H.30–H.39          |
| Telephone-type circuits used for facsimile telegraphy                                                      | H.40–H.49          |
| Characteristics of data signals                                                                            | H.50–H.99          |
| CHARACTERISTICS OF VISUAL TELEPHONE SYSTEMS                                                                | H.100–H.199        |
| INFRASTRUCTURE OF AUDIOVISUAL SERVICES                                                                     |                    |
| General                                                                                                    | H.200–H.219        |
| Transmission multiplexing and synchronization                                                              | H.220–H.229        |
| Systems aspects                                                                                            | H.230–H.239        |
| Communication procedures                                                                                   | H.240–H.259        |
| Coding of moving video                                                                                     | H.260–H.279        |
| Related systems aspects                                                                                    | H.280–H.299        |
| <b>Systems and terminal equipment for audiovisual services</b>                                             | <b>H.300–H.399</b> |

*For further details, please refer to ITU-T List of Recommendations.*

## ITU-T RECOMMENDATION H.323

### PACKET-BASED MULTIMEDIA COMMUNICATIONS SYSTEMS

#### Summary

This Recommendation describes terminals and other entities that provide multimedia communications services over Packet Based Networks (PBN) which may not provide a guaranteed Quality of Service. H.323 entities may provide real-time audio, video and/or data communications. Support for audio is mandatory, while data and video are optional, but if supported, the ability to use a specified common mode of operation is required, so that all terminals supporting that media type can interwork.

The packet based network over which H.323 entities communicate may be a point-to-point connection, a single network segment, or an internetwork having multiple segments with complex topologies.

H.323 entities may be used in point-to-point, multipoint, or broadcast (as described in Recommendation H.332) configurations. They may interwork with H.310 terminals on B-ISDN, H.320 terminals on N-ISDN, H.321 terminals on B-ISDN, H.322 terminals on Guaranteed Quality of Service LANs, H.324 terminals on GSTN and wireless networks, V.70 terminals on GSTN, and voice terminals on GSTN or ISDN through the use of Gateways.

H.323 entities may be integrated into personal computers or implemented in stand-alone devices such as videotelephones.

Products claiming compliance with Version 1 of H.323 shall comply with all of the mandatory requirements of H.323 (1996) which references Recommendations H.225.0 (1996) and H.245 (1996). Version 1 products can be identified by H.225.0 messages containing a **protocolIdentifier** = {itu-t (0) recommendation (0) h (8) 2250 version (0) 1} and H.245 messages containing a **protocolIdentifier** = {itu-t (0) recommendation (0) h (8) 245 version (0) 2}. Products claiming compliance with Version 2 of H.323 shall comply with all of the mandatory requirements of this Recommendation, H.323 (1998), which references Recommendations H.225.0 (1998) and H.245 (1998). Version 2 products can be identified by H.225.0 messages containing a **protocolIdentifier** = {itu-t (0) recommendation (0) h (8) 2250 version (0) 2} and H.245 messages containing a **protocolIdentifier** = {itu-t (0) recommendation (0) h (8) 245 version (0) 3}.

Note that the title of H.323 (1996) was "Visual telephone systems and equipment for local area networks which provide a non-guaranteed quality of service". The title has been changed in this version to be consistent with its expanded scope.

#### Source

ITU-T Recommendation H.323 was revised by ITU-T Study Group 16 (1997-2000) and was approved under the WTSC Resolution No. 1 procedure on the 6th of February 1998.

## FOREWORD

ITU (International Telecommunication Union) is the United Nations Specialized Agency in the field of telecommunications. The ITU Telecommunication Standardization Sector (ITU-T) is a permanent organ of the ITU. The ITU-T is responsible for studying technical, operating and tariff questions and issuing Recommendations on them with a view to standardizing telecommunications on a worldwide basis.

The World Telecommunication Standardization Conference (WTSC), which meets every four years, establishes the topics for study by the ITU-T Study Groups which, in their turn, produce Recommendations on these topics.

The approval of Recommendations by the Members of the ITU-T is covered by the procedure laid down in WTSC Resolution No. 1.

In some areas of information technology which fall within ITU-T's purview, the necessary standards are prepared on a collaborative basis with ISO and IEC.

## NOTE

In this Recommendation, the expression "Administration" is used for conciseness to indicate both a telecommunication administration and a recognized operating agency.

## INTELLECTUAL PROPERTY RIGHTS

The ITU draws attention to the possibility that the practice or implementation of this Recommendation may involve the use of a claimed Intellectual Property Right. The ITU takes no position concerning the evidence, validity or applicability of claimed Intellectual Property Rights, whether asserted by ITU members or others outside of the Recommendation development process.

As of the date of approval of this Recommendation, the ITU had received notice of intellectual property, protected by patents, which may be required to implement this Recommendation. However, implementors are cautioned that this may not represent the latest information and are therefore strongly urged to consult the TSB patent database.

© ITU 1998

All rights reserved. No part of this publication may be reproduced or utilized in any form or by any means, electronic or mechanical, including photocopying and microfilm, without permission in writing from the ITU.

# CONTENTS

|                                                                        | <b>Page</b> |
|------------------------------------------------------------------------|-------------|
| 1 Scope .....                                                          | 1           |
| 2 Normative references.....                                            | 2           |
| 3 Definitions .....                                                    | 4           |
| 4 Symbols and abbreviations.....                                       | 9           |
| 5 Conventions.....                                                     | 12          |
| 6 System description.....                                              | 12          |
| 6.1 Information streams.....                                           | 12          |
| 6.2 Terminal characteristics .....                                     | 13          |
| 6.2.1 Terminal elements outside the scope of this Recommendation ..... | 14          |
| 6.2.2 Terminal elements within the scope of this Recommendation .....  | 14          |
| 6.2.3 Packet based network interface .....                             | 14          |
| 6.2.4 Video codec.....                                                 | 15          |
| 6.2.5 Audio codec.....                                                 | 15          |
| 6.2.6 Receive path delay .....                                         | 17          |
| 6.2.7 Data channel.....                                                | 17          |
| 6.2.8 H.245 control function .....                                     | 19          |
| 6.2.9 RAS signalling function.....                                     | 23          |
| 6.2.10 Call signalling function .....                                  | 24          |
| 6.2.11 H.225.0 layer .....                                             | 24          |
| 6.3 Gateway characteristics.....                                       | 24          |
| 6.4 Gatekeeper characteristics.....                                    | 27          |
| 6.5 Multipoint controller characteristics .....                        | 28          |
| 6.6 Multipoint processor characteristics.....                          | 29          |
| 6.7 Multipoint control unit characteristics .....                      | 30          |
| 6.8 Multipoint capability .....                                        | 30          |
| 6.8.1 Centralized multipoint capability .....                          | 30          |
| 6.8.2 Decentralized multipoint capability .....                        | 31          |
| 6.8.3 Hybrid multipoint – Centralized audio .....                      | 31          |
| 6.8.4 Hybrid multipoint – Centralized video .....                      | 31          |
| 6.8.5 Establishment of common mode .....                               | 32          |
| 6.8.6 Multipoint rate matching.....                                    | 32          |
| 6.8.7 Multipoint lip synchronization.....                              | 32          |
| 6.8.8 Multipoint encryption.....                                       | 33          |
| 6.8.9 Cascading multipoint control units .....                         | 33          |

|        | <b>Page</b>                                                     |
|--------|-----------------------------------------------------------------|
| 7      | Call signalling..... 33                                         |
| 7.1    | Addresses..... 33                                               |
| 7.1.1  | Network address..... 33                                         |
| 7.1.2  | TSAP identifier ..... 33                                        |
| 7.1.3  | Alias address ..... 33                                          |
| 7.2    | Registration, Admission and Status (RAS) channel..... 34        |
| 7.2.1  | Gatekeeper discovery ..... 34                                   |
| 7.2.2  | Endpoint registration..... 35                                   |
| 7.2.3  | Endpoint location ..... 37                                      |
| 7.2.4  | Admissions, bandwidth change, status and disengage..... 37      |
| 7.2.5  | Access tokens ..... 38                                          |
| 7.3    | Call signalling channel ..... 38                                |
| 7.3.1  | Call signalling channel routing ..... 38                        |
| 7.3.2  | Control channel routing..... 39                                 |
| 7.4    | Call reference value..... 40                                    |
| 7.5    | Call ID ..... 41                                                |
| 7.6    | Conference ID and Conference Goal ..... 41                      |
| 8      | Call signalling procedures ..... 41                             |
| 8.1    | Phase A – Call setup..... 41                                    |
| 8.1.1  | Basic call setup – Neither endpoint registered..... 42          |
| 8.1.2  | Both endpoints registered to the same Gatekeeper ..... 42       |
| 8.1.3  | Only calling endpoint has Gatekeeper ..... 44                   |
| 8.1.4  | Only called endpoint has Gatekeeper..... 45                     |
| 8.1.5  | Both endpoints registered to different Gatekeepers ..... 47     |
| 8.1.6  | Optional Called Endpoint Signalling ..... 51                    |
| 8.1.7  | Fast Connect Procedure..... 53                                  |
| 8.1.8  | Call setup via gateways..... 55                                 |
| 8.1.9  | Call setup with an MCU..... 56                                  |
| 8.1.10 | Call forwarding ..... 57                                        |
| 8.1.11 | Broadcast call setup..... 57                                    |
| 8.1.12 | Overlapped Sending ..... 57                                     |
| 8.1.13 | Call setup to conference alias..... 57                          |
| 8.2    | Phase B – Initial communication and capability exchange..... 59 |
| 8.2.1  | Encapsulation of H.245 messages within Q.931 messages..... 59   |
| 8.2.2  | Tunneling through intermediate signalling entities..... 60      |
| 8.2.3  | Switching to a separate H.245 connection ..... 60               |

