Telecommunication Standardization Sector (TSS)

AVC-893v1
Original: English

SG15/WP1

January 16-24, 1996

Question 2/15

STUDY GROUP 15 CONTRIBUTION

Source: D.Skran, Editor H.225.0

Title: Annex B to H.225.0 (RTP Profile)

Date: January 12, 1996

Summary: This document contains Annex B to H.225.0, which is the text of the RTP profile as produced by the IETF. It is being included here by order of the SG15 Plenary on the grounds that a reference to an RFC cannot be normative. Thanks are extended by the editor to Bob Webber of PictureTel for placing this material into windows format, and to the authors for their kind assistance, especially Steve Casner.

Notes on reading: ***{Editors notes are generally in braces}*** while underlined text is new. ~~The deleted text has strikethrough font.~~ Please ignore all references to other sections both inside and outside this document; these will be updated in the final version. All changes are relative to the basic text as taken from the IETF document dated Nov. 20, 1995.

Annex B: RTP Profile

See the introduction to Annex A; all the warnings mentioned there apply to this Annex as well. An informative reference to the full IETF document can be found in Appendix B; however, this Annex contains all information needed for the implmentation of H.323.

~~Internet Engineering Task Force Audio-Video Transport Working Group draft-ietf-avt-profile-06.txt GMD Fokus November 20, 1995 Expires: 4/1/96~~

~~RTP Profile for Audio and Video Conferences with Minimal Control~~

~~STATUS OF THIS MEMO~~

~~This document is an Internet-Draft. Internet-Drafts are working documents of the Internet Engineering Task Force (IETF), its areas, and its working groups. Note that other groups may also distribute working documents as Internet-Drafts.~~

~~Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as ``work in progress''.~~

~~To learn the current status of any Internet-Draft, please check the ``1id-abstracts.txt'' listing contained in the Internet-Drafts Shadow Directories on ftp.is.co.za (Africa), nic.nordu.net (Europe), munnari.oz.au (Pacific Rim), ds.internic.net (US East Coast), or ftp.isi.edu (US West Coast).~~

~~Distribution of this document is unlimited.~~

~~ABSTRACT~~

~~This memo describes a profile for the use of the real-time transport protocol (RTP), version 2, and the associated control protocol, RTCP, within audio and video multiparticipant conferences with minimal control. It provides interpretations of generic fields within the RTP specification suitable for audio and video conferences. In particular, this document defines a set of default mappings from payload type numbers to encodings.~~

~~The document also describes how audio and video data may be carried within RTP. It defines a set of standard encodings and their names when used within RTP. However, the encoding definitions are independent of the particular transport mechanism used. The descriptions provide pointers to reference implementations and the detailed standards. This document is meant as an aid for implementors of audio, video and other real-time multimedia applications.~~

1 Introduction

This profile defines aspects of RTP left unspecified in ~~the~~ Annex A~~RTP Version 2 protocol definition (RFC TBD)~~. This profile is intended for the use within audio and video conferences with minimal session control. In particular, no support for the negotiation of parameters or membership control is provided. The profile is expected to be useful in sessions where no negotiation or membership control are used (e.g., using the static payload types and the membership indications provided by RTCP), but this profile may also be useful in conjunction with a higher-level control protocol.

Use of this profile occurs by use of the appropriate applications; there is no explicit indication by port number, protocol identifier or the like.

Other profiles may make different choices for the items specified here.

2 RTP and RTCP Packet Forms and Protocol Behavior

The section "RTP Profiles and Payload Format Specification" enumerates a number of items that can be specified or modified in a profile. This section addresses these items. Generally, this profile follows the default and/or recommended aspects of the RTP specification.

RTP data header: The standard format of the fixed RTP data header is used (one marker bit).

Payload types: Static payload types are defined in Section 6.

RTP data header additions: No additional fixed fields are appended to the RTP data header.

RTP data header extensions: No RTP header extensions are defined, but applications operating under this profile may use such extensions. Thus, applications should not assume that the RTP header X bit is always zero and should be prepared to ignore the header extension. If a header extension is defined in the future, that definition must specify the contents of the first 16 bits in such a way that multiple different extensions can be identified.

