INTERNATIONAL TELECOMMUNICATION UNION

ITU-T

TELECOMMUNICATION
STANDARDIZATION SECTOR
OF ITU

H.332
(09/98)

SERIES H: AUDIOVISUAL AND MULTIMEDIA SYSTEMS

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

---

H.323 extended for loosely coupled conferences

ITU-T Recommendation H.332

(Previously CCITT Recommendation)

ITU-T H-SERIES RECOMMENDATIONS

AUDIOVISUAL AND MULTIMEDIA SYSTEMS

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

ITU-T RECOMMENDATION H.332

H.323 EXTENDED FOR LOOSELY COUPLED CONFERENCES

Summary

This Recommendation describes terminals, equipment, and services for multimedia communication over packet-based networks. H.332 terminals and equipment may carry real-time voice, data, and video in any combination. The goal is to work within the framework of Recommendation H.323, yet provide scalability well beyond the limits of Recommendation H.323 to enable applications involving hundreds and thousands of participants.

H.332 terminals may be integrated into personal computers or implemented in stand-alone devices such as videotelephones. Support for voice 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. Like Recommendation H.323, this Recommendation allows more than one channel of each type to be in use. Other Recommendations in the H.332 series include H.323 visual telephone equipment for Local Area Networks, H.225.0 packet and synchronization, H.245 control, H.261 and H.263 video codecs, G.711, G.722, G.728, G.729, and G.723.1 audio codecs, and the T.120 series of multimedia communications protocols.

H.332 terminals are based on Recommendation H.323; therefore, they interwork via gateways 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, and V.70 terminals on GSTN.

Source

ITU-T Recommendation H.332 was prepared by ITU-T Study Group 16 (1997-2000) and was approved under the WTSC Resolution No. 1 procedure on the 25th of September 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 not 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.

CONTENTS

Page

Recommendation H.332

H.323 EXTENDED FOR LOOSELY COUPLED CONFERENCES

(Geneva, 1998)

1 Scope

This Recommendation covers highly scalable conferencing on packet-based networks involving hundreds and thousands of participants. This type of conference is defined in Recommendation H.323 as broadcast and broadcast-panel conferences. This Recommendation is based on Recommendation H.323, thus providing interoperability with circuit-switched endpoints (H.320, H.321, H.324) via Gateways as shown in Figure 1. Use of data, as opposed to real-time streams like audio and video, is for further study.

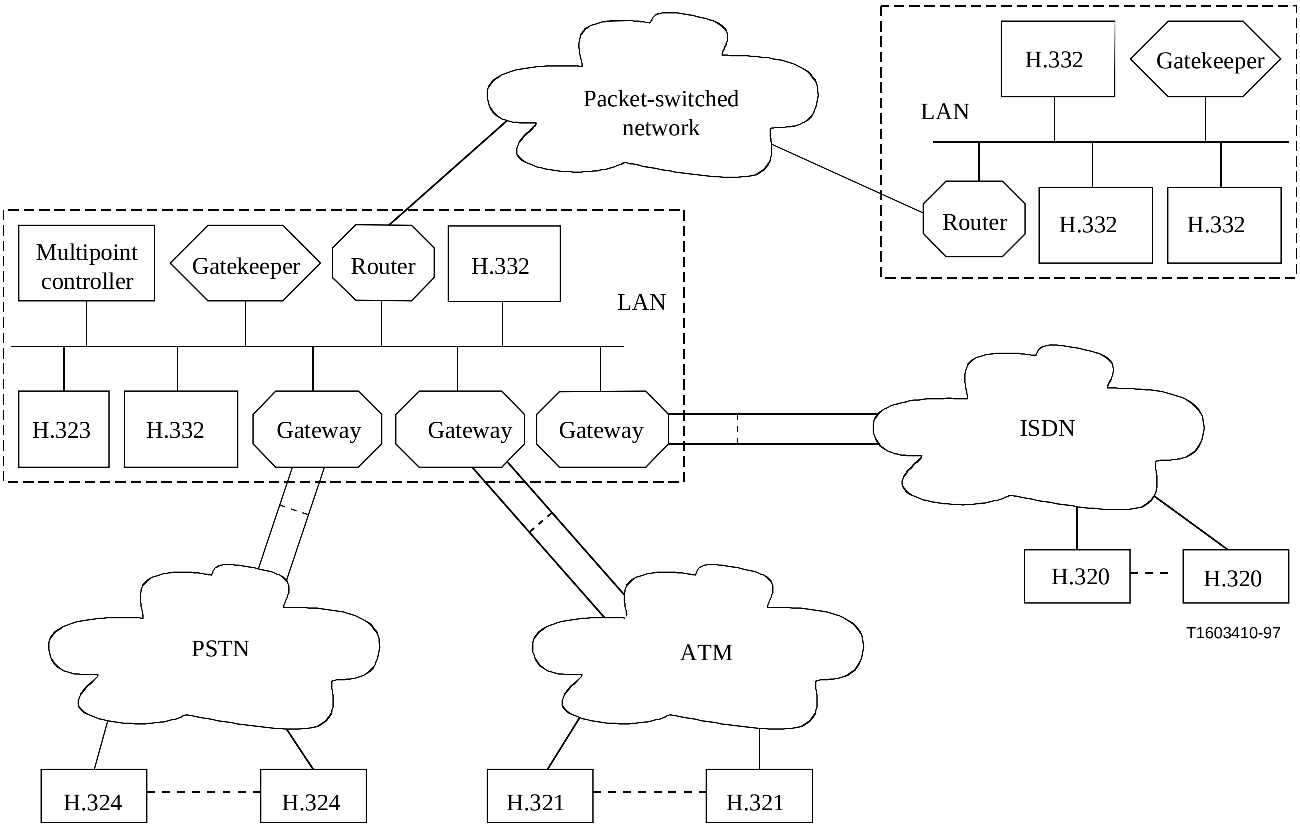

Figure 1/H.332 – Scope of H.332 conference

2 Conference model

The goal of this Recommendation is to provide scalability of the number of terminals in a conference well beyond the limits of Recommendation H.323. Recommendation H.323 is limited in scalability due to the requirements for tightly coupled conferencing. It requires all participants of the conference to be known and a set of procedures for conference setup, capability negotiation, creation and control of audio/video/data streams, and conference tear down. Such a procedure for conference setup and capability negotiation at the start of the conference, though essential for tightly coupled conferences, is cumbersome and impractical for conferences involving an arbitrarily large number of participants. Information required to setup a large conference must be disseminated well before the start of the conference.