|                                                                                      | <b>Page</b> |
|--------------------------------------------------------------------------------------|-------------|
| 8.3 Phase C – Establishment of audiovisual communication .....                       | 61          |
| 8.3.1 Mode changes.....                                                              | 61          |
| 8.3.2 Exchange of video by mutual agreement .....                                    | 61          |
| 8.3.3 Media stream address distribution .....                                        | 62          |
| 8.3.4 Correlation of media streams in multipoint conferences .....                   | 62          |
| 8.3.5 Communication Mode Command Procedures .....                                    | 62          |
| 8.4 Phase D – Call services .....                                                    | 63          |
| 8.4.1 Bandwidth changes .....                                                        | 63          |
| 8.4.2 Status .....                                                                   | 65          |
| 8.4.3 Ad hoc conference expansion .....                                              | 66          |
| 8.4.4 Supplementary services.....                                                    | 74          |
| 8.4.5 Multipoint cascading.....                                                      | 74          |
| 8.4.6 Third party initiated pause and re-routing.....                                | 75          |
| 8.5 Phase E – Call termination .....                                                 | 75          |
| 8.5.1 Call clearing without a Gatekeeper .....                                       | 76          |
| 8.5.2 Call clearing with a Gatekeeper .....                                          | 76          |
| 8.5.3 Call clearing by Gatekeeper .....                                              | 76          |
| 8.6 Protocol failure handling .....                                                  | 77          |
| 9 Interoperation with other terminal types .....                                     | 78          |
| 9.1 Speech-only terminals .....                                                      | 78          |
| 9.2 Visual telephone terminals over the ISDN (H.320) .....                           | 78          |
| 9.3 Visual telephone terminals over GSTN (H.324) .....                               | 78          |
| 9.4 Visual telephone terminals over mobile radio (H.324/M) .....                     | 79          |
| 9.5 Visual telephone terminals over ATM (H.321 and H.310 RAST).....                  | 79          |
| 9.6 Visual telephone terminals over guaranteed quality of service LANs (H.322) ..... | 79          |
| 9.7 Simultaneous voice and data terminals over GSTN (V.70) .....                     | 80          |
| 9.8 T.120 terminals on the packet based network .....                                | 80          |
| 10 Optional enhancements .....                                                       | 80          |
| 10.1 Encryption .....                                                                | 80          |
| 10.2 Multipoint operation.....                                                       | 80          |
| 10.2.1 H.243 Control and Indication.....                                             | 80          |
| 11 Maintenance .....                                                                 | 80          |
| 11.1 Loopbacks for maintenance purposes .....                                        | 80          |
| 11.2 Monitoring methods .....                                                        | 81          |

|                                                                                | <b>Page</b> |
|--------------------------------------------------------------------------------|-------------|
| Annex A – H.245 messages used by H.323 endpoints .....                         | 82          |
| Annex B – Procedures for layered video codecs .....                            | 87          |
| B.1 Scope .....                                                                | 87          |
| B.2 Introduction .....                                                         | 87          |
| B.3 Scalability methods .....                                                  | 88          |
| B.4 Call establishment .....                                                   | 88          |
| B.5 Use of RTP sessions and codec layers .....                                 | 88          |
| B.5.1 Associate base to audio for lip synchronization.....                     | 88          |
| B.5.2 Enhancement layer dependency .....                                       | 89          |
| B.6 Possible layering models .....                                             | 90          |
| B.6.1 Multiple logical channels and RTP sessions for a layered stream.....     | 90          |
| B.6.2 Impact of one layer per logical channel and per RTP session .....        | 90          |
| B.7 Impact on multipoint conferences .....                                     | 91          |
| B.7.1 MC Impartial model .....                                                 | 91          |
| B.7.2 MC Decision model .....                                                  | 91          |
| B.7.3 Multipoint conference containing endpoints on different bandwidths ..... | 91          |
| B.8 Use of network QOS for layered video streams .....                         | 93          |
| Annex C – H.323 on ATM .....                                                   | 94          |
| C.1 Introduction .....                                                         | 94          |
| C.2 Scope .....                                                                | 94          |
| C.2.1 Point-to-point conferencing .....                                        | 94          |
| C.2.2 MCU-based multipoint.....                                                | 94          |
| C.2.3 H.323 interoperability with endpoints using IP .....                     | 94          |
| C.3 Architecture .....                                                         | 94          |
| C.3.1 Overview of system.....                                                  | 95          |
| C.3.2 Interoperation with other ITU-T H-Series endpoints.....                  | 95          |
| C.3.3 H.225.0 on IP over ATM .....                                             | 95          |
| C.3.4 H.245 on TCP/IP over ATM .....                                           | 95          |
| C.3.5 Addressing for A/V streams .....                                         | 96          |
| C.3.6 Transport Capabilities added to Terminal Capability Set.....             | 96          |
| C.3.7 Elements of ATM signalling .....                                         | 96          |
| C.3.8 A/V streams on RTP on AAL5 .....                                         | 96          |
| C.3.9 QOS considerations (Optional) .....                                      | 97          |
| C.4 Protocol section .....                                                     | 99          |
| C.4.1 ATM signalling information elements .....                                | 99          |
| C.4.2 H.245 Usage .....                                                        | 101         |

|                                                                      | <b>Page</b> |
|----------------------------------------------------------------------|-------------|
| C.4.3 RTP usage .....                                                | 102         |
| C.4.4 Interoperation with H.323 on IP.....                           | 102         |
| Appendix I – Sample MC to Terminal Communication Mode Command.....   | 103         |
| I.1 Sample conference Scenario A.....                                | 103         |
| I.2 CommunicationModeTable sent to all Endpoints .....               | 103         |
| I.3 Sample conference Scenario B.....                                | 104         |
| I.4 CommunicationModeTable sent to all Endpoints .....               | 104         |
| Appendix II – Transport level resource reservation procedures.....   | 105         |
| II.1 Introduction .....                                              | 105         |
| II.2 QOS Support for H.323 .....                                     | 106         |
| II.3 RSVP background .....                                           | 107         |
| II.4 The H.245 capability exchange phase .....                       | 108         |
| II.5 Open logical channel and setting up reservations .....          | 109         |
| II.6 Close logical channel and tearing down reservations .....       | 111         |
| II.7 Resource reservation for multicast H.323 logical channels ..... | 111         |
| Appendix III – Gatekeeper based user location.....                   | 112         |
| III.1 Introduction .....                                             | 112         |
| III.2 Signalling.....                                                | 112         |

|     |                         |
|-----|-------------------------|
| UCF | Unregister Confirmation |
| UDP | User Datagram Protocol  |
| URJ | Unregister Reject       |
| URQ | Unregister Request      |
| VBR | Variable Bit Rate       |
| VC  | Virtual Channel         |
| VC  | Virtual Circuit         |

## 5 Conventions

In this Recommendation, the following conventions are used:

"Shall" indicates a mandatory requirement.

"Should" indicates a suggested but optional course of action.

"May" indicates an optional course of action rather than a recommendation that something take place.

References to clauses, subclauses, Annexes and Appendices refer to those items within this Recommendation unless another specification is explicitly listed. For example, 1.4 refers to 1.4 of this Recommendation; 6.4/H.245 refers to 6.4 in Recommendation H.245.

Throughout this Recommendation, the term "network" is used to indicate any packet based network regardless of the underlying physical connection or the geographic scope of the network. This includes Local Area Networks, internetworks, and other packet based networks. The term "Switched Circuit Network" or "SCN" is used explicitly when referring to switched circuit networks such as GSTN and ISDN.

Where items exist on both the packet based network and the SCN, references to the SCN item will be explicit. For example, an MCU is an H.323 MCU on the packet based network, an SCN-MCU is an MCU on the SCN.

This Recommendation describes the use of three different message types: H.245, RAS and Q.931. To distinguish between the different message types, the following convention is followed. H.245 message and parameter names consist of multiple concatenated words highlighted in bold typeface (**maximumDelayJitter**). RAS message names are represented by three letter abbreviations (ARQ). Q.931 message names consist of one or two words with the first letters capitalized (Call Proceeding).

## 6 System description

This Recommendation describes the elements of the H.323 components. These are Terminals, Gateways, Gatekeepers, MCs and MCUs. These components communicate through the transmission of Information Streams. The characteristics of these components are described in this clause.

### 6.1 Information streams

Visual telephone components communicate through the transmission of Information Streams. These Information Streams are classified into video, audio, data, communications control and call control as follows.

Audio signals contain digitized and coded speech. In order to reduce the average bit rate of audio signals, voice activation may be provided. The audio signal is accompanied by an audio control signal.

Video signals contain digitized and coded motion video. Video is transmitted at a rate no greater than that selected as a result of the capability exchange. The video signal is accompanied by a video control signal.

Data signals include still pictures, facsimile, documents, computer files and other data streams.

Communications control signals pass control data between remote-like functional elements and are used for capability exchange, opening and closing logical channels, mode control and other functions that are part of communications control.

Call control signals are used for call establishment, disconnect and other call control functions.

The information streams described above are formatted and sent to the network interface as described in Recommendation H.225.0.

## 6.2 Terminal characteristics

An example of an H.323 terminal is shown in Figure 4. The diagram shows the user equipment interfaces, video codec, audio codec, telematic equipment, H.225.0 layer, system control functions and the interface to the packet based network. All H.323 terminals shall have a System Control Unit, H.225.0 layer, Network Interface and an Audio Codec Unit. The Video Codec Unit and User Data Applications are optional.