RTCP packet types: No additional RTCP packet types are defined by this profile specification.

RTCP report interval: The suggested constants are to be used for the RTCP report interval calculation.

SR/RR extension: No extension section is defined for the RTCP SR or RR packet.

SDES use: Applications may use any of the SDES items described. While CNAME information is sent every reporting interval, other items should be sent only every fifth reporting interval.

Security: The RTP default security services are also the default under this profile.

String-to-key mapping: See Appendix B for this informative information.~~A user-provided string ("pass phrase") is hashed with the MD5 algorithm to a 16-octet digest. An n-bit key is extracted from the digest by taking the first n bits from the digest. If several keys are needed with a total length of 128 bits or less (as for triple DES), they are extracted in order from that digest. The octet ordering is specified in RFC 1423, Section 2.2. (Note that some DES implementations require that the 56-bit key be expanded into 8 octets by inserting an odd parity bit in the most significant bit of the octet to go with each 7 bits of the key.)~~

~~It is suggested that pass phrases are restricted to ASCII letters, digits, the hyphen, and white space to reduce the the chance of transcription errors when conveying keys by phone, fax, telex or email.~~

~~The pass phrase may be preceded by a specification of the encryption algorithm. Any characters up to the first slash (ASCII 0x2f) are taken as the name of the encryption algorithm. The encryption format specifiers should be drawn from RFC 1423 or any additional identifiers registered with IANA. If no slash is present, DES-CBC is assumed as default. The encryption algorithm specifier is case sensitive.~~

~~The pass phrase typed by the user is transformed to a canonical form before applying the hash algorithm. For that purpose, we define return, tab, or vertical tab as well as all characters contained in the Unicode space characters table. The transformation consists of the following steps: (1) convert the input string to the ISO 10646 character set, using the UTF-8 encoding as specified in Annex P to ISO/IEC 10646-1:1993 (ASCII characters require no mapping, but ISO 8859-1 characters do); (2) remove leading and trailing white space characters; (3) replace one or more contiguous white space characters by a single space (ASCII or UTF-8 0x20); (4) convert all letters to lower case and replace sequences of characters and non-spacing accents with a single character, where possible. A minimum length of 16 key characters (after applying the transformation) should be enforced by the application, while applications must allow up to 256 characters of input.~~

Underlying protocol: The profile specifies the use of RTP over unicast and multicast UDP. (This does not preclude the use of these definitions when RTP is carried by other lower-layer protocols.)

Transport mapping: The standard mapping of RTP and RTCP to transport-level addresses is used.

Encapsulation: No encapsulation of RTP packets is specified.

3 ~~Registering~~ Payload Types

See Appendix B for information on registering new payload types.

~~This profile defines a set of standard encodings and their payload types when used within RTP. Other encodings and their payload types are to be registered with the Internet Assigned Numbers Authority (IANA). When registering a new encoding/payload type, the following information should be provided:~~

~~o name and description of encoding, in particular the RTP timestamp clock rate; the names defined here are 3 or 4 characters long to allow a compact representation if needed;~~

~~o indication of who has change control over the encoding (for example, ISO, CCITT/ITU, other international standardization bodies, a consortium or a particular company or group of companies);~~

~~o any operating parameters or profiles;~~

~~o a reference to a further description, if available, for example (in order of preference) an RFC, a published paper, a patent filing, a technical report, documented source code or a computer manual;~~

~~o for proprietary encodings, contact information (postal and email address);~~

~~o the payload type value for this profile, if necessary (see below).~~

Note that not all encodings to be used by RTP need to be assigned a static payload type. Non-RTP means beyond the scope of this memo (such as directory services or invitation protocols) may be used to establish a dynamic mapping between a payload type drawn from the range 96-127 and an encoding. For implementor convenience, this profile contains descriptions of encodings which do not currently have a static payload type assigned to them.