Small conferences, like telephone calls, are usually ad hoc in nature. In comparison, large conferences are usually planned and pre-announced. Examples include presentation to a large geographically dispersed audience, distance learning, etc. When a conference is pre-announced, conference capabilities can also be pre-announced. Some form of capability negotiation is also desirable to accommodate users at different network bandwidth and endpoint resources (Processor power, display resolution, etc.). Network bandwidth may span from low-speed modem links of 14.4 kbit/s to high speed B-ISDN links of 622 Mbit/s. Layered video can be used to accommodate participants at different bandwidth links and their need for varying picture quality.

Loosely coupled conferencing is already part of RTP/RTCP. It is designed to scale to thousands of participants. Over a period of time, the identity of each participant is known through the RTCP messages. Floor control is provided through human or social control. Human control works well for small conferences or in a setting where everybody participating can see each other. But when conferences consist of hundreds of participants, some form of automatic chair-control mechanism is desirable. RTCP is constrained to occupy a small percentage of the total conference bandwidth and any chair-control mechanism implemented in RTCP will not provide service in a timely manner. Therefore, the concept of an H.323 panel is introduced here.

The basic model of this Recommendation is shown in Figure 2. The panel consists of a small H.323 conference connected to a large number of RTP receiving terminals via RTP/RTCP. These RTP receiving terminals may be H.332 terminals, or other RTP/RTCP capable terminals that have external means to understand how to connect to the conference. Within the panel, full interaction is allowed. Interaction could be through social- or chair- control. Outside the panel, the participants are passive; they are essentially receivers who are, by default, not allowed to interact. If they wish to interact, they have to join the panel or get invited by the panel. Inside the panel, any H.323 model – centralized, decentralized, or hybrid – can be used. However, outside the panel multicast is used in order to provide the scalability required for the H.332 conference. This can be achieved either by using the H.323 decentralized model, or, when the centralized model is used, by using a Multipoint Processor (MP) to multicast media streams to the RTP receiving terminals.

The panel consists of permanent and temporary members and is limited in size only by the amount of resources available at the Multipoint Controller (MC) for the conference. The permanent members consist of those who are essential to the conference such as the teacher in a distance-learning virtual class or the presenter in a virtual auditorium full of people. The temporary members come from the RTP receiving terminals who want to participate with questions, discussions, etc. Temporary members change with time as new members join and old members leave voluntarily or are requested to leave to make room for new ones.

The panel employs social or automatic control. If social control is used, then all in the panel can potentially talk and send their video on the audio and video RTP sessions. But, in practice, through social control, one participant usually talks at a time. Automatic control, on the other hand, is through H.323 chair-control. H.323 chair-control gives special privileges to the chair. Any panel member who wants to talk and send video must first request the floor from the chair. When the floor is granted by the chair, the participant can use the audio and video RTP sessions.

The sequence of events in an H.332 conference is as follows:

1) The conference is pre-announced with sufficient information to enable discovery and participation. The IETF Session Description Protocol (SDP) shall be used to encode the conference announcement. Any mechanism can be used to carry the announcement. Some suggested mechanisms are: email (SMTP) and web (HTTP).

2) If the conference limits participation for security, registration or fee-paying reasons, then the public announcement shall contain information on how to register and obtain a private announcement with the encryption key(s) and algorithm(s), and any other private information.

3) Capability negotiation may be done prior to the conference. If capabilities change, a new announcement shall be created and sent. Even though the capabilities of the conference have been determined prior to the conference, the panel may change the capabilities during the conference. However, changing capabilities during the conference is not advisable as it may force some RTP receiving terminals out of the conference because of the mismatch in capabilities.

4) Before the start of the conference, permanent members of the panel either join or are invited according to the rules of Recommendation H.323. The method through which permanent members are selected is beyond the scope of this Recommendation.

5) The conference starts with a small H.323 panel of permanent members. RTP receiving terminals join the conference to receive streams from the panel.

6) During the conference, RTP receiving terminals may either join the panel or be invited by the panel to participate as temporary panel members.

7) The panel participants shall join/leave the H.323 panel according to the rules of Recommendation H.323. The RTP receiving terminals join/leave the conference according to the rules of RTP/RTCP.

8) The H.332 conference ends when the H.323 panel conference ends according to the procedures of Recommendation H.323.

