Telecommunications Standardisation sector
Study Group 15
Experts Group for Video Coding and Systems
in ATM and Other Network environments

Document AVC-710
Version 1
24-27 January, Tokyo

Source: BT
Title: Control syntax and procedures for ATM-based multimedia systems
Purpose: Proposal

This paper has been distributed to the SG15 AVC group by Norman Kenyon, BT Labs, including the covering note.

Dear Colleagues,

I am pleased to inform you that Mike Nilsson has taken over from me the task of editing H.24X. This is a Good Thing, as Mike is surely better qualified than I to write down words about protocols, and he has also been involved in the broadband experts group. I will just look over his shoulder, to help try to avoid the pitfalls we encountered in H.221/242/230 - many of which were consolidated into immutable standards before we realised the problems they would cause.

I attach Mike's redraft (he has gone off to have Christmas now), and you will find it considerably changed. Instead of simple lists of codepoints we now have the beginnings of an ASN.1 description - I am rather hoping it would be legitimate to add a "comments" column to it so that the uninitiated can follow what is going on in there! Mike has also removed (at the behest of the last meeting of Mr Okubo's group, I believe) most of the non-capability C&I which, in the narrowband equivalent Recs., were housed in Table A.1/H.221 and in H.230. **I personally believe this to be unwise**, making the following points.

1. There are several functions that are not capabilities, nor so intimately related to an information stream (I mean Audio, Video, or Data) that they can be embedded therein: loopbacks (LCA, LCV, LCD, LCO, random numbers for digital loop detection); encryption management messages and IVs; numbers and alphanumerics - useful for a variety of standard and non-standard purposes; Freeze-picture and Fast-Update (VCF, VCU); non-standard messages; and there will undoubtedly be others.

2. It will be confusing if there is more than one "C&I" channel involved in the communication (and we will have long arguments as which items are Capabilities and which are not!). It is true that the T.120 protocol could carry these functions (including the Caps, actually), but we probably will not stipulate that T.120 shall always be present and that it shall be activated before all else in the call. It is true that a "CS" protocol (DSM-CC-ext, perhaps?) will be needed for dialogue between a client and a server, and probably we would put that in a "data" path - indeed we would do the same to a T.120 stream; I doubt if we would want both simultaneously. Thus I would say a reasonable communication includes:

Audio stream(s) + Video stream(s) + {either T.120 or CS protocol but not both} + other Data stream(s) + C&I

though of course one or more of these could be omitted, and the transmit and receive directions at an endpoint would not need to have the same mix (communication can even be quite unidirectional). Even with only one C&I stream we will have to take great care over the relationship between the C&I and T.120/CS protocol if present.

3. Even if it is decided that capabilities and "other" C&I must have two distinct channels, I would advocate at least dealing with them both in H.24X. It does not seem to me to be right to use H.32X for what would essentially be a toolbox of C&I functions which could be used or omitted in the great variety of multimedia terminals which can be foreseen.

The purpose of H.320 was to weld together the "infrastructure" Recs. H.221/242/230, H.261, G.711/722/728, to show how to specify a basic videoconference terminal or a videotelephone. H.32X should arguably have a similar aim, and should not introduce its own syntax etc.

4. For a while now we have recognised that real communications patterns will not adhere rigidly to the boundaries between "conversational", "storage/retrieval", "messaging" and other service classes. Nevertheless, it may be quite difficult to develop H.32X to cover all multimedia apps., including non-conversational as well as conversational. This needs more thought, particularly with regard to the call setup aspects. My suggestion is that we reach agreement on the **technical solution for call setup and other "communication phases"**, covering all multimedia, before we finalise how much of this should go into H.32X, how much into Q.2939, and so on. I will try to expand on the Annex to my previous H.24X draft in order to provide a discussion document on this.