The available payload type space is relatively small. Thus, new static payload types are assigned only if the following conditions are met:

o The encoding is of interest to the Internet community at large.

o It offers benefits compared to existing encodings and/or is required for interoperation with existing, widely deployed conferencing or multimedia systems.

o The description is sufficient to build a decoder.

4 Audio

4.1 Encoding-Independent Recommendations

For applications which send no packets during silence, the first packet of a talkspurt (first packet after a silence period) is distinguished by setting the marker bit in the RTP data header. Applications without silence suppression set the bit to zero.

The RTP clock rate used for generating the RTP timestamp is independent of the number of channels and the encoding; it equals the number of sampling periods per second. For N-channel encodings, each sampling period (say, 1/8000 of a second) generates N samples. (This terminology is standard, but somewhat confusing, as the total number of samples generated per second is then the sampling rate times the channel count.)

If multiple audio channels are used, channels are numbered left-to- right, starting at one. In RTP audio packets, information from lower-numbered channels precedes that from higher-numbered channels. For more than two channels, the convention followed by the AIFF-C audio interchange format should be followed [1], *{Editor's Note: we can't refer to an Apple document; this must be changed somehow}* using the following notation:

```
l    left
r    right
c    center
S    surround
F    front
R    rear
```

channels	description	channel 1	2	3	4	5	6
2	stereo	l	r				
3		l	r	c			
4	quadrophonic	Fl	Fr	Rl	Rr		
4		l	c	r	S		
5		Fl	Fr	Fc	Sl	Sr	
6		l	lc	c	r	rc	S

Samples for all channels belonging to a single sampling instant must be within the same packet. The interleaving of samples from different channels depends on the encoding. General guidelines are given in Section 4.2 and 4.3.

The sampling frequency should be drawn from the set: 8000, 11025, 16000, 22050, 24000, 32000, 44100 and 48000 Hz. (The Apple Macintosh computers have native sample rates of 22254.54 and 11127.27, which can be converted to 22050 and 11025 with acceptable quality by dropping 4 or 2 samples in a 20 ms frame.) However, most audio encodings are defined for a more restricted set of sampling frequencies. Receivers should be prepared to accept multi-channel audio, but may choose to only play a single channel.

The following recommendations are default operating parameters. Applications should be prepared to handle other values. The ranges given are meant to give guidance to application writers, allowing a set of applications conforming to these guidelines to interoperate without additional negotiation. These guidelines are not intended to restrict operating parameters for applications that can negotiate a set of interoperable parameters, e.g., through a conference control protocol.

For packetized audio, the default packetization interval should have a duration of 20 ms, unless otherwise noted when describing the encoding. The packetization interval determines the minimum end-to- end delay; longer packets introduce less header overhead but higher delay and make packet loss more noticeable. For non-interactive applications such as lectures or links with severe bandwidth constraints, a higher packetization delay may be appropriate. A receiver should accept packets representing between 0 and 200 ms of audio data. This restriction allows reasonable buffer sizing for the receiver.

4.2 Guidelines for Sample-Based Audio Encodings

In sample-based encodings, each audio sample is represented by a fixed number of bits. Within the compressed audio data, codes for individual samples may span octet boundaries. An RTP audio packet may contain any number of audio samples, subject to the constraint that the number of bits per sample times the number of samples per packet yields an integral octet count. Fractional encodings produce less than one octet per sample.

The duration of an audio packet is determined by the number of samples in the packet.

For sample-based encodings producing one or more octets per sample, samples from different channels sampled at the same sampling instant are packed in consecutive octets. For example, for a two-channel encoding, the octet sequence is (left channel, first sample), (right channel, first sample), (left channel, second sample), (right channel, second sample), ... For multi-octet encodings, octets are transmitted in network byte order (i.e., most significant octet first).

The packing of sample-based encodings producing less than one octet per sample is encoding-specific.

4.3 Guidelines for Frame-Based Audio Encodings

Frame-based encodings encode a fixed-length block of audio into another block of compressed data, typically also of fixed length. For frame-based encodings, the sender may choose to combine several such frames into a single message. The receiver can tell the number of frames contained in a message since the frame duration is defined as part of the encoding.