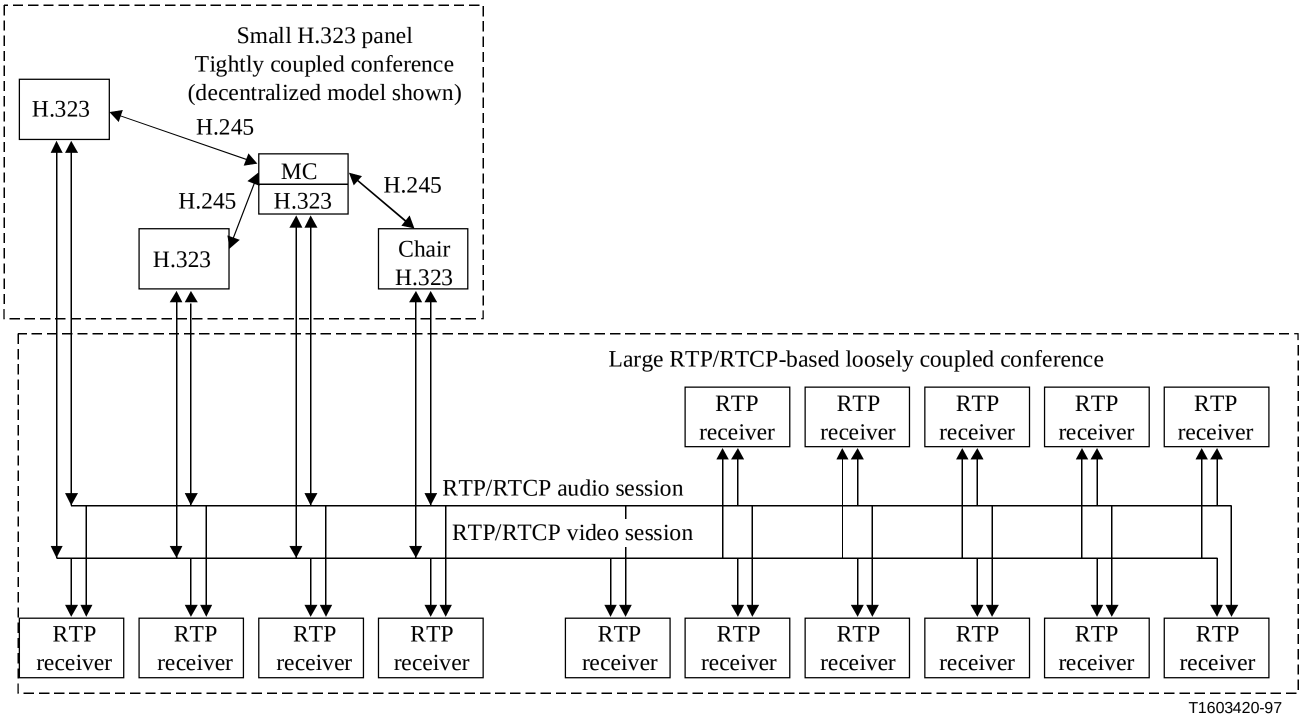

Figure 2/H.332 – A large conference consisting of an H.323 panel and RTP/RTCP-based receivers

3 Normative references

The following ITU-T Recommendations and other references contain provisions which, through reference in this text, constitute provisions of this Recommendation. At the time of publication, the editions indicated were valid. All Recommendations and other references are subject to revision; all users of this Recommendation are therefore encouraged to investigate the possibility of applying the most recent edition of the Recommendations and other references listed below. A list of the currently valid ITU-T Recommendations is regularly published.

[1] ITU-T Recommendation H.225.0 (1998), *Call signalling protocols and media stream packetization for packet-based multimedia communication systems.*

[2] ITU-T Recommendation H.245 (1998), *Control protocol for multimedia communication*.

[3] ITU-T Recommendation H.323 (1998), *Packet-based multimedia communications systems.*

[4] IETF RFC 2327 – *SDP*: *Session Description Protocol*.

[5] ITU-T Recommendation Q.931 (1998), *User-network interface layer 3 specification for basic call control*.

4 Definitions

For the purposes of this Recommendation the definitions given in clause 3 of Recommendations H.225.0 [1], H.245 [2], and H.323 [3] apply. These definitions apply to the Local Area Network (LAN) side only. Other terms may be appropriate when referring to the Circuit-Switched Network (CSN) side.

5 Symbols and abbreviations

This Recommendation uses the following symbols and abbreviations:

B-ISDN	Broadband Integrated Services Digital Network
CID	Conference Identifier
CSN	Circuit-Switched Network
GSTN	General Switched Telephone Network
HTTP	Hypertext Transfer Protocol
IETF	Internet Engineering Task Force
IGMP	Internet Group Management Protocol
IP	Internet Protocol
IP v4	Internet Protocol Version 4
LAN	Local Area Network
MC	Multipoint Controller
MCU	Multipoint Control Unit
MP	Multipoint Processor
RAS	Registration Admission Status
RFC	Request for Comments
RSVP	Resource Reservation Protocol
RTCP	Real-time Transport Control Protocol
RTP	Real-time Transport Protocol

SDES	Source Description
SDP	Session Description Protocol
SMTP	Simple Mail Transfer Protocol
SSRC	Synchronization Source
TTL	Time To Live
URI	Universal Resource Identifier
URL	Universal Resource Locator
UUID	Universally Unique Identifier

6 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.

Where items exist on both the LAN and on the CSN, references to the CSN item will be explicit. For example, an MCU is an H.323 MCU on the LAN, an CSN MCU is an MCU on the CSN.

This Recommendation describes the use of four different message types: H.245, RAS, Q.931, and RTCP. 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 in bold and italics typeface with the first letters capitalized (***Call Proceeding***). RTCP messages are capitalized with italics typeface (*CNAME*).

7 Announcement encoding

The IETF Session Description Protocol (SDP) is used to encode the conference announcement.

7.1 Extensions

Since SDP does not provide all the information needed to encode an H.332 announcement, it is extended here using the guidelines for extension specified in SDP as follows:

• a=type:<conference type>

This record specifies the type of the conference. Suggested values in the SDP specification are "broadcast", "meeting", and "moderated". This Recommendation adds a new value "H332" to specify that the type of the conference is H.332. H.332 conferences shall default to receive-only for the RTP receiving terminals. This record shall be present in SDP announcements to allow the SDP parser to determine if the announcement is for an H.332 conference.

• m=<media> <port> <transport> <fmt list>

This record specifies the media announcements. It is extended in this Recommendation to provide the control to do capability negotiation before the conference and be able to join the panel during the conference.