Norman Kenyon

INTERNATIONAL TELECOMMUNICATION UNION

ITU-T

DRAFT H.24x

TELECOMMUNICATION
STANDARDIZATION SECTOR
OF ITU

(03/01/95)

LINE TRANSMISSION OF NON-TELEPHONE SIGNALS

MULTIMEDIA SYSTEM CONTROL

DRAFT ITU-T Recommendation H.24x

FOREWORD

The ITU Telecommunication Standardization Sector (ITU-T) is a permanent organ of the International Telecommunication Union. 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, established the topics for study by the ITU-T Study Groups which, in their turn, produce Recommendations on these topics.

ITU-T Recommendation H.24x was prepared by the ITU-T Study Group 15 (199x-199x) and was approved by the WTSC (Place, Month xx-xx, 199x).

NOTES

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

© ITU 199x

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

MULTIMEDIA SYSTEM CONTROL

(Place, 199x)

{Ed. Text within curly brackets beginning "Ed" is not part of the Recommendation, but serves to hold notes, questions etc by the editor.}

1 Summary

{Ed. This section is included to provide background to this draft: why it is the way it is. It is assumed that this section will be removed when the document matures.}

In-band procedures and protocol are needed for both ATM-based multimedia systems and for PSTN videophone/multimedia; there is no obvious reason why these two digital systems, despite their bit-rate difference, should not have essentially the same scheme, at least at the "effective" layer - though this may sit on different transport protocols.

This paper makes a proposal, for further consideration. It draws on the experience of H.242, but seeks to avoid the deficiencies thereof: there are transmit-capabilities as well as receive-caps (H.242 had only the latter); multiple audio and video streams and other data can be declared, and the capsets are shorter. There are no commands (on the assumption that audio, video and data have their own headers), and one end-point can control what is sent to it by means of "mode requests".

2 Scope

This recommendation covers in-band procedures and protocol, including the capability exchanges and mode-setting, for ATM-based multimedia systems. It covers a wide range of applications, including storage/retrieval, messaging and "distribution" services as well as conversational. There is independence of the two directions of transmission.

Transmission is in a packet-based logical channel carrying C&I (control and indications) that provides a highly reliable transmission service.

Note: there should be no confusion with the T.120 management system, which is carried within the data stream, and covers different functionalities from those described here - the C&I stream and the T.120-data stream are complementary.

3 References

[1] Reference to ASN.1

{Ed. Necessary references to be added.}

4 Definitions

{Ed. Necessary definitions to be added.}

5 General

The procedures in this Recommendation are intended to ensure that only those signals are transmitted that can be received and appropriately treated by the remote terminal without ambiguity. This requires that the capabilities of each terminal to receive and decode be known to the other terminal. It is not necessary that a terminal understand or store all incoming capabilities: those that are not understood, or can not be used (because the terminal has no means to transmit corresponding information) shall be ignored, and no fault shall be considered to have occurred.

The total capability of a terminal to receive and decode various signals is made known to the other terminal by transmission of its capability set.

All control and indication signals defined in this Recommendation are specified using the notation defined in ASN.1 [1]. It is assumed that a lower layer provides a highly reliable transmission service. This Recommendation does not specify error correction or ACK/NACK procedures.

{Ed. This was proposed in AVC-699, where X.224 was suggested}

There are separate capability messages for ability to receive and transmit: receiveCapSet and transmitCapSet.

These capability sets provide for more than one stream of a given medium type to be sent "simultaneously" - of course not at exactly the same time, but multiplexed within H.222.1 sub channels. For example, a terminal may declare its ability to receive (or send) two independent H.262 video streams and two independent G.722 audio streams at the same time. The capability sets also provide for terminals that are software-oriented, and have capabilities that are interdependent; for example, they might be able to decode pictures faster (lower MPI) when the audio algorithm is less computation-intensive (G.722) and vice versa.

Capabilities are indicated by a sequence of TermCapSet records. Each record is indicated as being an independent record or a dependent record. Independent capability records simply indicate the data types that can be processed (G.711 audio, G.728 audio, QCIF H.261 video etc.), without any implication of simultaneous capability, except that one of each data type (video, audio, etc.) can be processed at the same time. Dependent capability records explicitly state the simultaneous capabilities.

Means are provided to control the mode transmitted by the far-end terminal whenever this has not been explicitly prevented by that terminal. Messages are defined to allow the transfer of data when T.120 capability is absent, and for encryption and key management.