For frame-based codecs, the channel order is defined for the whole block. That is, for two-channel audio, right and left samples are coded independently, with the encoded frame for the left channel preceding that for the right channel.

All frame-oriented audio codecs should be able to encode and decode several consecutive frames within a single packet. Since the frame size for the frame-oriented codecs is given, there is no need to use a separate designation for the same encoding, but with different number of frames per packet.

4.4 Audio Encodings

encoding	sample/frame	bits/sample	ms/frame
~~1016~~	~~frame~~	~~N/A~~	~~30~~
~~DVI4~~	~~sample~~	~~4~~	
G721	sample	4	
G722	sample	8	
G728	frame	N/A	2.5
~~GSM~~	~~frame~~	~~N/A~~	~~20~~
~~L8~~	~~sample~~	~~8~~	
~~L16~~	~~sample~~	~~16~~	
~~LPC~~	~~frame~~	~~N/A~~	~~20~~
~~MPA~~	~~frame~~	~~N/A~~	
PCMA	sample	8	
~~PCMU~~	~~sample~~	~~8~~	
~~VDVI~~	~~sample~~	~~var.~~	

Table 1/Annex B-H.225.0: Properties of Audio Encodings

The characteristics of standard audio encodings are shown in Table 1 and their payload types are listed in Table 2.

See Appendix B for information on any coding not listed in Table 1. Support for such codings is not part of the H.323 recommendation.

~~4.4.1 1016~~

~~Encoding 1016 is a frame based encoding using code-excited linear prediction (CELP) and is specified in Federal Standard FED-STD 1016 [2,3,4,5].~~

~~The U. S. DoD's Federal-Standard-1016-based 4800 bps code-excited linear prediction voice coder version 3.2 (CELP 3.2) Fortran and C simulation source codes are available for worldwide distribution at no charge (on DOS diskettes, but configured to compile on Sun SPARC stations) from: Bob Fenichel, National Communications System, Washington, D.C. 20305, phone +1-703-692-2124, fax +1-703-746-4960.~~

~~4.4.2 DVI4~~

~~DVI4 is specified, with pseudo-code, in [6] as the IMA ADPCM wave type. A specification titled "DVI ADPCM Wave Type" can also be found in the Microsoft Developer Network Development Library CD~~

~~ROM published quarterly by Microsoft. The relevant section is found under Product Documentation, SDKs, Multimedia Standards Update, New Multimedia Data Types and Data Techniques, Revision 3.0, April 15, 1994. However, the encoding defined here as DVI4 differs in two respects from these recommendations:~~

~~o The header contains the predicted value rather than the first sample value.~~

~~o IMA ADPCM blocks contain odd number of samples, since the first sample of a block is contained just in the header (uncompressed), followed by an even number of compressed samples. DVI4 has an even number of compressed samples only, using the 'predict' word from the header to decode the first sample.~~

~~Each packet contains a single DVI block. The profile only defines the 4-bit-per-sample version, while IMA also specifies a 3-bit-per-sample encoding.~~

~~The "header" word for each channel has the following structure:~~

~~int16 predict; /* predicted value of first sample from the previous block (L16 format) */ u_int8 index; /* current index into stepsize table */ u_int8 reserved; /* set to zero by sender, ignored by receiver */~~

~~Packing of samples for multiple channels is for further study.~~

~~The document IMA Recommended Practices for Enhancing Digital Audio Compatibility in Multimedia Systems (version 3.0) contains the algorithm description. It is available from Interactive Multimedia Association 48 Maryland Avenue, Suite 202 Annapolis, MD 21401-8011 USA phone: +1 410 626-1380~~

~~4.4.3 G721~~

~~G721 is specified in ITU recommendation G.721. Reference implementations for G.721 are available as part of the CCITT/ITU-T Software Tool Library (STL) from the ITU General Secretariat, Sales Service, Place du Nations, CH-1211 Geneve 20, Switzerland. The library is covered by a license.~~