The first sub-field is the media type. Suggested values in the SDP specification are "audio", "video", "whiteboard", "text" and "data". This Recommendation adds a new value "control" to specify an external control.

The second sub-field is the transport port to which the control applies. It could be the well-known call-signalling port of Recommendation H.323 if the call-signalling address of the MC or the capability negotiation server will be provided in the connection record.

The third sub-field is the transport protocol. This Recommendation adds a new value "H323" to specify that the procedures of Recommendation H.323 will be used for control.

The fourth and subsequent sub-fields are media formats. This Recommendation adds two new values: "mc" and "caps".

It is possible that a server may have means to allow clients to negotiate the capabilities required to access the conference. This negotiation is done using methods in 8.3. If the clients are allowed to negotiate their capabilities with the server, the call-signalling address of the capability negotiation server shall be provided in the connection "c" record, as defined in SDP. If the capabilities change as a result of the negotiation, a new announcement shall be made.

If the RTP receiving terminals are allowed to join the panel during the conference, the call-signalling address of the MC shall be provided in the connection "c" record.

As an example, the following two SDP records specify the call-signalling address of the capability negotiation server:

m=control 1720 H323 caps

c=IN IP4 134.134.157.81

In another example, the following two SDP records specify the call-signalling address of the MC:

m=control 1720 H323 mc

c=IN IP4 134.134.157.81

- The "origin" record in SDP "o=<username> <session id> <version> <network type> <address type> <address>" contains a session identifier field which is text based; it contains any printable 8-bit ISO 8859-1 character with the exception of 0x0a (newline) and 0x0d (carriage return). The conference identification (CID) of Recommendation H.323 shall be used in the session identifier field. The SDP parser shall consider the session identifier field as the CID when the announcement is for the H.332 conference as specified in the "a=type:H332" record. The CID is in UUID form as shown in Recommendation H.225.0. The methods for creating a UUID is also shown in Recommendation H.225.0. The UUID must be converted to human readable text, using 8-bit ISO 8859-1 characters to form a string representation. A UUID string representation is specified as a sequence of fields, some of which are separated by single dashes. Each field is treated as an integer and has its value printed as a zero-filled hexadecimal digit string with the most significant digit first. The hexadecimal values "a" to "f" inclusive are output as lower case characters, and are case insensitive on input. The sequence is the same as the UUID constructed type.

7.2 Grammar

The records of the H.332 announcement are shown below with the SDP-extended records in italics. The "type" record is the only mandatory SDP-extended record. This subclause contains changes to the grammar in the SDP document; therefore it is intended to be used along with the grammar section in the SDP document:

```
announcement ::=             proto-version
                             origin-field
                             session-name-field
                             information-field
                             uri-field
                             email-fields
                             phone-fields
                             connection-field
                             bandwidth-fields
                             time-fields
                             key-field
                             type-field
                             attribute-fields
                             media-descriptions

type-field ::=               "a=type:" conferencetype [newline]

conferencetype ::=           (ALPHA)+
                             ;typically "broadcast", "meeting", "moderated", or "H332"

key-data ::=                 encryption-algorithm: encryption-key

encryption-algorithm ::=     printable-ascii
                             ;syntax is that of the object identifier

encryption-key ::=           printable-ascii

media-field ::=              "m=" media space port ["/" integer] space proto (space fmt)+ newline

media ::=                    (alpha-numeric)+
                             ;typically "audio", "video", "whiteboard", "text", or "control"

proto ::=                    (alpha-numeric)+
                             ;typically "RTP/AVP", "VAT", "UDP" for IP4, or "H323"

fmt ::=                      (alpha-numeric)+
                             ;typically an RTP payload type, "mc", or "caps"

sess-id ::=                  <time_low> <hyphen> <time_mid> <hyphen>
                             <time_high_and_version> <hyphen>

                             <clock_seq_and_reserved> <clock_seq_low> <hyphen> <node>

time_low ::=                 <hexOctet> <hexOctet> <hexOctet> <hexOctet>

time_mid ::=                 <hexOctet> <hexOctet>

time_high_and_version ::=    <hexOctet> <hexOctet>

clock_seq_and_reserved ::=<hexOctet>

clock_seq_low ::=            <hexOctet>

node ::=                     <hexOctet><hexOctet><hexOctet><hexOctet><hexOctet><hexOctet>

hexOctet ::=                 <hexDigit> <hexDigit>

hexDigit ::=                 <digit> | <a> | <b> | <c> | <d> | <e> | <f>

digit ::=                    "0" | "1" | "2" | "3" | "4" | "5" | "6" | "7" | "8" | "9"

hyphen ::=                   "-"
```

a ::=	"a" | "A"
b ::=	"b" | "B"
c ::=	"c" | "C"
d ::=	"d" | "D"
e ::=	"e" | "E"
f ::=	"f" | "F"

7.3 Default records

The default records in SDP are as follows:

- RTP sessions shall be receive-only (a=recvonly) unless explicitly specified otherwise. The RTP receiving terminals shall not send any information on these RTP sessions until they join the panel and follow the procedures of Recommendation H.323 to determine when they are allowed to send information.

8 Procedures for conference establishment

The sequence of events to establish an H.332 conference are:

1) Announcement: The conference is pre-announced with sufficient information to enable discovery and participation. SDP shall be used to encode the announcement.
2) Registration: If the conference limits participation, then the announcement shall contain information on how to register and obtain a private announcement.
3) Capability Negotiation: The conference mode is specified prior to the conference. If the mode changes, a new announcement shall be created and sent. Even though the capabilities of the conference have been determined prior to the conference, the panel may change the capabilities during the conference. However, changing capabilities during the conference is not advisable as it may force some RTP receiving terminals out of the conference because of the mismatch in capabilities.
4) Invitation: Before the start of the conference, permanent members of the panel either join or are invited.