This Recommendation does not specify means for transmitting the current mode of operation, frame synchronous C&I and keyboard/user input.

6 Messages: syntax

This section specifies the syntax of messages using the notation defined in ASN.1. Messages shall be encoded according to ASN.1 packed encoding rules [1].

6.1 Transmit and Receive capabilities

```
DeclaredTermCapSet          ::=SEQUENCE
{
    independentTermCapSet   SEQUENCE OF TermCapSet OPTIONAL,
    dependentTermCapSet     SEQUENCE OF TermCapSet OPTIONAL,
    symmetricalCapSets      BOOLEAN OPTIONAL,
    ...
}

TermCapSet                  ::=SET
{
    receiveCapSet           CapSet OPTIONAL,
    transmitCapSet          CapSet OPTIONAL,
    ...
}

CapSet                      ::=SET
{
    videoCap                SET OF VideoCap,
    audioCap                SET OF AudioCap,
    networkAdaptCap         SET OF NetworkAdaptCap,
    dataCap                 SET OF DataCap,
    ...
}

VideoCap                    ::=CHOICE
{
    h261VideoCap            H261VideoCap,
    h262VideoCap            H262VideoCap,
    h26pVideoCap            H26pVideoCap,
    ...
}

H26pVideoCap                   ::=SEQUENCE
{
    ...
}

H261VideoCap                   ::=SEQUENCE
{
    qcifMPI                    INTEGER (1..4),
    cifCapability              BOOLEAN,
    cifMPI                     INTEGER (1..4) OPTIONAL,
    ...
}

H262VideoCap                   ::=SEQUENCE
{
    profileAndLevel            ProfileAndLevel,
    videoBitRate               INTEGER OPTIONAL,
    vbvBufferSize              INTEGER OPTIONAL,
    ...
}
```

{Ed. Bit Rate and VBV Buffer Size included particularly for levels above Main Level where it is expected that equipment will be built between profile and level points. Concerning compatibility with H.261, is spatial scalability without the need for high resolution capability a useful thing to be able to signal? if so, how?}

```
ProfileAndLevel                ::=SEQUENCE
{
    SP@ML                      BOOLEAN,
    MP@LL                      BOOLEAN,
    MP@ML                      BOOLEAN,
    MP@H-14                    BOOLEAN,
    MP@HL                      BOOLEAN,
    SNR@LL                     BOOLEAN,
    SNR@ML                     BOOLEAN,
    Spat@H-14                  BOOLEAN,
    HP@ML                      BOOLEAN,
    HP@H-14                    BOOLEAN,
    HP@HL                      BOOLEAN,
    ...
}

AudioCap                       ::=SEQUENCE
{
    g711Alaw                   BOOLEAN,
    g711Ulaw                   BOOLEAN,
    g722                       BOOLEAN,
    g728                       BOOLEAN,
    av25y                      Av25y OPTIONAL,
    mpegAudio                  MpegAudio OPTIONAL
    ...
}

MpegAudio                      ::=SEQUENCE
{
    audioLayer                 AudioLayer,
    audioSampling              AudioSampling,
    asynchronousCap            BOOLEAN,
    AudioCorrectionModes       AudioCorrectionModes,
    bitRate                    INTEGER,
    ...
}

AudioLayer                     ::=ENUMERATED
```

```
{
    layer1                      (0),
    layers1&2                   (1),
    layers1&2&3                 (2),
    ...
}

AudioSampling                   ::=SEQUENCE
{
    32k                         BOOLEAN,
    44k1                        BOOLEAN,
    48k                         BOOLEAN,
    ...
}

AudioCorrectionModes            ::=ENUMERATED
{
    mode1                       (0),
    mode2                       (1),
    mode3                       (2),
    allThreeModes               (3),
    ...
}

Av25y                           ::=SEQUENCE
{
    ...
}

NetworkAdaptCap                 ::=SEQUENCE
{
    aal                         Aal,
    h222Multiplex               H222Multiplex,
    bitRate                     INTEGER,
    numberOfVCs                 INTEGER,
    ...
}

Aal                             ::=SEQUENCE
{
    Aal1                        BOOLEAN,
    Aal5                        BOOLEAN,
    ...
}

H222Multiplex                   ::=SEQUENCE
{
    transportStream             BOOLEAN,
    programStream               BOOLEAN,
    ...
}

DataCap                         ::=SEQUENCE
{
    t120                        BOOLEAN,
    lapm                        BOOLEAN,
    ...
}
```