~~4.4.4 G722~~

~~G722 is specified in ITU-T recommendation G.722, "7 kHz audio-coding within 64 kbit/s".~~

~~4.4.5 G728~~

~~G728 is specified in ITU-T recommendation G.728, "Coding of speech at 16 kbit/s using low-delay code excited linear prediction".~~

~~4.4.6 GSM~~

~~GSM (group speciale mobile) denotes the European GSM 06.10 provisional standard for full-rate speech transcoding, prI-ETS 300 036, which is based on RPE/LTP (residual pulse excitation/long term prediction) coding at a rate of 13 kb/s [7,8,9]. The standard can be obtained from ETSI (European Telecommunications Standards Institute) ETSI Secretariat: B.P.152 F-06561 Valbonne Cedex France Phone: +33 92 94 42 00 Fax: +33 93 65 47 16~~

~~4.4.7 L8~~

~~L8 denotes linear audio data, using 8-bits of precision with an offset of 128, that is, the most negative signal is encoded as zero.~~

~~4.4.8 L16~~

~~L16 denotes uncompressed audio data, using 16-bit signed representation with 65535 equally divided steps between minimum and maximum signal level, ranging from -32768 to 32767. The value is represented in two's complement notation and network byte order.~~

~~4.4.9 LPC~~

~~LPC designates an experimental linear predictive encoding contributed by Ron Frederick, Xerox PARC, which is based on an implementation written by Ron Zuckerman, Motorola, posted to the Usenet group comp.dsp on June 26, 1992.~~

~~4.4.10 MPA~~

~~MPA denotes MPEG-I or MPEG-II audio encapsulated as elementary streams. The encoding is defined in ISO standards ISO/IEC 11172-3 and 13818-3. The encapsulation is specified in work in progress [10], Section 3. The authors can be contacted at Don Hoffman Sun Microsystems, Inc. Mail-stop UMPK14-305 2550 Garcia Avenue Mountain View, California 94043-1100 USA electronic mail: don.hoffman@eng.sun.com~~

~~Sampling rate and channel count are contained in the payload. MPEG-I audio supports sampling rates of 32000, 44100, and 48000 Hz (ISO/IEC 11172-3, section 1.1; "Scope"). MPEG-II additionally supports ISO/IEC 11172-3 Audio...").~~

4.4.11 PCMA

PCMA is specified in CCITT/ITU-T recommendation G.711. Audio data is encoded as eight bits per sample, after logarithmic scaling. ~~Code to convert between linear and A-law companded data is available in [6]. A detailed description is given by Jayant and Noll [11].~~

4.4.12 PCMU

PCMU is specified in CCITT/ITU-T recommendation G.711. Audio data is encoded as eight bits per sample, after logarithmic scaling. ~~Code to convert between linear and mu-law companded data is available in [6]. PCMU is the encoding used for the Internet media type audio/basic. A detailed description is given by Jayant and Noll [11].~~

~~4.4.13 VDVI~~

~~VDVI is a variable-rate version of DVI4, yielding speech bit rates of between 10 and 25 kb/s. It is specified for single-channel operation only. It uses the following encoding:~~

~~DVI4 codeword~~	~~VDVI bit pattern~~
~~0~~	~~00~~
~~1~~	~~010~~
~~2~~	~~1100~~
~~3~~	~~11100~~
~~4~~	~~111100~~
~~5~~	~~1111100~~
~~6~~	~~11111100~~
~~7~~	~~11111110~~
~~8~~	~~10~~
~~9~~	~~011~~
~~10~~	~~1101~~
~~11~~	~~11101~~
~~12~~	~~111101~~
~~13~~	~~1111101~~
~~14~~	~~11111101~~
~~15~~	~~11111111~~

5 Video

The following video encodings are currently defined, with their abbreviated names used for identification:

See Appendix B for any coding not described here. Such coding are not part of the H.323 recommendation.