8.1 Announcement

The information in the conference announcement shall have sufficient information to allow H.332 terminals to join the conference if the media capabilities match. Even though the announcement is meant for all terminals both inside and outside of the panel, it shall be written from the perspective of the RTP receiving terminal due to the fact that such terminals do not have access to the control channel during the conference. The announcement consists of the following:

- conference identification;
- conference start and stop time;
- mode of each RTP and Data sessions which include the media type (video, audio, data, etc.), the coder (G.723.1, H.263, etc.), and so on;
- optional information on how to register and obtain the encryption key(s) and algorithm(s)Addresses and ports for each RTP session;
- optional URL to more information on the conference, such as slides for downloading, documents, payment method, etc.;

- optional address of the capability negotiation server (MC);
- optional MC address to provide back-channel to the RTP receiving terminals.

The SDP announcements shall have sufficient information so that non-H.332 terminals which are SDP and RTP/RTCP compliant can join the H.332 conference as RTP receiving terminals. Such non-H.323 terminals will ignore extended SDP attributes that are not understood. Also, being non-H.323 terminals, they will not be able to join the panel.

8.2 Registration

If the conference is limited to registered participants, the streams may be encrypted. The public announcement shall then contain information on how to register and obtain a private announcement. It is up to the announcer to determine how much information is revealed in the public announcement. For example, the MC address may not be revealed in the public announcement. The private announcement shall have all the information about the conference including the encryption key(s) and algorithm(s) as well as the public announcement fields.

The public announcement would contain key-field of the form:

key-field ::= "k=" key-type

key-type ::= "uri:" uri | "prompt"

The user accesses the uri to perform the necessary registration. After registration, the user is sent a private announcement on a secure channel (for example, on a secure http session, or via secure e-mail). The private announcement contains the key(s) required to decipher the streams, along with the name of the encryption algorithm(s) used and their mode(s) of operation. Thus:

key-field ::= "k=" key-type

key-type ::= "clear:" key-data | "base64:" key-data

This information could be common for the conference, or specified on a per RTP session basis but shall not be both.

8.3 Capability negotiation

H.332 clients may do capability negotiation if the announcement provides the address of the capability negotiation server. The capability negotiation server may also be the MC of the announced conference.

The procedures outlined in Recommendation H.323 shall be followed. In summary, the client shall establish a call-signalling connection with the Server and in ***Setup*** specify conferenceGoal = capability_negotiation. Then, after the control connection is established with the server, the client shall send **terminalCapabilitySet** to the server. The Server shall respond with **terminalCapabilitySetAck** or **terminalCapabilitySetReject** followed by endSession which will terminate the conference. A response of **terminalCapabilitySetReject** indicates that the server was unable to understand or store the capabilities. A response of **terminalCapabilitySetAck** indicates that the capabilities were received and stored by the server. It does not imply that the conference capabilities will be changed. To find that out, the client must listen for a revised announcement of the conference with the same CID as before. The server shall send a revised announcement if it has changed the capabilities of the conference.

8.4 Invitation

The procedures outlined in Recommendation H.323 shall be used to invite permanent members of the panel at the conference start time. Permanent members can also join the conference at the conference start time by following the procedures outlined in Recommendation H.323.

9 Procedures during the conference

The RTP receiving terminals are usually passive; they can only receive information. If they want a back-channel to ask questions or become part of the panel discussion, they have to either join the panel or get invited by the panel. Once the RTP receiving terminals join the panel, they become H.323 terminals and participate according to the rules of the panel; when they leave they become RTP receiving terminals again. Because the panel is limited in size, usually due to the resource constraints of the MC, only a small subset of RTP receiving terminals will be able to join the panel at one time. Others will have to wait for their turn. Note that it is possible for an SDP announcement to allow back-channel communication from outside the panel but this is not expected to be normally used for H.332 conferences.

In order to join the panel, the RTP receiving terminals need to know the address of the MC. The conference announcement provides the MC address if the terminals are allowed to join the panel.

All terminals in the conference are periodically sending the user's real name in the RTCP SDES item *NAME* which enables each terminal to build up a roster over a period of time. The RTP receiving terminals are also sending their callable address in the RTCP SDES item *H323-CADDR* which allows the panel to invite them.

The panel could either have social or automatic control. Automatic control is provided through H.323 chair-control. The conference capabilities (**chairControlCapability**) of the MC provide information about the type of control used.

9.1 Building a roster for an RTP session

The RTCP SDES item *NAME* provides the real name and association (e.g. John Doe, XYZ Corporation) of the user. All H.332 terminals in the conference shall periodically send *NAME* in each of the RTP sessions that they belong to so that all terminals can, over a period of time, build up a roster of participants in the RTP session. In a 7 kbit/s audio RTP session consisting of 100 participants, it is expected that the roster will be updated within eight minutes from the time the change occurs; a change such as an RTP receiving terminal joining the conference. Due to this fact, the state of the roster shall not be considered as an accurate reflection of all the participants in the conference at any point in time. The algorithm in the RTP specification to calculate the RTCP transmission interval was designed to provide a fast response for small RTP sessions where, for example, identification of all participants is important, and still be able to automatically adapt to large RTP sessions. The RTCP transmission interval is expected to usefully scale to about two to five minutes.

All terminals shall send RTCP *BYE* before leaving an RTP session to allow terminals in the conference to update their roster immediately. The H.323 terminals in the panel shall, in addition, comply with the H.245 procedures by closing the logical channels for that RTP session.

9.2 Being invited to the panel

H.332 terminals shall use the SDES item H323-CADDR, as shown in Figure 3, to provide the complete callable address of the user. This will allow the H.323 panel to invite the RTP receiving terminal in the panel through the use of the H.323 protocol. The SDES item H323-CADDR is H.332

application specific and will not be implemented by non-H.332 terminals. H.332 terminals that can not provide a callable address, primarily due to being inside of a certain type of firewall, shall not use H323-CADDR. Terminals that do not provide H323-CADDR can not be invited to the panel.