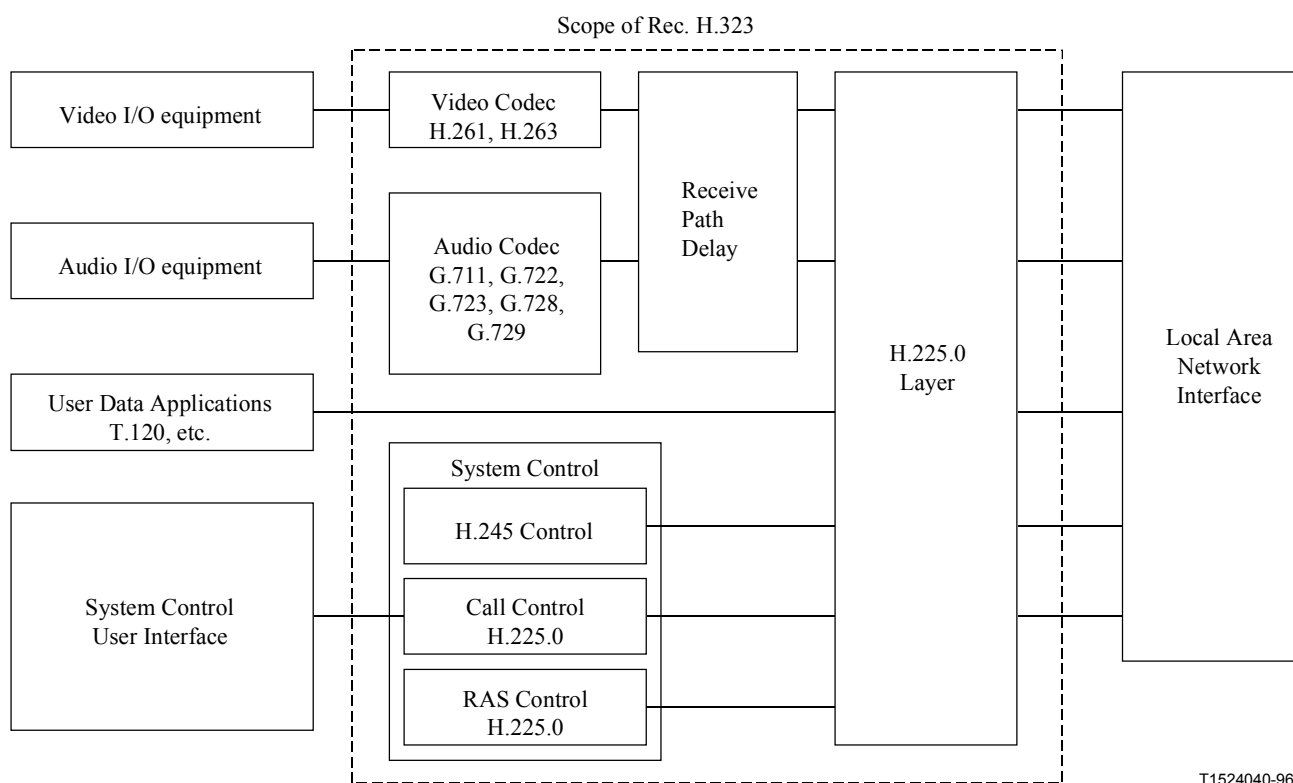


**Figure 4/H.323 – H.323 terminal equipment**

### **6.2.1 Terminal elements outside the scope of this Recommendation**

The following elements are not within the scope of this Recommendation and are therefore not defined within this Recommendation:

- Attached audio devices providing voice activation sensing, microphone and loudspeaker, telephone instrument or equivalent, multiple microphones mixers, and acoustic echo cancellation.
- Attached video equipment providing cameras and monitors, and their control and selection, video processing to improve compression or provide split screen functions.
- Data applications and associated user interfaces which use T.120 or other data services over the data channel.
- Attached Network Interface, which provides the interface to the packet based network, supporting appropriate signalling and voltage levels, in accordance with national and international standards.
- Human user system control, user interface and operation.

### **6.2.2 Terminal elements within the scope of this Recommendation**

The following elements are within the scope of this Recommendation, and are therefore subject to standardization and are defined within this Recommendation:

- The Video Codec (H.261, etc.) encodes the video from the video source (i.e. camera) for transmission and decodes the received video code which is output to a video display.
- The Audio Codec (G.711, etc.) encodes the audio signal from the microphone for transmission and decodes the received audio code which is output to the loudspeaker.
- The Data Channel supports telematic applications such as electronic whiteboards, still image transfer, file exchange, database access, audiographics conferencing, etc. The standardized data application for real-time audiographics conferencing is Recommendation T.120. Other applications and protocols may also be used via H.245 negotiation as specified in 6.2.7.
- The System Control Unit (H.245, H.225.0) provides signalling for proper operation of the H.323 terminal. It provides for call control, capability exchange, signalling of commands and indications, and messages to open and fully describe the content of logical channels.
- H.225.0 Layer (H.225.0) formats the transmitted video, audio, data and control streams into messages for output to the network interface and retrieves the received video, audio, data, and control streams from messages which have been input from the network interface. In addition, it performs logical framing, sequence numbering, error detection and error correction as appropriate to each media type.

### **6.2.3 Packet based network interface**

The packet based network interface is implementation-specific and is outside the scope of this Recommendation. However, the network interface shall provide the services described in Recommendation H.225.0. This includes the following: Reliable (e.g. TCP, SPX) end-to-end service is mandatory for the H.245 Control Channel, the Data Channels, and the Call Signalling Channel. Unreliable (e.g. UDP, IPX) end-to-end service is mandatory for the Audio Channels, the Video Channels, and the RAS Channel. These services may be duplex or simplex, unicast or multicast depending on the application, the capabilities of the terminals, and the configuration of the network.

#### 6.2.4 Video codec

The video codec is optional. If video capability is provided, it shall be provided according to the requirements of this Recommendation. All H.323 terminals providing video communications shall be capable of encoding and decoding video according to H.261 QCIF. Optionally, a terminal may also be capable of encoding and decoding video according to the other modes of H.261 or H.263. If a terminal supports H.263 with CIF or higher resolution, it shall also support H.261 CIF. All terminals which support H.263 shall support H.263 QCIF. The H.261 and H.263 codecs, on the network, shall be used without BCH error correction and without error correction framing.

Other video codecs, and other picture formats, may also be used via H.245 negotiation. More than one video channel may be transmitted and/or received, as negotiated via the H.245 Control Channel. The H.323 terminal may optionally send more than one video channel at the same time, for example, to convey the speaker and a second video source. The H.323 terminal may optionally receive more than one video channel at the same time, for example, to display multiple participants in a distributed multipoint conference.

The video bit rate, picture format and algorithm options that can be accepted by the decoder are defined during the capability exchange using H.245. The encoder is free to transmit anything that is within the decoder capability set. The decoder should have the possibility to generate requests via H.245 for a certain mode, but the encoder is allowed to simply ignore these requests if they are not mandatory modes. Decoders which indicate capability for a particular algorithm option shall also be capable of accepting video bit streams which do not make use of that option.

H.323 terminals shall be capable of operating in asymmetric video bit rates, frame rates, and, if more than one picture resolution is supported, picture resolutions. For example, this will allow a CIF capable terminal to transmit QCIF while receiving CIF pictures.

When each video logical channel is opened, the selected operating mode to be used on that channel is signalled to the receiver in the H.245 **OpenLogicalChannel** message. The header within the video logical channel indicates which mode is actually used for each picture, within the stated decoder capability.

The video stream is formatted as described in Recommendation H.225.0.

##### 6.2.4.1 Terminal-based continuous presence

H.323 terminals may receive more than one video channel, particularly for multipoint conferencing. In these cases, the H.323 terminal may need to perform a video mixing or switching function in order to present the video signal to the user. This function may include presenting the video from more than one terminal to the user. The H.323 terminal shall use H.245 simultaneous capabilities to indicate how many simultaneous video streams it is capable of decoding. The simultaneous capability of one terminal should not limit the number of video streams which are multicast in a conference (this choice is made by the MC).

#### 6.2.5 Audio codec

All H.323 terminals shall have an audio codec. All H.323 terminals shall be capable of encoding and decoding speech according to Recommendation G.711. All terminals shall be capable of transmitting and receiving A-law and  $\mu$ -law. A terminal may optionally be capable of encoding and decoding speech using Recommendations G.722, G.728, G.729, MPEG 1 audio, and G.723.1. The audio algorithm used by the encoder shall be derived during the capability exchange using H.245. The H.323 terminal should be capable of asymmetric operation for all audio capabilities it has declared within the same capability set, e.g. it should be able to send G.711 and receive G.728 if it is capable of both.

If G.723.1 audio is provided, the audio codec shall be capable of encoding and decoding according to both the 5.3 kbit/s mode and the 6.3 kbit/s mode.

The audio stream is formatted as described in Recommendation H.225.0.

The H.323 terminal may optionally send more than one audio channel at the same time, for example, to allow two languages to be conveyed.

Audio packets should be delivered to the transport layer periodically at an interval determined by the audio codec Recommendation in use (audio frame interval). The delivery of each audio packet shall occur no later than 5 ms after a whole multiple of the audio frame interval, measured from delivery of the first audio frame (audio delay jitter). Audio coders capable of further limiting their audio delay jitter may so signal using the H.245 **maximumDelayJitter** parameter of the **h2250Capability** structure contained within a terminal capability set message, so that receivers may optionally reduce their jitter delay buffers. This is not the same as the RTCP interarrival jitter field.

NOTE – The testing point for the maximum delay jitter is at the input to network transport layer. Network stack, network, driver, and interface card jitter are not included.

#### **6.2.5.1 Audio mixing**

H.323 terminals may receive more than one audio channel, particularly for multipoint conferencing. In these cases, the H.323 terminal may need to perform an audio mixing function in order to present a composite audio signal to the user. The H.323 terminal shall use H.245 simultaneous capabilities to indicate how many simultaneous audio streams it is capable of decoding. The simultaneous capability of one terminal should not limit the number of audio streams which are multicast in a conference.