~~5.1 CelB~~

~~The CELL-B encoding is a proprietary encoding proposed by Sun Microsystems. The byte stream format is described in work in progress [12]. The author can be contacted at Michael F. Speer Sun Microsystems Computer Corporation 2550 Garcia Ave MailStop UMPK14-305 Mountain View, CA 94043 United States electronic mail: michael.speer@eng.sun.com~~

~~5.2 JPEG~~

~~The encoding is specified in ISO Standards 10918-1 and 10918-2. The RTP payload format is as specified in work in progress [13]. Further information can be obtained from Steven McCanne Lawrence Berkeley National Laboratory M/S 46A-1123 One Cyclotron Road Berkeley, CA 94720 United States Phone: +1 510 486 7520 electronic mail: mccanne@ee.lbl.gov~~

5.3 H261

The encoding is specified in CCITT/ITU-T standard H.261. The packetization and RTP-specific properties are described in Annex C~~work in progress [14]~~. Further information can be obtained from Thierry Turletti Office NE 43-505 Telemedia, Networks and Systems Laboratory for Computer Science Massachusetts Institute of Technology, 545 Technology Square Cambridge, MA 02139 United States electronic mail: turletti@clove.lcs.mit.edu

~~5.4 MPV~~

~~MPV designates the use MPEG-I and MPEG-II video encoding elementary streams as specified in ISO Standards ISO/IEC 11172 and 13818-2, respectively. The RTP payload format is as specified in work in progress [10], Section 3. See the description of the MPA audio encoding for contact information.~~

~~5.5 MP2T~~

~~MP2T designates the use of MPEG-II transport streams, for either audio or video. The encapsulation is described in work in progress, [10], Section 2. See the description of the MPA audio encoding for contact information.~~

~~5.6 nv~~

~~The encoding is implemented in the program 'nv', version 4, developed at Xerox PARC by Ron Frederick. Further information is available from the author: Ron Frederick Xerox Palo Alto Research Center 3333 Coyote Hill Road Palo Alto, CA 94304 United States electronic mail: frederic@parc.xerox.com~~

6 Payload Type Definitions

Table 2 defines this profile's static payload type values for the PT field of the RTP data header. A new RTP payload format specification may be registered with the IANA by name, and may also be assigned a static payload type value from the range marked in Section 3.

In addition, payload type values in the range 96-127 may be defined dynamically through a conference control protocol, which is beyond the scope of this document. For example, a session directory could specify that for a given session, payload type 96 indicates PCMU encoding, 8,000 Hz sampling rate, 2 channels. The payload type range marked 'reserved' has been set aside so that RTCP and RTP packets can be reliably distinguished (see Section "Summary of Protocol Constants" of the RTP protocol specification).

An RTP source emits a single RTP payload type at any given time; the interleaving of several RTP payload types in a single RTP session is not allowed, but multiple RTP sessions may be used in parallel to send multiple media. The payload types currently defined in this profile carry either audio or video, but not both. However, it is allowed to define payload types that combine several media, e.g., audio and video, with appropriate separation in the payload format. Session participants agree through mechanisms beyond the scope of this specification on the set of payload types allowed in a given session. This set may, for example, be defined by the capabilities of the applications used, negotiated by a conference control protocol or established by agreement between the human participants.

Audio applications operating under this profile should, at minimum, be able to send and receive payload types 0 (PCMU) and 5 (DVI4). This allows interoperability without format negotiation and successful negotation with a conference control protocol.

All current video encodings use a timestamp frequency of 90,000 Hz, the same as the MPEG presentation time stamp frequency. This frequency yields exact integer timestamp increments for the typical 24 (HDTV), 25 (PAL), and 29.97 (NTSC) and 30 Hz (HDTV) frame rates and 50, 59.94 and 60 Hz field rates. While 90 kHz is the recommended rate for future video encodings used within this profile, other rates are possible. However, it is not sufficient to use the video frame rate (typically between 15 and 30 Hz) because that does not provide adequate resolution for typical synchronization requirements when calculating the RTP timestamp corresponding to the NTP timestamp in an RTCP SR packet [See Annex A]~~[15].~~ The timestamp resolution must also be sufficient for the jitter estimate contained in the receiver reports.