The SDES item H323-CADDR consists of the following fields:

- H323-CADDR [8 bits]: Specifies the constant 9 as the SDES item for the Callable Address.
- Length [8 bits]: Specifies the length in bytes of the terminal's callable address field. If the length is equal to zero, then the CNAME shall have the valid callable address. The Null H323-CADDR is preferred since it saves bytes that can be used to send other SDES items more frequently.
- Terminal Callable Address [less than 256 bytes]: Specifies the complete callable address of the user. The form of the address depends on whether the user terminal has a firewall.
  - If the firewall is not present or if the firewall is transparent, then the user terminal shall have the form "user@terminal" or "terminal". The user name "user" is typically the login name (e.g. jdoe) rather than the personal name (e.g. John Doe). The terminal name "terminal" is either the fully qualified domain name of the host or the standard ASCII representation of the terminal numeric address (e.g. 134.134.157.81). For the multi-user terminal, the ***Setup*** message from the panel shall contain the user name in the "destinationAddress" field.
  - If the firewall is present then the form of the callable address is for further study.

Note that the format of the RTCP SDES packet is defined in the RTP standard. Refer to the standard for more details.

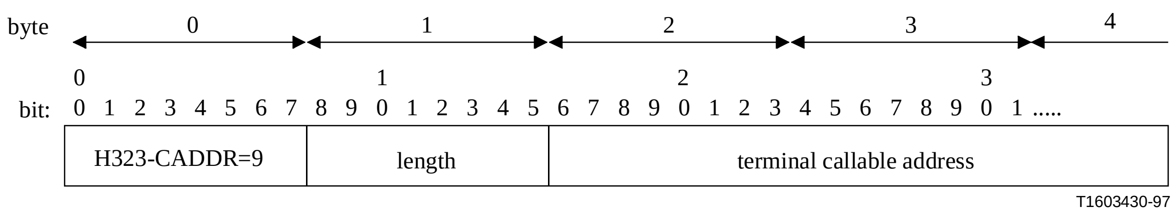

Figure 3/H.332 – Format of the SDES item H323-CADDR

9.3 Joining the panel

To join the panel, the RTP receiving terminals shall retrieve the address of the MC from the SDP announcement. It shall then follow the procedures of Recommendation H.323 to join the panel.

If the MC address is not advertised in the SDP announcement, then the RTP receiving terminals shall not be allowed to join the panel; they may only be invited to the panel.

Since the MC address is known to all the RTP receiving terminals, there is a possibility of a large number of RTP receiving terminals trying to join the panel and congesting the MC. To avoid this possibility, there may need to be an MC Congestion Control algorithm to prevent the MC from getting congested with connections from the RTP receiving terminals. This algorithm is currently a research topic and would require much experimentation before it could become part of this Recommendation. Therefore, this algorithm is for further study and will be considered for standardization in this Recommendation when it has been properly tested.

9.4 Leaving the panel

Once an RTP receiving terminal has joined the panel, it can stay until it decides to leave or until the MC terminates its H.323 call. The MC allows a limited number of connections based on the amount of resources allocated for the panel. When the resources have been exhausted and some RTP receiving terminal wants to join the panel, the MC may drop some temporary panel member. The decision on which temporary panel member to drop is based on some policy of the MC that is not part of this Recommendation. The dropped member shall leave the panel by following the procedures of Recommendation H.323. After being dropped, the temporary panel member becomes an RTP receiving terminal.

9.5 Periodicity of RTCP SDES items

Three RTCP SDES items are used: *CNAME*, *H323-CADDR*, and *NAME*. *CNAME* shall be sent every RTCP interval. One extra SDES item shall be sent at least every third interval. The extra items are *NAME*, and *H323-CADDR*. Both *NAME* and *H323-CADDR* shall be sent at least every sixth interval. Note that with the use of the Null H323-CADDR, the frequency of NAME can be increased.

9.6 Generating SSRC in RTP/RTCP

The RTP/RTCP standard mandates that all terminals, both sources and sinks of RTP streams, generate a globally unique SSRC. Even though in the H.332 conference there are a few sources in the panel and thousands of sinks of the RTP streams, each terminal – sink or source – still has to generate a unique SSRC within an RTP session which amounts to thousands of unique SSRCs within an RTP session. The rules specified in the RTP/RTCP standard shall be followed by all H.332 receiving terminals to generate a random SSRC.

The H.323 terminals in the panel generate SSRC by using the 8 bits of the terminal number as the low 8 bits of SSRC. This mapping guarantees that SSRC collisions will never happen in the panel. It also limits the number of globally unique SSRCs to 256 (actually 192 as mentioned in Recommendation H.323).

Due to the inclusion of RTP receiving terminals outside the panel, the H.323 terminals in the panel cannot assume that SSRC collisions will not happen. When an H.323 terminal detects a collision, it shall change the high 24 bits of the SSRC. The H.323 terminals can still use the low 8 bits of the SSRC in the RTP (not RTCP) packet as a terminal identifier because all sources are H.323 terminals in the panel.

10 Security

If the H.332 conference is secure, then the media streams must be encrypted. Since the encryption key(s) and algorithm(s) are distributed securely before the start of the conference, they should not be redistributed within the panel to prevent a potential disclosure of key(s) to unregistered interlopers.

11 Mandatory features of H.332 terminals

The H.332 terminals come in two flavours: H.332 receivers (RTP receiving terminals) and H.332 transmitters/receivers (H.323 terminals). It is possible to only implement an H.332 receiver.

An H.332 receiver shall comply with the mandatory requirements of RTP/RTCP, SDP and of Recommendation H.332.

An H.332 transmitter/receiver shall comply with the mandatory requirements of Recommendations H.323 and H.332.