#### **6.2.5.2 Maximum audio-video transmit skew**

To allow H.323 terminals to appropriately set their receive buffer(s) size, H.323 terminals shall transmit the **h2250MaximumSkewIndication** message to indicate the maximum skew between the audio and video signals as delivered to the network transport. **h2250MaximumSkewIndication** shall be sent for each pair of associated audio and video logical channels. This is not required for audio only or hybrid conferences. Lip synchronization, if desired, shall be achieved via use of time-stamps.

#### **6.2.5.3 Low bit rate operation**

G.711 audio cannot be used in an H.323 conference being carried over low bit rate (< 56 kbit/s) links or segments. An endpoint used for multimedia communications over such low bit rate links or segments should have an audio codec capable of encoding and decoding speech according to Recommendation G.723.1. An endpoint used for audio-only communications over such low bit rate links or segments should have an audio codec capable of encoding and decoding speech according to Recommendation G.729. An endpoint may support several audio codecs in order to provide the widest possible interoperability with those endpoints which only support one low bitrate audio codec. The endpoint shall indicate in the H.245 Capability Exchange procedures at the beginning of each call the capability to receive audio according to the available audio Recommendations which can be supported within the known bit rate limitations of the connection. An endpoint which does not have this low bit rate audio capability may not be able to operate when the end-to-end connection contains one or more low bit rate segments.

The endpoint shall also comply with the requirement of 6.2.5 to be capable of encoding and decoding speech according to Recommendation G.711. However, the endpoint need not indicate this capability if it is sure that it is communicating through a low bit rate segment. If an endpoint is unaware of the presence, in the end-to-end connection, of any links or segments with insufficient

capacity to support G.711 audio (along with other intended media streams, if any), then the endpoint shall declare the capability to receive audio according to Recommendation G.711.

### 6.2.6 Receive path delay

Receive path delay includes delay added to a media stream in order to maintain synchronization and to account for network packet arrival jitter. Media streams may optionally be delayed in the receiver processing path to maintain synchronization with other media streams. Further, a media stream may optionally be delayed to allow for network delays which cause packet arrival jitter. An H.323 terminal shall not add delay for this purpose in its transmitting media path.

Intermediate processing points such as MCUs or Gateways may alter the video and audio time tag information, and shall transmit appropriately modified audio and video time tags and sequence numbers, reflecting their transmitted signals. Receiving endpoints may add appropriate delay in the audio path to achieve lip synchronization.

### 6.2.7 Data channel

One or more data channels are optional. The data channel may be unidirectional or bidirectional depending on the requirements of the data application.

Recommendation T.120 is the default basis of data interoperability between an H.323 terminal and other H.323, H.324, H.320, or H.310 terminals. Where any optional data application is implemented using one or more of the ITU-T Recommendations which can be negotiated via H.245, the equivalent T.120 application, if any, shall be one of those provided. A terminal that provides far-end camera control using H.281 and H.224 is not required to also support a T.120 far-end camera control protocol.

Note that non-standard data applications (**dataApplicationCapability.application = non-standard application**) and Transparent User Data (**dataApplicationCapability.application = userData application**, **dataProtocolCapability = transparent**) may be used whether the equivalent T.120 application is provided or not.

T.120 capability shall be signalled using **dataApplicationCapability.application = t120** application, **dataProtocolCapability = separateLANStack**.

Within the **MediaDistributionCapability**, the **distributedData** structure shall be used if multicast T.120 is available and/or the **centralizedData** structure if unicast T.120 is available. Any node that supports the T.120 data capability shall support the standard T.123 unicast stack.

In the **OpenLogicalChannel** message, the **distribution** choice of the **NetworkAccessParameters** structure is set to **unicast** if T.123 is to be used or **multicast** if Annex A/T.125 is to be used. The **networkAddress** choice is set to **localAreaAddress**, which should always be **unicastAddress**. Within the **IPAddress** sequence, the **network** field is set to the binary IP address and the **tsapIdentifier** is set to the dynamic port on which the T.120 stack will be calling or listening.

The Data channel is formatted as described in Recommendation H.225.0.

#### 6.2.7.1 T.120 data channels

The T.120 connection is established during an H.323 call as an inherent part of the call. Procedures for establishing the T.120 connection prior to the H.323 connection are for further study.

The normal call setup procedures of 8.1 are followed. After the capability exchange takes place, a bidirectional logical channel shall be opened for the T.120 connection according to the normal H.245 procedures indicating that a new connection shall be created as described below.

The opening of a bidirectional logical channel for T.120 may be initiated by either entity sending **openLogicalChannel**, and then following the bidirectional logical channel procedures of Recommendation H.245.

To actually open the logical channel, the initiating entity shall send an **openLogicalChannel** message indicating that a T.120 data channel is to be opened in the **forwardLogicalChannelParameters** as well as in the **reverseLogicalChannelParameters**. The initiator shall include a transport address in the **openLogicalChannel** message. The peer endpoint may choose to ignore the transport address. An endpoint may use a dynamic port number for the T.120 transport address instead of using port 1503 as specified in Recommendation T.123. If the peer (the responder) accepts this logical channel, it shall confirm the opening of the logical channel using **openLogicalChannelAck**. In the **openLogicalChannelAck**, the responder shall include a transport address even if it expects the initiator to originate the T.120 call. In all cases, the transport address for the T.120 connection shall be carried in the **separateStack** parameter, and shall remain valid for the duration of the logical channel.

In the **openLogicalChannel**, the **t120SetupProcedure** choice of the **NetworkAccessParameters** structure can optionally be set to indicate to the responder how the initiator would like to establish the T.120 call. The responder is free to override this preference. **originateCall** indicates that the initiator would like the responder to place the call. **waitForCall** indicates that the initiator would like the responder to receive the call. **issueQuery** is not used when indicating a preference.

In the **openLogicalChannelAck**, the **t120SetupProcedure** choice of the **NetworkAccessParameters** structure should be set to indicate to the initiator how the T.120 call will be established. If neither endpoint has a preference, the T.120 call should be established in the same direction as the H.323 call. **originateCall** tells the initiator to place the call. **waitForCall** tells the initiator that it will receive the call. Whoever originates the call will issue either a join request or an invite request, depending on which endpoint won master/slave determination (the master is always hierarchically higher in the T.120 conference). **issueQuery** can be used by a Gateway to tell the initiator that it must originate the call, and issue a query request to the remote endpoint. It must then set up the T.120 conference in accordance with the contents of the query response (as described in Recommendation T.124).

When possible, the T.120 call should be established in the same direction as the H.323 call. The OLC initiator should not indicate a preference unless there is a need to modify this default behavior. When the initiator indicates a preference, the responder should not override it unless necessary. When no preference is indicated, the responder should specify the default unless there is a need to do otherwise.

In both the **openLogicalChannel** and the **openLogicalChannelAck** messages, the **associateConference** parameter shall be set to **false**.

Recommendation T.120 shall follow the procedures of Recommendation T.123 for the protocol stack indicated in the **dataProtocolCapability** except that the transport addresses as described above shall be employed for connection setup.

If an endpoint is the Active MC or master in a conference which includes T.120, it should also be in control of the T.120 top provider node.

If an endpoint intends to create a conference which includes audio and/or video plus T.120 data, then the H.245 Control Channel shall be established before the T.120 connection is made. This applies to conference create, join, and invite and the actions of an MC. The H.323 call setup procedures shall be used to establish the Active MC (if any), before a T.120 connection is made.

In order to establish a T.120 connection using a GCC-Join request, endpoints are required to know the T.120 conference name. If an alias exists which represents an H.323 conference name (**conferenceAlias**), then the same text which is used for the conference alias should be used as the text

portion of the T.120 conference name. Likewise, the H.323 CID should be used as the numeric T.120 conference name as follows. Each byte of the H.323 CID is converted into a series of three ASCII characters which represent the decimal value of the byte being converted. Note that this requires the value of some CID bytes to be converted such that "0" characters are used for padding. The result will be a string of 48 ASCII characters.

A T.120 MP may be queried for a list of existing conferences. The H.323 CID may be available by converting from the T.120 Numeric Conference name back into the 16-byte octet string. Likewise, the Text Conference name may be used as the H.323 conference alias. Note that a T.124 Conference Query may happen out-of-band from H.323 and prior to an endpoint setting up an H.323 call.

The termination of the associated T.120 conference does not imply the termination of the H.323 call. In other words, closing the T.120 channel shall only affect the Data stream of an H.323 call and shall not affect any other part of the H.323 call. By contrast, when an H.323 call or conference is terminated, then the associated T.120 conference shall also be terminated.

NOTE – The T.120 operation after completion of the connection setup is beyond the scope of this Recommendation.

### **6.2.8 H.245 control function**

The H.245 Control Function uses the H.245 Control Channel to carry end-to-end control messages governing operation of the H.323 entity, including capabilities exchange, opening and closing of logical channels, mode preference requests, flow control messages, and general commands and indications.

H.245 signalling is established between two endpoints, an endpoint and an MC, or an endpoint and a Gatekeeper. The endpoint shall establish exactly one H.245 Control Channel for each call that the endpoint is participating in. This channel shall use the messages and procedures of Recommendation H.245. Note that a terminal, MCU, Gateway, or Gatekeeper may support many calls, and thus many H.245 Control Channels. The H.245 Control Channel shall be carried on logical channel 0. Logical channel 0 shall be considered to be permanently open from the establishment of the H.245 Control Channel until the termination of this channel. The normal procedures for opening and closing logical channels shall not apply to the H.245 Control Channel.