6.2 Request Transmit and Receive capabilities

```
RequestCapability               ::=SEQUENCE
{
    requestTxCap                BOOLEAN,
    requestRxCap                BOOLEAN,
        cause                   To be studied,
        ...
    }
```

6.3 Request Mode (Remote control)

```
    RequestMode                 ::=SEQUENCE
    {
        requestedMode           TermCapSet,
        ...
    }

    RequestModeFailure          ::=SEQUENCE
    {
        cause                   To be studied,
        ...
    }
```

6.4 Downloadable software

```
    DownloadableSoftware        ::=ENUMERATED
    {
        requestRxSoftware       (0),
        readyForSoftware        (1),
        nackRxSoftware          (2),
        ...
    }
```

6.5 Encryption

```
    Encryption                  ::=CHOICE
    {
        encryptionSE            EncryptionSE,
        encryptionIV            EncryptionIV,
        ...
    }

    EncryptionSE                ::=SEQUENCE
    {
        messageIdentifier       INTEGER (0..255),
        content                 SEQUENCE INTEGER (0..255),
        ...
    }

    EncryptionIV                ::=SEQUENCE
    {
        initializationVector    SEQUENCE INTEGER (0..255),
        ...
    }
```

{Ed. Mode indication, frame synchronous commands, like freeze picture, and various data transfers, like alphanumeric, have been omitted as they will be defined elsewhere. For example, mode indication will be done using PSI/PSM defined in H.222.0. Additional descriptors may need to be defined.}

6.6 Change or end session

```
    ChangeOrEndSession          ::=SEQUENCE
    {
        command                 ENUMERATED
        {
            changeSession       (0),
            endSession          (1),
            ...
        },
        ...
    }
```

7 Messages: semantic definitions

This section provides semantic definitions and constraints on the syntax elements defined in the previous section.

7.1 Transmit and Receive capabilities

DeclaredTermCapSet: a sequence of capabilities. The semantic definition of independent and dependent capabilities are different. A DeclaredTermCapSet shall include at least one TermCapSet. It is possible to include a mix of independent and dependent TermCapSets.

Independent capabilities: the indication of more than one capability of the same type (video, audio, adaptation or data), such as H.261 and H.262, does not indicate the ability to perform multiple simultaneous processing of the same data type, but merely the independent capabilities. A terminal shall at least be capable of the simultaneous processing of any one indicated independent capability of each type.

Dependent capabilities: this explicitly states the simultaneous capabilities. A terminal shall be capable of the simultaneous processing of all the dependent capabilities indicated in a single TermCapSet.

Note: terminals that can not vary the allocation of resources can indicate their capability completely by use of a single independent TermCapSet.

{Ed. Example for clarification of intentions: A terminal that can process G.711 and QCIF, G.728 and QCIF, or G.711 and CIF simultaneously, but not G.728 and CIF simultaneous could indicate any of the following.

independent TermCapSet 1	:	G.711, G.728, QCIF
independent TermCapSet 2	:	G.711, QCIF, CIF
dependent TermCapSet 1 :	G.711, CIF	
dependent TermCapSet 2 :	G.728, QCIF	

Note: the inclusion of, say, G.711 and G.728 in a dependent capability would indicate that both can be performed simultaneously.}

symmetricalCapSets indicates whether the transmit capabilities are the same as the receive capabilities; it shall not be present when any transmit capabilities are explicitly indicated.

TermCapSet: a set of receive and transmit capabilities. In general, both are optional, but at least one shall be present.

CapSet: a set of unidirectional capabilities. At least one of each of VideoCap, AudioCap, NetworkAdaptCap and DataCap shall be present.

VideoCap: indicates video capabilities.