The baseline codecs for this Recommendation shall be the same as defined in Recommendation H.323.

12 Bibliography

– CCITT Recommendation G.711 (1988), *Pulse Code Modulation (PCM) of voice frequencies*.

– CCITT Recommendation G.722 (1988), *7 kHz audio-coding within 64 kbit/s*.

– ITU-T Recommendation G.723.1 (1996), *Dual rate speech coder for multimedia communications transmitting at 5.3 and 6.3 kbit/s*.

– CCITT Recommendation G.728 (1992), *Coding of speech at 16 kbit/s using low-delay code excited linear prediction.*

– ITU-T Recommendation H.261 (1993), *Video codec for audiovisual services at p × 64 kbit/s*.

– ITU-T Recommendation H.263 (1998), *Video coding for low bit rate communication*.

– ITU-T Recommendation H.320 (1997), *Narrow-band visual telephone systems and terminal equipment*.

– ITU-T Recommendation H.321 (1998), *Adaptation of H.320 visual telephone terminals to B-ISDN environments*.

– ITU-T Recommendation H.322 (1996), *Visual telephone systems and terminal equipment for local area networks which provide a guaranteed quality of service.*

– ITU-T Recommendation H.324 (1998), *Terminal for low Bit-rate multimedia communication*.

– ITU-T Recommendation T.120 (1996), *Data protocols for multimedia conferencing*.

APPENDIX I

Examples of SDP announcements

Examples of announcements using this Recommendation are described here.

I.1 Example for an RTP/RTCP audio conference using social control

This example is about an audio conference which uses social control within RTP/RTCP. The announcement for this conference is as follows:

```
v=0
o=vkumar f81d4fae-7dec-11d0-a765-00a0c91e6bf6 1 IN IP4 134.134.157.81
s=Discussion on Environmental Pollution
t=<3034423619> 0
r=7d 1h 0

m=audio 5004 RTP/AVP 120
c=IN IP4 228.2.1.1/63
b=9.8
a=rtpmap:120 G729/8000
a=sendrecv
```

The explanation of the records is as follows:

– *Conference records*

v version 0 of SDP

o login name of the originator; universally unique ID of the conference; version of the announcement; Internet as the network type of the host; IP v4 as the address type of the host; IP address of the originator's host

s name of the conference

t start time is this Monday at 10 am; end time of zero implies that the conference never ends

r conference is held every week; duration is 1 hour; conference starts at an offset of 0 from the start time

– *Audio RTP session records*

m audio RTP session; RTP port of 5004, RTCP port of 5005 implied; Audio Video Profile is used; dynamic payload type is 120

c connection address includes Internet as the network type; IP v4 as the address type; a multicast address for connection, and a TTL of 63 implying that the conference is bounded within the region

b this RTP session uses a bandwidth of 10 kbit/s

a=rtpmap a dynamic payload type of 120 is used; details of the codec include G.729 with a clock rate of 8000 Hz

a=sendrecv can send and receive on this RTP session, i.e. this RTP session uses social control

I.2 Example for distance education

This example is about an announcement for Distance Learning that will be multicast in two languages. Clients are allowed to negotiate capabilities before the conference. During the conference the clients are allowed to join the panel using the MC address provided in the announcement.

Note that the announcement has the capability negotiation server and the MC as the same machine. If an MCU – instead of an MC – is used, then within the panel the media is centralized; outside the panel the media is multicast. Also, the MCU provides the mixed audio and the video of the focus to all the terminals. If an MC is used, then chair-control can be used to allow media from one terminal at a time.

I.2.1 Public announcement

The public announcement does not contain the MC address and the encryption keys. This extra information is part of the private announcement. The public announcement is shown below:

```
v=0
o=vkumar f81d4fae-7a13-11d0-a7bc-00a0c91e6bf6 1 IN IP4 134.134.157.81
s=CS 506: In-depth study of Video Conferencing Architecture, Hardware, and Software
u=http://www.university.edu/cs506
t=4921749382 8311296431
r=7d 1h 0 48h 48h
k=http://www.university.edu/cs/registration.html
a=type:H332
```

m=control 1720 H323 caps
c=IN IP4 134.134.157.81
m=audio 5004 RTP/AVP 4
i=Audio in English
c=IN IP4 224.2.1.1/127
b=5.6
a=recvonly

m=audio 5006 RTP/AVP 4
i=Audio in Hindi
c=IN IP4 224.60.1.1/127
b=5.6
a=recvonly

m=video 5008 RTP/AVP 100
c=IN IP4 224.40.1.1/127
b=16
a=rtpmap: 100 H263/90000
a=recvonly

The explanation of the records is as follows:

– *Conference records*

v	version 0 of SDP
o	login name of the originator; universally unique ID of the conference; version of announcement; Internet as the network type of the host; IP v4 as the address type of the host; IP address of the originator's host
s	name of the conference
u	URL is provided for details about the class
t	start time is September 8, 1997; end time is December 19, 1997
r	class is held every week; duration is 1 hour; class is held Monday; class is also held Wednesday; and class is also held Friday
k	URL is provided for the private announcement
a=type	H.332 protocol is used for conferencing

– *Control records for capability negotiation*

m	control; H.323 call-signalling well-known port; H.323 protocol, capability negotiation is allowed
c	connection address includes Internet as the network type; IP v4 as the address type; server's call-signalling address for connection

– *Audio RTP session records*

m	audio RTP session; RTP port of 5004, RTCP port of 5005; Audio Video Profile is used; static payload type for G.723.1
i	description to point out that this RTP session is in English
c	connection address includes Internet as the network type; IP v4 as the address type; multicast address for connection, and a TTL of 127 implies that the class is open to all in the world
b	this RTP session uses a bandwidth of 5.3 kbit/s
a= recvonly	can only receive on this RTP session when outside the panel; this is the default