Recommendation H.245 specifies a number of independent protocol entities which support endpoint-to-endpoint signalling. A protocol entity is specified by its syntax (messages), semantics, and a set of procedures which specify the exchange of messages and the interaction with the user. H.323 endpoints shall support the syntax, semantics, and procedures of the following protocol entities:

- Master/slave determination.
- Capability Exchange.
- Logical Channel Signalling.
- Bidirectional Logical Channel Signalling.
- Close Logical Channel Signalling.
- Mode Request.
- Round Trip Delay Determination.
- Maintenance Loop Signalling.

General commands and indications shall be chosen from the message set contained in Recommendation H.245. In addition, other command and indication signals may be sent which have been specifically defined to be transferred in-band within video, audio or data streams (see the appropriate Recommendation to determine if such signals have been defined).

H.245 messages fall into four categories: Request, Response, Command, and Indication. Request and Response messages are used by the protocol entities. Request messages require a specific action by the receiver, including an immediate response. Response messages respond to a corresponding request. Command messages require a specific action, but do not require a response. Indication messages are informative only, and do not require any action or response. H.323 terminals shall respond to all H.245 commands and requests as specified in Annex A, and shall transmit indications reflecting the state of the terminal.

H.323 terminals shall be capable of parsing all H.245 **MultimediaSystemControlMessage** messages, and shall send and receive all messages needed to implement required functions and those optional functions which are supported by the terminal. Annex A contains a table showing which H.245 messages are mandatory, optional, or forbidden for H.323 terminals. H.323 terminals shall send the **functionNotSupported** message in response to any unrecognized request, response, or command messages that it receives.

An H.245 indication, **userInputIndication**, is available for transport of user input alphanumeric characters from a keypad or keyboard, equivalent to the DTMF signals used in analogue telephony, or SBE number messages in Recommendation H.230. This may be used to manually operate remote equipment such as voice mail or video mail systems, menu-driven information services, etc. H.323 terminals shall support the transmission of user input characters 0-9, "\*", and "#". Transmission of other characters is optional.

Three H.245 request messages conflict with RTCP control packets. The H.245 **videoFastUpdatePicture**, **videoFastUpdateGOB** and **videoFastUpdateMB** requests should be used instead of the RTCP control packets Full Intra Request (FIR) and Negative Acknowledgement (NACK). The ability to accept FIR and NACK is signalled during the H.245 capability exchange.

#### 6.2.8.1 Capabilities exchange

Capabilities exchange shall follow the procedures of Recommendation H.245, which provides for separate receive and transmit capabilities, as well as a method by which the terminal may describe its ability to operate in various combinations of modes simultaneously.

Receive capabilities describe the terminal's ability to receive and process incoming information streams. Transmitters shall limit the content of their transmitted information to that which the receiver has indicated it is capable of receiving. The absence of a receive capability indicates that the terminal cannot receive (is a transmitter only).

Transmit capabilities describe the terminal's ability to transmit information streams. Transmit capabilities serve to offer receivers a choice of possible modes of operation, so that the receiver may request the mode which it prefers to receive. The absence of a transmit capability indicates that the terminal is not offering a choice of preferred modes to the receiver (but it may still transmit anything within the capability of the receiver).

The transmitting terminal assigns each individual mode the terminal is capable of operating in a number in a **capabilityTable**. For example, G.723.1 audio, G.728 audio, and CIF H.263 video would each be assigned separate numbers.

These capability numbers are grouped into **alternativeCapabilitySet** structures. Each **alternativeCapabilitySet** indicates that the terminal is capable of operating in exactly one mode listed in the set. For example, an **alternativeCapabilitySet** listing {G.711, G.723.1, G.728} means that the terminal can operate in any one of those audio modes, but not more than one.

These **alternativeCapabilitySet** structures are grouped into **simultaneousCapabilities** structures. Each **simultaneousCapabilities** structure indicates a set of modes the terminal is capable of using simultaneously. For example, a **simultaneousCapabilities** structure containing the two

**alternativeCapabilitySet** structures {H.261, H.263} and {G.711, G.723.1, G.728} means that the terminal can operate either of the video codecs simultaneously with any one of the audio codecs. The **simultaneousCapabilities** set { {H.261}, {H.261, H.263}, {G.711, G.723.1, G.728} } means the terminal can operate two video channels and one audio channel simultaneously: one video channel per H.261, another video channel per either H.261 or H.263, and one audio channel per either G.711, G.723.1, or G.728.

NOTE – The actual capabilities stored in the **capabilityTable** are often more complex than presented here. For example, each H.263 capability indicates details including the ability to support various picture formats at given minimum picture intervals, and the ability to use optional coding modes. For a complete description, see Recommendation H.245.

The terminal's total capabilities are described by a set of **capabilityDescriptor** structures, each of which is a single **simultaneousCapabilities** structure and a **capabilityDescriptorNumber**. By sending more than one **capabilityDescriptor**, the terminal may signal dependencies between operating modes by describing different sets of modes which it can simultaneously use. For example, a terminal issuing two **capabilityDescriptor** structures, one { {H.261, H.263}, {G.711, G.723.1, G.728} } as in the previous example, and the other { {H.262}, {G.711} }, means the terminal can also operate the H.262 video codec, but only with the low-complexity G.711 audio codec.

Terminals may dynamically add capabilities during a communication session by issuing additional **capabilityDescriptor** structures, or remove capabilities by sending revised **capabilityDescriptor** structures. All H.323 terminals shall transmit at least one **capabilityDescriptor** structure.

Non-standard capabilities and control messages may be issued using the **nonStandardParameter** structure defined in Recommendation H.245. Note that while the meaning of non-standard messages is defined by individual organizations, equipment built by any manufacturer may signal any non-standard message, if the meaning is known.

Terminals may re-issue capability sets at any time, according to the procedures of Recommendation H.245.

### 6.2.8.2 Logical channel signalling

Each logical channel carries information from a transmitter to one or more receivers, and is identified by a logical channel number which is unique for each direction of transmission.

Logical channels are opened and closed using the **openLogicalChannel** and **closeLogicalChannel** messages and procedures of Recommendation H.245. When a logical channel is opened, the **openLogicalChannel** message fully describes the content of the logical channel, including media type, algorithm in use, any options, and all other information needed for the receiver to interpret the content of the logical channel. Logical channels may be closed when no longer needed. Open logical channels may be inactive, if the information source has nothing to send.

Most logical channels in this Recommendation are unidirectional, so asymmetrical operation, in which the number and type of information streams is different in each direction of transmission, is allowed. However, if a receiver is capable only of certain symmetrical modes of operation, it may send a receive capability set that reflects its limitations, except where noted elsewhere in this Recommendation. Terminals may also be capable of using a particular mode in only one direction of transmission. Certain media types, including data protocols such as T.120, inherently require a bidirectional channel for their operation. In such cases a pair of unidirectional logical channels, one in each direction, may be opened and associated together to form a bidirectional channel using the bidirectional channel opening procedures of Recommendation H.245. Such pairs of associated channels need not share the same logical channel number, since logical channel numbers are independent in each direction of transmission.

Logical channels shall be opened using the following procedure:

The initiating terminal shall send an **OpenLogicalChannel** message as described in Recommendation H.245. If the logical channel is to carry a media type using RTP (audio or video), the **OpenLogicalChannel** message shall include the **mediaControlChannel** parameter containing the transport address for the reverse RTCP channel.

The responding terminal shall respond with an **OpenLogicalChannelAck** message as described in Recommendation H.245. If the logical channel is to carry a media type using RTP, the **OpenLogicalChannelAck** message shall include both the **mediaTransportChannel** parameter containing the RTP transport address for the media channel and the **mediaControlChannel** parameter containing the transport address for the forward RTCP channel.

Media types (such as T.120 data) which do not use RTP/RTCP shall omit the **mediaControlChannel** parameters.

If a corresponding reverse channel is opened for a given existing RTP session (identified by the RTP **sessionID**), the **mediaControlChannel** transport addresses exchanged by the **OpenLogicalChannel** process shall be identical to those used for the forward channel. Should a collision occur where both ends attempt to establish conflicting RTP sessions at the same time, the master endpoint shall reject the conflicting attempt as described in Recommendation H.245. The rejected **OpenLogicalChannel** attempt may then be retried at a later time.

#### 6.2.8.3 Mode preferences

Receivers may request transmitters to send a particular mode using the H.245 **requestMode** message, which describes the desired mode. Transmitters should comply if possible.

An endpoint receiving the **multipointModeCommand** from the MC shall then comply with all **requestMode** commands, if they are within its capability set. Note that in a decentralized conference, as in a centralized conference, all terminal **requestMode** commands are directed to the MC. The MC may grant the request or not; the basis for this decision is left to the manufacturer.

#### 6.2.8.4 Master-slave determination

The H.245 Master-slave determination procedures are used to resolve conflicts between two endpoints which can both be the MC for a conference, or between two endpoints which are attempting to open a bidirectional channel. In this procedure, two endpoints exchange random numbers in the H.245 **masterSlaveDetermination** message, to determine the master and slave endpoints. H.323 endpoints shall be capable of operating in both master and slave modes. The endpoints shall set **terminalType** to the value specified in Table 1 below and set **statusDeterminationNumber** to a random number in the range 0 to  $2^{24}-1$ . Only one random number shall be chosen by the endpoint for each call, except in the case of identical random numbers, as described in Recommendation H.245.

**Table 1/H.323 – H.323 terminal types for H.245 master-slave determination**

| TerminalType value table                         | H.323 entity |         |            |     |
|--------------------------------------------------|--------------|---------|------------|-----|
|                                                  | Terminal     | Gateway | Gatekeeper | MCU |
| Entity with No MC                                | 50           | 60      | NA         | NA  |
| Entity contains an MC but no MP                  | 70           | 80      | 120        | 160 |
| Entity contains MC with data MP                  | NA           | 90      | 130        | 170 |
| Entity contains MC with data and audio MP        | NA           | 100     | 140        | 180 |
| Entity contains MC with data, audio and video MP | NA           | 110     | 150        | 190 |

The Active MC in a conference shall use a value of 240.