H261VideoCap: indicates H.261 capabilities. qcifMPI indicates the minimum picture interval in units of 1/29.97 for the processing of QCIF pictures and shall take one of the values 1, 2, 3, 4. cifCapability indicates capability of processing CIF resolution pictures; it does not indicate the capability of simultaneous QCIF and CIF capability. Simultaneous processing capability can be indicated by multiple instances of VideoCap in a dependent TermCapSet. cifMPI indicates the minimum picture interval in units of 1/29.97 for the processing of QCIF pictures and shall take one of the values 1, 2, 3, 4. cifMPI shall be present only when CIF capability is indicated.

H262VideoCap, ProfileAndLevel: indicates H.262 capabilities. The indication of more than a single capability within a single H262VideoCap hierarchy does not indicate simultaneous processing capability. Simultaneous processing capability can be indicated by multiple instances of VideoCap in a dependent TermCapSet.

AudioCap: indicates audio capabilities.

Av25y: indicates capability for processing of Av25y coded audio.

MpegAudio, AudioLayer, AudioSampling, AudioCorrectionModes: indicates the specifics of MPEG audio capabilities, when present. The indication of more than a single capability within a single MPEGAudio hierarchy does not indicate simultaneous processing capability. Simultaneous processing capability can be indicated by multiple instances of AudioCap in a dependent TermCapSet.

NetworkAdaptCap, Aal, H222Multiplex: indicates network adaptation capabilities.

DataCap: indicates data capabilities.

7.2 Request Transmit and Receive capabilities

RequestCapability: this message requests the far end terminal to transmit its transmit and/or receive capabilities. The inclusion of a field to indicate the cause of the request is for future study.

7.3 Request Mode (Remote control)

RequestMode: this message requests the far end terminal to change its transmission mode to the mode indicated. requestedMode is a dependent TermCapSet that is consistent with the most recently received DeclaredTermCapSet.

RequestModeFailure: this message indicates that the terminal was unable to change its transmission mode to the requested mode. The definition of the field to indicate the cause of the failure is for future study.

7.4 Downloadable Software

DownloadableSoftware: this message allows for software transfers in case T.120 is not in use. The facility is intended for server-to-terminal purposes, but could be applicable also for terminal-to-terminal cases. The transfer is managed by use of requestRxSoftware, readyForSoftware and nackRxSoftware.

{Ed. How is this to be used?}

7.5 Encryption

Encryption: this message concerns encryption. messageIdentifier is defined in H.233 and H.234; content is defined in H.233/4.

{Ed. More needs to be written here, but I don't know what!}

7.6 Change or end session

ChangeOrEndSession: this message indicates a change of session or the end of a session.

{Ed. What happens on change of session?}

8 In-band procedures

This section defines in-band procedures.

8.1 Point-to-point: conversational services

The following sequence takes place during a point-to-point session.

1. Prior to receipt of the first DeclaredTermCapSet containing receiveCapSet from the remote end-point, transmission is on the basis of call class as determined in call set-up procedures [Ref xxx].

2. For conversational and speech services, both ends shall send DeclaredTermCapSet messages (one or more receiveCapSets and zero or more transmitCapSets).

{Ed. Acknowledgements removed on assumption of reliable transmission.}

3. Action related to any capability value received may not be taken until the complete message is received and validated. Such action may be to transmit a corresponding Mode, or send a request for the remote end-point to do so.

{Ed. The paragraphs here relating to mode indication and associate information have been deleted. In H.32x, mode indication and associative information will be provided by PSI/PSM defined by MPEG, but possibly extended by ITU-T.}

4. All terminals intended for use in connection-oriented applications shall be able to identify a receiveCapSet and its structure, and such capability values therein as are mandatory for those applications; any unrecognised capability values shall be ignored, and no fault shall be implied.

5. At any time after receipt of a DeclaredTermCapSet containing one or more receiveCapSets from the remote end-point, a mode switch may be made - other/changed media streams may be transmitted, provided that they are processable by the remote end-point, as notified by the receiveCapSets within the latest received DeclaredTermCapSet.