The standard video encodings and their payload types are listed in Table 2.

7 Port Assignment

As specified in the RTP protocol definition, RTP data is to be carried on an even UDP port number and the corresponding RTCP packets are to be carried on the next higher (odd) port number.

Applications operating under this profile may use any such UDP port pair. For example, the port pair may be allocated randomly by a session management program. A single fixed port number pair cannot be required because multiple applications using this profile are likely to run on the same host, and there are some operating systems that do not allow multiple processes to use the same UDP port with different multicast addresses. ***{Editor's Note: the deleted payload types will be listed as reserved, with a reference to Appendix B}***

PT	encoding name	audio/video (A/V)	clock rate (Hz)	channels (audio)
0	PCMU	A	8000	1
~~1~~	~~1016~~	~~A~~	~~8000~~	~~1~~
~~2~~	~~G721~~	~~A~~	~~8000~~	~~1~~
~~3~~	~~GSM~~	~~A~~	~~8000~~	~~1~~
~~4~~	~~unassigned~~	~~A~~	~~8000~~	~~1~~
~~5~~	~~DVI4~~	~~A~~	~~8000~~	~~1~~
~~6~~	~~DVI4~~	~~A~~	~~16000~~	~~1~~
~~7~~	~~LPC~~	~~A~~	~~8000~~	~~1~~
8	PCMA	A	8000	1
9	G722	A	8000	1
~~10~~	~~L16~~	~~A~~	~~44100~~	~~2~~
~~11~~	~~L16~~	~~A~~	~~44100~~	~~1~~
~~12~~	~~unassigned~~	~~A~~		
~~13~~	~~unassigned~~	~~A~~		
~~14~~	~~MPA~~	~~A~~	~~90000~~	~~(see text)~~
15	G728	A	8000	1
~~16—23~~	~~unassigned~~	~~A~~		
~~24~~	~~unassigned~~	~~V~~		
~~25~~	~~CelB~~	~~V~~	~~90000~~	
~~26~~	~~JPEG~~	~~V~~	~~90000~~	
~~27~~	~~unassigned~~	~~V~~		
~~28~~	~~nv~~	~~V~~	~~90000~~	
~~29~~	~~unassigned~~	~~V~~		

~~30~~	~~unassigned~~	~~V~~		
31	H261	V	90000	
~~32~~	~~MPV~~	~~V~~	~~90000~~	
~~33~~	~~MP2T~~	~~AV~~	~~90000~~	
~~34--71~~	~~unassigned~~	~~?~~		
~~72--76~~	~~reserved~~	~~N/A~~	~~N/A~~	~~N/A~~
~~77--95~~	~~unassigned~~	~~?~~		
96--127	dynamic	?		

Table 2/Annex B-H.225.0: Payload types (PT) for standard audio and video encodings

However, port numbers 5004 and 5005 have been registered for use with this profile for those applications that choose to use them as the default pair. Applications that operate under multiple profiles may use this port pair as an indication to select this profile if they are not subject to the constraint of the previous paragraph. Applications need not have a default and may require that the port pair be explicitly specified. The particular port numbers were chosen to lie in the range above 5000 to accomodate port number allocation practice within the Unix operating system, where port numbers below 1024 can only be used by privileged processes and port numbers between 1024 and 5000 are automatically assigned by the operating system.

~~8 Bibliography~~

~~[1] Apple Computer, "Audio interchange file format AIFF-C," Aug. 1991. (also ftp://ftp.sgi.com/sgi/aiff-c.9.26.91.ps.Z).~~

~~[2] Office of Technology and Standards, "Telecommunications: Analog to digital conversion of radio voice by 4,800 bit/second code excited linear prediction (celp)," Federal Standard FS-1016, GSA, Room 6654; 7th & D Street SW; Washington, DC 20407 (+1-202-708-9205), 1990.~~