If a single H.323 entity can take part in multiple calls, then the value used for **terminalType** in the master-slave determination process shall be based on the features that the H.323 entity has assigned or will assign to the call in which it is being signalled.

An MC that is already acting as an MC shall always remain the active MC. Therefore, once an MC has been selected as the active MC in a conference, it shall use the Active MC value for all subsequent connections to the conference.

If no MC is active and the entities are of the same type, then the H.323 entity with the highest feature set (as shown in Table 1) shall win the master-slave determination. If no MC is active and the entities are of different types, then an MC that is located in an MCU shall have priority over an MC that is located in a Gatekeeper, which shall have priority over an MC that is located in a Gateway, which in turn shall have priority over an MC located in a terminal.

If an H.323 entity can be associated with two or more of the classifications shown in Table 1, then it should use the highest value for which it qualifies.

#### **6.2.8.5 Timer and counter values**

All timers defined in Recommendation H.245 should have periods of at least as long as the maximum data delivery time allowed by the data link layer carrying the H.245 Control Channel, including any retransmissions.

The H.245 retry counter N100 should be at least 3.

Procedures relating to H.245 protocol error handling are covered in 8.6.

#### **6.2.9 RAS signalling function**

The RAS signalling function uses H.225.0 messages to perform registration, admissions, bandwidth changes, status, and disengage procedures between endpoints and Gatekeepers. The RAS Signalling Channel is independent from the Call Signalling Channel and the H.245 Control Channel. H.245 open logical channel procedures are not used to establish the RAS Signalling Channel. In network environments that do not have a Gatekeeper, the RAS Signalling Channel is not used. In network environments which contain a Gatekeeper (a Zone), the RAS Signalling Channel is opened between the endpoint and the Gatekeeper. The RAS Signalling Channel is opened prior to the establishment of any other channels between H.323 endpoints. This channel is described in detail in clause 7.

### 6.2.10 Call signalling function

The call signalling function uses H.225.0 call signalling to establish a connection between two H.323 endpoints. The Call Signalling Channel is independent from the RAS Channel and the H.245 Control Channel. H.245 open logical channel procedures are not used to establish the Call Signalling Channel. The Call Signalling Channel is opened prior to the establishment of the H.245 Channel and any other logical channels between H.323 endpoints. In systems that do not have a Gatekeeper, the Call Signalling Channel is opened between the two endpoints involved in the call. In systems which contain a Gatekeeper, the Call Signalling Channel is opened between the endpoint and the Gatekeeper, or between the endpoints themselves as chosen by the Gatekeeper. This channel is described in detail in clause 7.

### 6.2.11 H.225.0 layer

Logical channels of video, audio, data or control information are established according to the procedures of Recommendation H.245. Logical channels are unidirectional, and are independent in each direction of transmission. Some logical channels, such as for data, may be bidirectional, and are associated through the bidirectional open logical channel procedure of Recommendation H.245. Any number of logical channels of each media type may be transmitted, except for the H.245 Control Channel of which there shall be one per call. In addition to the logical channels, H.323 endpoints use two signalling channels for call control, and Gatekeeper related functions. The formatting used for these channels shall conform to Recommendation H.225.0.

#### 6.2.11.1 Logical channel numbers

Each logical channel is identified by a Logical Channel Number (LCN), in the range 0 to 65 535, which serves only to associate logical channels with the transport connection. Logical channel numbers are selected arbitrarily by the transmitter, except that logical channel 0 shall be permanently assigned to the H.245 Control Channel. The actual Transport Address that the transmitter shall transmit to shall be returned by the receiver in the **openLogicalChannelAck** message.

#### 6.2.11.2 Logical channel bit rate limits

A logical channel's bandwidth shall have an upper limit specified by the minimum of the endpoint's transmit capability (if present) and the receiving endpoint's receive capability. Based on this limit, an endpoint shall open a logical channel at a bit rate at or below this upper limit. A transmitter shall transmit an information stream within the logical channel at any bit rate at or below the open logical channel bit rate. The limit applies to the information streams which are the content of the logical channel(s), not including RTP headers, RTP payload headers and network headers and other overhead.

H.323 endpoints shall obey to the **flowControlCommand** message of H.245, which commands a limit to the bit rate of a logical channel or the aggregate bit rate of all logical channels. H.323 endpoints that want to limit the bit rate of a logical channel, or the aggregate bit rate of all logical channels should send the **flowControlCommand** message to the transmitting endpoint.

When the terminal has no information to send in a given channel, the terminal shall send no information. Fill data shall not be sent on the network in order to maintain a specific data rate.

## 6.3 Gateway characteristics

The Gateway shall provide the appropriate translation between transmission formats (for example H.225.0 to/from H.221) and between communications procedures (for example H.245 to/from H.242). This translation is specified in Recommendation H.246. The Gateway shall also perform call setup and clearing on both the network side and the SCN side. Translation between video, audio, and

data formats may also be performed in the Gateway. In general, the purpose of the Gateway (when not operating as an MCU) is to reflect the characteristics of a network endpoint to an SCN endpoint, and the reverse, in a transparent fashion.

An H.323 endpoint may communicate with another H.323 endpoint on the same network directly and without involving a Gateway. The Gateway may be omitted if communications with SCN terminals (terminals not on the network) are not required. It may also be possible for a terminal on one segment of the network to call out through one Gateway and back onto the network through another Gateway in order to bypass a router or a low bandwidth link.

The Gateway has the characteristics of an H.323 Terminal or MCU on the network, and of the SCN terminal or MCU on the SCN. The choice of terminal or MCU is left to the manufacturer. The Gateway provides the necessary conversion between the different terminal types. Note that the Gateway may initially operate as a terminal, but later using H.245 signalling begin to operate as an MCU for the same call that was initially point-to-point. Gatekeepers are aware of which terminals are Gateways since this is indicated when the terminal/Gateway registers with the Gatekeeper.

A Gateway which passes T.120 data between the SCN and the network may contain a T.120 MCS Provider which connects the T.120 MCS Providers on the network to the T.120 MCS Providers on the SCN.

Four examples of an H.323 Gateway are shown in Figure 5. The diagrams show the H.323 terminal or MCU function, the SCN terminal or MCU function, and the conversion function. The H.323 terminal function has the characteristics described in 6.2. The H.323 MCU function has the characteristics described in 6.5. The Gateway appears to the other H.323 terminals on the network as one or more H.323 terminals, or an H.323 MCU. It communicates with the other H.323 terminals using the procedures in this Recommendation.

The SCN terminal or MCU function has the characteristics described in the appropriate Recommendation (H.310, H.320, H.321, H.322, H.324, V.70, GSTN or ISDN speech only terminals). The Gateway appears to the terminals on the SCN as one or more of the same terminal types or MCUs. It communicates to another terminal on the SCN using the procedures described in the appropriate Recommendation for that terminal. SCN signalling procedures are beyond the scope of this Recommendation, including such topics as whether the H.323 Gateway appears as a terminal or a network to the SCN. Note that a Gateway may convert H.323 directly to H.324 or H.310 without going to H.320.

Gateways supporting interworking with speech only terminals on GSTN or ISDN should generate and detect DTMF signals corresponding to H.245 **userInputIndications** for 0-9, \*, and #.

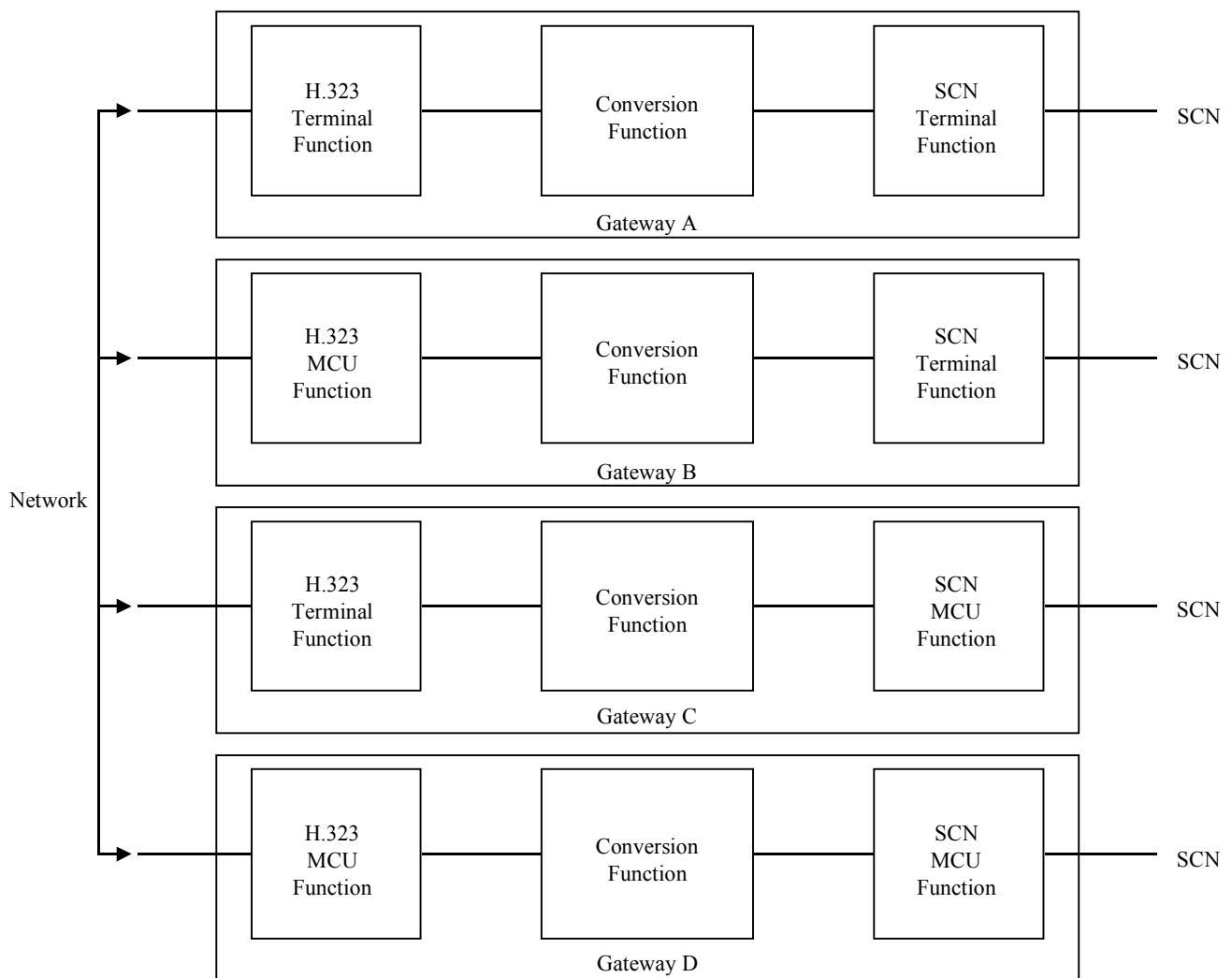
The conversion function provides the necessary conversion of transmission format, control, audio, video, and/or data streams between the different terminal Recommendations. At a minimum, the Gateway shall provide a conversion function for the transmission format, call setup signals and procedures, and communications control signals and procedures. When required, the Gateway shall provide for H.242 to H.245 conversion. The Gateway performs the appropriate conversion between the H.225.0 Call Signalling and the SCN signalling system (Q.931, Q.2931, etc.). The conversion between Q.931 messages on the network and Q.931 messages on the SCN is described in Recommendation H.246.