Note. From receipt of the first DeclaredTermCapSet containing receiveCapSet until receipt of endSession, all transmitted signals shall be in the range indicated by the receiveCapSets in the latest received DeclaredTermCapSet and by the transmitCapSets, if any, in the latest transmitted DeclaredTermCapSet - that is, valid capsets always replace the previous ones. An end-point shall store the latest received capsets as the "currently valid capsets".

{Ed. Is the last sentence needed? this looks untestable.}

6. If the currently valid received DeclaredTermCapSet contains one or more transmitCapSets, and if the currently valid received DeclaredTermCapSet contains one or more receiveCapSets, an endpoint may select a mode which it prefers to have transmitted to it, by sending a RequestMode message. An end-point whose latest transmission of DeclaredTermCapSet contains one or more transmitCapSets and which is in receipt of RequestMode shall comply; if for any reason it is unable to do so, it shall send RequestModeFailure.

A terminal that has not indicated transmit capabilities, or one whose most recently sent DeclaredTermCapSet did not indicate transmit capability, shall not receive RequestMode, that is, when transmit capabilities are not sent, the terminal does not wish to, and shall not, be remotely controlled.

Note. For the purpose of this paragraph, a transmitCapSet indicated by symmetricalCapSets shall be considered exactly as if it had been explicitly sent.

7. Change of capabilities: a changed DeclaredTermCapSet may be sent at any time. An unchanged capset shall not be sent unless:

incoming signals are not processable, that is, they lie outside the range of the transmitted receiveCapSets;

a RequestCapability message has been received; or

a fault has occurred (for further study).

8. A RequestCapability message may be sent at any time to elicit these details from the remote end-point. This may contain a "cause" field - for future study.

{Ed. For example, the cause field could indicate that the request follows receipt of an errored capset; other causes may be identified...}

A RequestCapability message may be sent if necessary, following an interruption or other cause for uncertainty; such messages shall not be sent repetitively or otherwise without strong cause.

9. If it is desired to change from conversational to another class, a ChangeSession message shall be sent. Subsequent action shall conform to something.

{Ed. Need to define what happens here.}

10. Prior to clearing down, an end-point shall transmit EndSession, and thereafter transmit no in-band messages, and ignore any received.

11. Downloading software - for further study.

{Ed. Await DAVIC results??}

{Ed. Is this still considered 'beyond the wit of Man', and if so, should it be defined?]

12. Encryption and Key Management - for further study.

{Ed. Same as for H.320 terminals??}

8.2 Point-to-point: connection-oriented non-conversational services

Storage/retrieval/messaging/distribution classes of service: for further study.

{Ed. Should there be provision for a text welcome while higher things are being negotiated?? The presumption here is that, just as speech is always available as starting-point for conversational applications, text should always be available as a default starting point for the other classes; an exception is that a fax-only device will go straight into its preamble.}

8.3 Point-to-point: connectionless non-conversational services

Storage/retrieval/messaging/distribution classes of service: for further study??.

{Ed. There seems to be little purpose in having this section within the inband signalling (capability exchange) part of the document as there is only unidirectional communication.}

8.4 Point-to-Multipoint working

Where one source is feeding several receivers, it will be unable (and does not try) to respond to any upstream signals, whether of the ReceiveCapSet type or requests to transmit in a particular medium/mode, RequestTxCap, etc. The source shall transmit its TransmitCapSets periodically (eg less than T1 intervals), including only the values actually in use in that session, together with RequestModeFailure and DeclaredTermCapSet containing no receiveCapSet. The TransmitCapSet shall not contain any options (NB: this does not mean that it cannot send two or more video signals in different formats, for example).

{Ed. RequestModeFailure should only be sent in case of receipt of RequestMode.}

{Ed. What does TransmitCapSet shall not contain any options mean?}

A receiving terminal shall recognise the DeclaredTermCapSet containing no receiveCapSet as indicative of a connectionless communication, and shall transmit no further messages. It shall, however, remain sensitive to incoming receiveCapSets which could indicate a change of orientation.

{Ed. All call control text deleted as outside the scope of this Recommendation.}

END