~~[3] J. P. Campbell, Jr., T. E. Tremain, and V. C. Welch, "The proposed Federal Standard 1016 4800 bps voice coder: CELP," Speech Technology , vol. 5, pp. 58--64, April/May 1990.~~

~~[4] J. P. Campbell, Jr., T. E. Tremain, and V. C. Welch, "The federal standard 1016 4800 bps CELP voice coder," Digital Signal Processing , vol. 1, no. 3, pp. 145--155, 1991.~~

~~[5] J. P. Campbell, Jr., T. E. Tremain, and V. C. Welch, "The dod 4.8 kbps standard (proposed federal standard 1016)," in Advances in Speech Coding (B. Atal, V. Cuperman, and A. Gersho, eds.), ch. 12, pp. 121--133, Kluwer Academic Publishers, 1991.~~

~~[6] IMA Digital Audio Focus and Technical Working Groups, "Recommended practices for enhancing digital audio compatibility in multimedia systems (version 3.00)," tech. rep., Interactive Multimedia Association, Annapolis, Maryland, Oct. 1992.~~

~~[7] M. Mouly and M.-B. Pautet, The GSM system for mobile communications Lassay-les-Chateaux, France: Europe Media Duplication, 1993.~~

~~[8] J. Degener, "Digital speech compression," Dr. Dobb's Journal , Dec. 1994.~~

~~[9] S. M. Redl, M. K. Weber, and M. W. Oliphant, An Introduction to GSM Boston: Artech House, 1995.~~

~~[10] D. Hoffman and V. Goyal, "RTP payload format for MPEG1/MPEG2 video," Internet Draft, Internet Engineering Task Force, June 1995. Work in progress.~~

~~[11] N. S. Jayant and P. Noll, Digital Coding of Waveforms--Principles and Applications to Speech and Video Englewood Cliffs, New Jersey: Prentice-Hall, 1984.~~

~~[12] M. F. Speer and D. Hoffman, "RTP payload format of CellB video encoding," Internet Draft, Internet Engineering Task Force, Aug. 1995. Work in progress.~~

~~[13] W. Fenner, L. Berc, R. Frederick, and S. McCanne, "RTP encapsulation of JPEG-compressed video," Internet Draft, Internet Engineering Task Force, Mar. 1995. Work in progress.~~

~~[14] T. Turletti and C. Huitema, "RTP payload format for H.261 video streams," Internet Draft, Internet Engineering Task Force, July 1995. Work in progress.~~

~~[15] H. Schulzrinne, S. Casner, R. Frederick, and V. Jacobson, "RTP: A transport protocol for real-time applications." Internet draft (work-in-progress) draft-ietf-avt-rtp-*.txt , Mar. 1995.~~

~~9 Acknowledgements~~

~~The comments and careful review of Steve Casner are gratefully acknowledged.~~

~~10 Address of Author~~

~~Henning Schulzrinne GMD Fokus Hardenbergplatz 2 D-10623 Berlin Germany electronic mail: schulzrinne@fokus.gmd.de~~

~~Current Locations of Related Resources~~

~~UTF-8~~

~~Information on the UCS Transformation Format 8 (UTF-8) is available at~~

~~http://www.stonehand.com/unicode/standard/utf8.html~~

~~1016~~

~~An implementation is available at~~

~~ftp://ftp.super.org/pub/speech/celp_3.2a.tar.Z~~

~~DVI4~~

~~An implementation is available from Jack Jansen at~~

~~ftp://ftp.cwi.nl/local/pub/audio/adpcm.shar~~

~~G721~~

~~An implementation is available at~~

~~ftp://gaia.cs.umass.edu/pub/hgschulz/ccitt/ccitt_tools.tar.Z~~

~~GSM~~

~~A reference implementation was written by Carsten Borman and Jutta Degener (TU Berlin, Germany). It is available at~~

~~ftp://ftp.cs.tu-berlin.de/pub/local/kbs/tubmik/gsm/~~

~~LPC~~

~~An implementation is available at~~

~~ftp://parcftp.xerox.com/pub/net-research/lpc.tar.Z~~