– *Audio RTP session records*

Similar to the previous audio RTP session except that the audio is in a different language

– *Video RTP session records*

m	video RTP session; RTP port of 5008, RTCP port of 5009; Audio Video Profile is used; dynamic payload type for H.263
c	connection address includes Internet as the network type; IP v4 as the address type; multicast address for connection, and a TTL of 127 implies that the class is open to all in the world
b	this RTP session uses a bandwidth of 16 kbit/s
a=rtpmap	a dynamic payload type of 100 is used; details of the codec include H.263 with a clock rate of 90 000 Hz
a=recvonly	can only receive on this RTP session when outside the panel; this is the default

I.2.2 Private announcement

The private announcement differs from the public announcement in that it has the MC address. The encryption algorithm and the key as shown below:

```
v=0
o=vkumar f81d4fae-7a13-11d0-a7bc-00a0c91e6bf6 2 IN IP4 134.134.157.81
s=CS 506: In-depth study of Video Conferencing Architecture, Hardware, and Software
u=http://www.university.edu/cs506
t=4921749382 8311296431
r=7d 1h 0 48h 48h
k=base64: des:a1AB07392hqiHC7Td283==BA
a=type:H332

m=control 1720 H323 caps
c=IN IP4 134.134.157.81

m=control 1720 H323 mc
c=IN IP4 134.134.157.81

m=audio 5004 RTP/AVP 4
i=Audio in English
c=IN IP4 224.2.1.1/127
b=5.6
a=recvonly

m=audio 5006 RTP/AVP 4
i=Audio in Hindi
c=IN IP4 224.60.1.1/127
b=5.6
a=recvonly

m=video 5008 RTP/AVP 100
c=IN IP4 224.40.1.1/127
b=16
a=rtpmap: 100 H263/90000
a=recvonly
```

The explanation of the records is as follows:

– *Conference records*

v	version 0 of SDP
o	login name of the originator; universally unique ID of the conference; version of the announcement; Internet as the network type of the host; IP v4 as the address type of the host; IP address of the originator's host
s	name of the conference
u	URL is provided for details about the class
t	start time is September 8, 1997; end time is December 19, 1997
r	class is held every week; duration is 1 hour; class is held Monday; class is also held Wednesday; and class is also held Friday
k	encryption algorithm and key in base 64 encoding
a=type	H.332 protocol is used for conferencing

– *Control records for capability negotiation*

m	control; H.323 call-signalling well-known port; H.323 protocol, capability negotiation is allowed
c	connection address includes Internet as the network type; IP v4 as the address type; server's call-signalling address for connection before the connection

– *Control records for back-channel*

m	control; H.323 call-signalling well-known port; H.323 protocol, joining the panel is allowed
c	connection address includes Internet as the network type; IP v4 as the address type; MC's call-signalling address for connection during the conference

– *Audio RTP session records*

m	audio RTP session; RTP port of 5004, RTCP port of 5005; Audio Video Profile is used; static payload type for G.723.1
i	description to point out that this RTP session is in English
c	connection address includes Internet as the network type; IP v4 as the address type; multicast address for connection, and a TTL of 127 implies that the class is open to all in the world
b	this RTP session uses a bandwidth of 5.3 kbit/s
a=recvonly	can only receive on this RTP session when outside the panel, this is the default

– *Audio RTP session records*

Similar to the previous audio RTP session expect that the audio is in a different language

– *Video RTP session records*

m	video RTP session; RTP port of 5008, RTCP port of 5009; Audio Video Profile is used; dynamic payload type for H.263
c	connection address includes Internet as the network type; IP v4 as the address type; multicast address for connection, and a TTL of 127 implies that the class is open to all in the world
b	this RTP session uses a bandwidth of 16 kbit/s

a=rtpmap	a dynamic payload type of 100 is used; details of the codec include H.263 with a clock rate of 90 000 Hz
a=recvonly	can only receive on this RTP session when outside the panel, this is the default.

APPENDIX II

Usage of RSVP

The information regarding each media stream to make QoS reservation using RSVP is included in the SDP announcement. The announcement for each stream includes, among other information, the multicast address, the transport port, and the bandwidth. To start receiving a media stream and to reserve resources for that stream, the H.332 terminal should follow the steps described below:

1) Join the multicast group by sending an IGMP "REPORT" message;
2) Register with RSVP to get notified when messages arrive;
3) Wait for the first RSVP "PATH" message;
4) Start sending RSVP "RESV" messages.

To release the reserved resources for a media stream and stop receiving that stream, the H.332 terminal should follow the steps described below:

1) Release the reserved resources by sending an RSVP "RESVTEAR" message;
2) Leave the multicast group by sending an IGMP "LEAVE" message.

ITU-T RECOMMENDATIONS SERIES

Series A	Organization of the work of the ITU-T
Series B	Means of expression: definitions, symbols, classification
Series C	General telecommunication statistics
Series D	General tariff principles
Series E	Overall network operation, telephone service, service operation and human factors
Series F	Non-telephone telecommunication services
Series G	Transmission systems and media, digital systems and networks
Series H	**Audiovisual and multimedia systems**
Series I	Integrated services digital network
Series J	Transmission of television, sound programme and other multimedia signals
Series K	Protection against interference
Series L	Construction, installation and protection of cables and other elements of outside plant
Series M	TMN and network maintenance: international transmission systems, telephone circuits, telegraphy, facsimile and leased circuits
Series N	Maintenance: international sound programme and television transmission circuits
Series O	Specifications of measuring equipment
Series P	Telephone transmission quality, telephone installations, local line networks
Series Q	Switching and signalling
Series R	Telegraph transmission
Series S	Telegraph services terminal equipment
Series T	Terminals for telematic services
Series U	Telegraph switching
Series V	Data communication over the telephone network
Series X	Data networks and open system communications
Series Y	Global information infrastructure
Series Z	Programming languages