All call signalling received by the Gateway from an SCN endpoint and not applicable to the Gateway should be passed through to the network endpoint, and vice versa. This signalling includes, but is not limited to, Q.932, Q.950, and H.450-Series messages. This will allow H.323 endpoints to implement the Supplementary Services defined in those Recommendations. The handling of other SCN call signalling systems is for further study.

This Recommendation describes the connection of one H.323 terminal on the network to one external terminal on the SCN through the Gateway. The actual number of H.323 terminals that can communicate through the Gateway is not subject to standardization. Similarly, the number of SCN connections, number of simultaneous independent conferences, audio/video/data conversion functions, and inclusion of multipoint functions is left to the manufacturer. If the Gateway includes an MCU function on the network side, that function shall be an H.323 MCU on the network side. If the Gateway includes an MCU function on the SCN side, it may appear as an H.231/H.243 MCU, or as an MCU for H.310 or H.324 systems (these MCUs are indicated as for further study in the respective Recommendations) on the SCN side.

A Gateway may be connected via the SCN to other Gateways to provide communication between H.323 terminals which are not on the same network.

Equipment which provides transparent interconnection between networks without using H-Series protocols (such as routers and remote dial in units) are not Gateways as defined within the scope of this Recommendation.



T1521240-96

**Figure 5/H.323 – H.323 gateway configurations**

## 6.4 Gatekeeper characteristics

The Gatekeeper, which is optional in an H.323 system, provides call control services to the H.323 endpoints. More than one Gatekeeper may be present and communicate with each other in an unspecified fashion. The Gatekeeper is logically separate from the endpoints; however, its physical implementation may coexist with a terminal, MCU, Gateway, MC, or other non-H.323 network device.

When it is present in a system, the Gatekeeper shall provide the following services:

- Address Translation – The Gatekeeper shall perform alias address to Transport Address translation. This should be done using a translation table which is updated using the Registration messages described in clause 7. Other methods of updating the translation table are also allowed.
- Admissions Control – The Gatekeeper shall authorize network access using ARQ/ACF/ARJ H.225.0 messages. This may be based on call authorization, bandwidth, or some other criteria which is left to the manufacturer. It may also be a null function which admits all requests.
- Bandwidth Control – The Gatekeeper shall support BRQ/BRJ/BCF messages. This may be based on bandwidth management. It may also be a null function which accepts all requests for bandwidth changes.
- Zone Management – The Gatekeeper shall provide the above functions for terminals, MCUs, and Gateways which have registered with it as described in 7.2.

The Gatekeeper may also perform other optional functions such as:

- Call Control Signalling – The Gatekeeper may choose to complete the call signalling with the endpoints and may process the call signalling itself. Alternatively, the Gatekeeper may direct the endpoints to connect the Call Signalling Channel directly to each other. In this manner, the Gatekeeper can avoid handling the H.225.0 call control signals. The Gatekeeper may have to act as the network as defined in Recommendation Q.931 in order to support supplementary services. This operation is for further study.
- Call Authorization – Through the use of the H.225.0 signalling, the Gatekeeper may reject calls from a terminal due to authorization failure. The reasons for rejection may include, but are not limited to, restricted access to/from particular terminals or Gateways, and restricted access during certain periods of time. The criteria for determining if authorization passes or fails is outside the scope of this Recommendation.
- Bandwidth Management – Control of the number of H.323 terminals permitted simultaneous access to the network. Through the use of the H.225.0 signalling, the Gatekeeper may reject calls from a terminal due to bandwidth limitations. This may occur if the Gatekeeper determines that there is not sufficient bandwidth available on the network to support the call. The criteria for determining if bandwidth is available is outside the scope of this Recommendation. Note that this may be a null function, i.e. all terminals are granted access. This function also operates during an active call when a terminal requests additional bandwidth.
- Call Management – For example, the Gatekeeper may maintain a list of ongoing H.323 calls. This information may be necessary to indicate that a called terminal is busy, and to provide information for the Bandwidth Management function.
- Gatekeeper management information data structure – For further study.
- Bandwidth reservation for terminals not capable of this function – For further study.
- Directory services – For further study.

In order to support ad hoc Multipoint Conferences, the Gatekeeper may choose to receive the H.245 Control Channels from the two terminals in a point-to-point conference. When the conference switches to a multipoint conference, the Gatekeeper can redirect the H.245 Control Channel to an MC. The Gatekeeper need not process the H.245 signalling; it only needs to pass it between the terminals or the terminals and the MC.

Networks which contain Gateways should also contain a Gatekeeper in order to translate incoming E.164 or **partyNumber** addresses into Transport Addresses.

H.323 entities that contain a Gatekeeper shall have a mechanism to disable the internal Gatekeeper so that when there are multiple H.323 entities that contain a Gatekeeper on a network, the H.323 entities can be configured into the same Zone.

## **6.5 Multipoint controller characteristics**

The MC provides control functions to support conferences between three or more endpoints in a multipoint conference. The MC carries out the capabilities exchange with each endpoint in a multipoint conference. The MC sends a capability set to the endpoints in the conference indicating the operating modes in which they may transmit. The MC may revise the capability set that it sends to the terminals as a result of terminals joining or leaving the conference, or for other reasons.

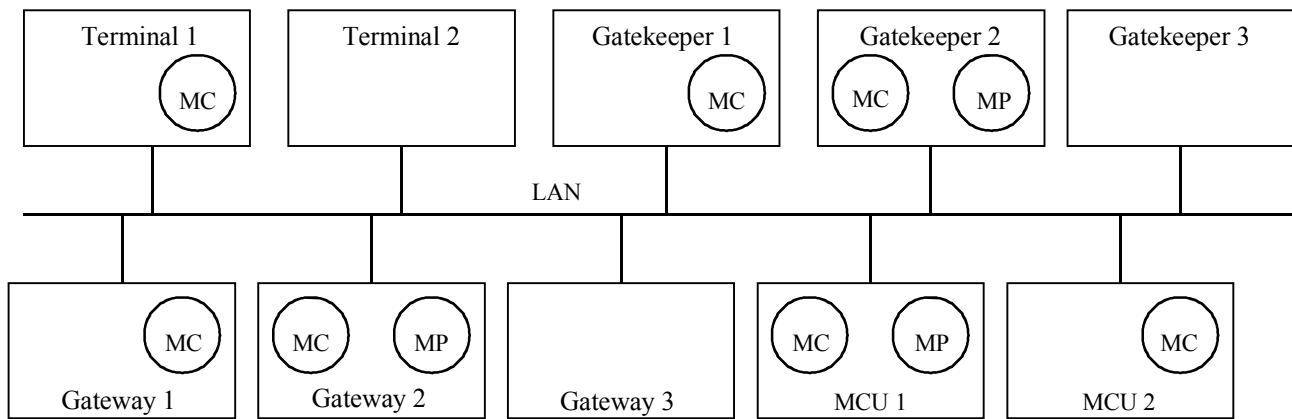
In this manner, the MC determines the Selected Communications Mode (SCM) for the conference. The SCM may be common for all endpoints in the conference. Alternatively, some endpoints may have a different SCM than other endpoints in the conference. The manner in which the MC determines an SCM is not within the scope of this Recommendation.

As part of multipoint conference setup, an endpoint will become connected to an MC on its H.245 Control Channel. This connection may occur:

- via an explicit connection with an MCU;
- via an implicit connection to the MC within a Gatekeeper;
- via an implicit connection to the MC within another terminal or Gateway in the multipoint conference;
- via an implicit connection through a Gatekeeper to an MCU.

The choice of conference mode (e.g. decentralized or centralized) occurs after connection with the MC using H.245 signalling. The choice of conference mode may be limited by the capability of the endpoints or the MC.

The MC may be located within a Gatekeeper, Gateway, terminal, or MCU. See Figure 6.



T1521250-96

NOTE – Gateway, Gatekeeper and MCU can be a single device.

**Figure 6/H.323 – Possible locations of MC and MP in H.323 system**

An MC within a terminal is not callable. It can be included in the call in order to process the H.245 signalling to support ad hoc multipoint conferences. In this case, there may be no distinction between the MC and the H.245 Control Function (see 6.2.8) of the terminal. Communications between them are outside the scope of this Recommendation.

An MC located with the Gatekeeper is not callable; however, an MCU located with a Gatekeeper may be callable. An MCU located with a Gatekeeper may function as an independent MCU. An MC located with a Gatekeeper may be used to support ad hoc multipoint conferences when the Gatekeeper receives the H.245 Control Channels from the endpoints. In this manner, the Gatekeeper can route the H.245 Control Channels to the MC at the start of the call or when the conference switches to multipoint.

The Gateway can function as a terminal or an MCU. When functioning as a terminal, the Gateway may contain an MC. This has the same characteristics as described above for an MC within a terminal.

An MCU always contains an MC. The MCU is callable and the MC processes the H.245 Control Channel from all of the endpoints.

When two or more endpoints are in a conference, the endpoints shall use the master slave resolution procedure of Recommendation H.245 to determine the MC that will control the conference.

After the capability exchange and master/slave determination, the MC may first assign a terminal number to a new endpoint using the **terminalNumberAssign**. The MC shall then notify the other endpoints of the new endpoint in the conference using **terminalJoinedConference**. The new endpoint may request a list of other endpoints in the conference using the **terminalListRequest**.

## 6.6 Multipoint processor characteristics

The MP receives audio, video and/or data streams from the endpoints involved in a centralized or hybrid multipoint conference. The MP processes these media streams and returns them to the endpoints.

Communications between the MC and the MP are not subject to standardization.

The MP may process one or more media stream types. When the MP processes video, it shall process the video algorithms and formats as described in 6.2.4. When the MP processes audio, it shall

process the audio algorithms as described in 6.2.5. When the MP processes data, it shall process data streams as described in 6.2.7.

An MP which processes video shall provide either video switching or video mixing. Video switching is the process of selecting the video that the MP outputs to the terminals from one source to another. The criteria used to make the switch may be determined through detection of a change in speaker (sensed by the associated audio level) or through H.245 control. Video mixing is the process of formatting more than one video source into the video stream that the MP outputs to the terminals. An example of video mixing is combining four source pictures into a two-by-two array in the video output picture. The criteria for which sources and how many are mixed is determined by the MC until other controls are defined. The use of the T.120-Series Recommendations for these control functions is for further study.

An MP which processes audio shall prepare N-audio outputs from M-audio inputs by switching, mixing, or a combination of these. Audio mixing requires decoding the input audio to linear signals (PCM or analogue), performing a linear combination of the signals and recoding the result to the appropriate audio format. The MP may eliminate or attenuate some of the input signals in order to reduce noise and other unwanted signals. Each audio output may have a different mix of input signals providing for private conversations. The terminals shall assume that their audio is not present in the audio stream returned to them. Terminal removal of its own audio from the MP audio output is for further study.

An MP which processes T.120 data shall be capable of acting as a non-leaf MCS provider and should be capable of acting as the Top MCS Provider. An MP may also process non-standard data, transparent user data and/or other types of data.

The MP may provide algorithm and format conversion, allowing terminals to participate in a conference at different SCMs.

The MP is not callable, the MCU which it is a part of is callable. The MP terminates and sources the media channels.

## **6.7 Multipoint control unit characteristics**

The MCU is an endpoint which provides support for multipoint conferences. The MCU shall consist of an MC and zero or more MPs. The MCU uses H.245 messages and procedures to implement features similar to those found in Recommendation H.243.

A typical MCU that supports centralized multipoint conferences consists of an MC and an audio, video and data MP. A typical MCU that supports decentralized multipoint conferences consists of an MC and a data MP supporting Recommendation T.120. It relies on decentralized audio and video processing.

The network side of a Gateway may be an MCU. A Gatekeeper may also include an MCU. In either case, they are independent functions that happen to be co-located.

The MCU shall be callable by other endpoints using the procedures of clause 8.

## **6.8 Multipoint capability**

### **6.8.1 Centralized multipoint capability**

All endpoints shall have centralized multipoint capability. In this mode of operation they communicate with the MC of the MCU in a point-to-point manner on the control channel and with the MP on the audio, video and data channels. In this mode, the MC performs H.245 multipoint control functions, while the MP performs video switching or mixing, audio mixing, and T.120

multipoint data distribution. The MP transmits the resulting video, audio and data streams back to the endpoints. The MP may have the capability to convert between different audio, video and data formats and bit rates, allowing the endpoints to participate in the conference using different communications modes.

The MCU may use multicast to distribute the processed media streams if the endpoints in the conference can receive multicast transmissions. Multicast distribution of data is for further study.

This mode is signalled by the following H.245 capabilities: **centralizedControl**, **centralizedAudio**, **centralizedVideo**, and **centralizedData**. Optionally, **distributedAudio** and **distributedVideo** may be used to indicate multicast distribution of media streams.

### 6.8.2 Decentralized multipoint capability

If the endpoints have decentralized multipoint capability, they communicate with the MC of an MCU, Gateway, Gatekeeper, or endpoint in a point-to-point mode on the H.245 Control Channel and optionally with an MP on data channels. The endpoints shall have the capability to multicast their audio and video channels to all other endpoints in the conference. The MC may control which endpoint or endpoints are actively multicasting audio and/or video (for example by using the **flowControlCommand** on either channel).

The endpoints receive multicast video channels and select one or more of the available channels for display to the user. The endpoints receive the multicast audio channels and perform an audio mixing function in order to present a composite audio signal to the user.

The MC may provide conference control functions such as chair control, video broadcast and video selection. This shall be done by receiving H.245 from an endpoint and then sending the appropriate control to other endpoints to enable or disable their video multicast. T.120 commands may optionally provide the same functions.

This mode is signalled by the following H.245 capabilities: **centralizedControl**, **distributedAudio**, **distributedVideo**, and **centralizedData**.

### 6.8.3 Hybrid multipoint – Centralized audio

If the endpoints and MCU have hybrid multipoint-centralized audio capability, they may use distributed multipoint for video and centralized multipoint for audio. In this mode, the endpoints communicate with the MC in a point-to-point mode on the H.245 Control Channel and optionally with an MP on data channels.

The endpoints shall have the capability to multicast their video channels to all other endpoints in the conference. The MC may control which endpoint or endpoints are actively multicasting video. The endpoints receive multicast video channels and select one or more of the available channels for display to the user.

All of the endpoints in the conference transmit their audio channels to the MP. The MP performs the audio mixing function and outputs the resulting audio streams to the endpoints. The MP may produce an exclusive audio sum for each endpoint in the conference. Multicast distribution of processed audio is for further study.

This mode is signalled by the following H.245 capabilities: **centralizedControl**, **centralizedAudio**, **distributedVideo**, and **centralizedData**.

### 6.8.4 Hybrid multipoint – Centralized video

If the endpoints and MCU have hybrid multipoint-centralized video capability, they may use distributed multipoint for audio and centralized multipoint for video. In this mode, the endpoints

communicate with the MC in a point-to-point mode on the H.245 Control Channel and optionally with an MP on data channels.

The endpoints shall have the capability to multicast their audio channels to all other endpoints in the conference. The MC may control which endpoint or endpoints are actively multicasting audio. The endpoints receive multicast audio channels and perform a mixing function in order to present a composite audio signal to the user.

All of the endpoints in the conference transmit their video channels to the MP. The MP performs the video switching, mixing, or format conversion functions and outputs the resulting video streams to the endpoints. The MP may produce an exclusive video stream for each endpoint in the conference, or it may multicast a video stream to all participating endpoints, in order to minimize the bandwidth used on the network.

This mode is signalled by the following H.245 capabilities: **centralizedControl**, **distributedAudio**, **centralizedVideo**, and **centralizedData**.

#### **6.8.5 Establishment of common mode**

The MC shall coordinate a common communications mode between the endpoints in the multipoint conference. The MC may force endpoints into a particular common mode of transmission (as allowed by their capability sets) by sending to the endpoint a receive capability set listing only the desired mode of transmission, or the MC may rely on **multipointModeCommand** and mode preference commands to enforce mode symmetry. The latter approach should be used since it allows the endpoints to know the full range of conference capabilities available that can be requested.

If the MCU has the capability to convert audio and/or video formats, it may not be necessary to force all endpoints into the same communications mode.

#### **6.8.6 Multipoint rate matching**

Since the endpoints on each link in a multipoint configuration may attempt to operate at different bit rates, the MC shall send H.245 **flowControlCommand** messages to limit the transmitted bit rates to those which can be sent to receivers.

#### **6.8.7 Multipoint lip synchronization**

An MP which is providing audio mixing in either the Centralized or Hybrid multipoint conferences shall modify the time tags of the audio and video streams, taking into account its own time base, in order to maintain audio and video synchronization. Further, when the MP processes the audio and/or video to generate a new stream sourced from the MP, the MP shall generate its own sequence numbers in the audio and video packets.

When mixing audio, the MP should synchronize each of the incoming audio streams to its own timing, mix the audio streams, and then shall generate a new audio stream based on its own timing with its own sequence numbers. If the MP is also switching video, the switched stream shall have its original time-stamp replaced with the MP time base to synchronize it with the mixed audio stream and shall have a new sequence number representing the stream from the MP.

In the case of distributed multipoint conferences, the receiving endpoint may be able to maintain lip synchronization by aligning the selected video stream and its associated audio by using the RTP time tags. Alignment of the other audio streams may not be necessary. If multiple video streams are displayed, the associated audio streams should be aligned.

It may not be possible to guarantee lip synchronization in hybrid multipoint conferences.