3GPP TSG-SA Plenary Meeting #25
Palm Springs, CA, USA, 13 - 16 September 2004

Tdoc ⌘SP-040656

CR-Form-v7

CHANGE REQUEST

⌘	**26.234**	**CR 075**	⌘ **rev**	**1** ⌘	Current version:	**6.0.0** ⌘

*For **HELP** on using this form, see bottom of this page or look at the pop-up text over the ⌘ symbols.*

Proposed change affects: UICC apps⌘ [ ] ME [X] Radio Access Network [ ] Core Network [ ]

Title:	⌘	Introduction of the H.264 (AVC) video codec into the PSS service
Source:	⌘	Apple Computer, AT&T Wireless Services, Ericsson (editor), France Telecom, Fraunhofer, Nokia, ORANGE, PacketVideo, Panasonic, Philips, RealNetworks, Sharp, STMicroelectronics, Texas Instruments, Toshiba, Vodafone
Work item code:	⌘	PSSrel6-Stage3 — ***Date:*** ⌘ 10/09/2004
Category:	⌘	**B** — ***Release:*** ⌘ Rel-6

Use one of the following categories:	*Use one of the following releases:*
F *(correction)*	*2 (GSM Phase 2)*
A *(corresponds to a correction in an earlier release)*	*R96 (Release 1996)*
B *(addition of feature),*	*R97 (Release 1997)*
C *(functional modification of feature)*	*R98 (Release 1998)*
D *(editorial modification)*	*R99 (Release 1999)*
Detailed explanations of the above categories can	*Rel-4 (Release 4)*
be found in 3GPP TR 21.900.	*Rel-5 (Release 5)*
	Rel-6 (Release 6)

Reason for change:	⌘	Improve video quality by introducing H.264 (AVC) video in Release 6.
Summary of change:	⌘	References to H.264 (AVC) and the corresponding RTP payload format are added. Buffer management of H.264 in the PSS service is defined.
Consequences if not approved:	⌘	Release-6 will not support the new video codec H.264 (AVC).

Clauses affected:	⌘	2, 3.2, 5.3.2.4, 5.3.3.2, 5.4, 6.2.4, 7.4, G.2			
		Y	**N**		
Other specs affected:	⌘		**X**	Other core specifications	⌘
			X	Test specifications	
			X	O&M Specifications	
Other comments:	⌘	Storage of H.264 in 3GP files is included in CR 26.244 004			

How to create CRs using this form:

Comprehensive information and tips about how to create CRs can be found at http://www.3gpp.org/specs/CR.htm. Below is a brief summary:

1) Fill out the above form. The symbols above marked ⌘ contain pop-up help information about the field that they are closest to.

2) Obtain the latest version for the release of the specification to which the change is proposed. Use the MS Word "revision marks" feature (also known as "track changes") when making the changes. All 3GPP specifications can be

downloaded from the 3GPP server under ftp://ftp.3gpp.org/specs/ For the latest version, look for the directory name with the latest date e.g. 2001-03 contains the specifications resulting from the March 2001 TSG meetings.

3) With "track changes" disabled, paste the entire CR form (use CTRL-A to select it) into the specification just in front of the clause containing the first piece of changed text. Delete those parts of the specification which are not relevant to the change request.

2 References

The following documents contain provisions which, through reference in this text, constitute provisions of the present document.

- References are either specific (identified by date of publication, edition number, version number, etc.) or non-specific.
- For a specific reference, subsequent revisions do not apply.
- For a non-specific reference, the latest version applies. In the case of a reference to a 3GPP document (including a GSM document), a non-specific reference implicitly refers to the latest version of that document *in the same Release as the present document*.

[1] 3GPP TS 22.233: "Transparent End-to-End Packet-switched Streaming Service; Stage 1".

[2] 3GPP TS 26.233: "Transparent end-to-end packet switched streaming service (PSS); General description".

[3] 3GPP TR 21.905: "Vocabulary for 3GPP Specifications".

[4] IETF RFC 1738: "Uniform Resource Locators (URL)", Berners-Lee T., Masinter L. and McCahill M., December 1994.

[5] IETF RFC 2326: "Real Time Streaming Protocol (RTSP)", Schulzrinne H., Rao A. and Lanphier R., April 1998.

[6] IETF RFC 2327: "SDP: Session Description Protocol", Handley M. and Jacobson V., April 1998.

[7] IETF STD 0006: "User Datagram Protocol", Postel J., August 1980.

[8] IETF STD 0007: "Transmission Control Protocol", Postel J., September 1981.

[9] IETF RFC 3550: "RTP: A Transport Protocol for Real-Time Applications", Schulzrinne H. et al., July 2003.

[10] IETF RFC 3551: "RTP Profile for Audio and Video Conferences with Minimal Control", Schulzrinne H. and Casner S., July 2003.

[11] IETF RFC 3267: "Real-Time Transport Protocol (RTP) Payload Format and File Storage Format for the Adaptive Multi-Rate (AMR) Adaptive Multi-Rate Wideband (AMR-WB) Audio Codecs", Sjoberg J. et al., June 2002.

[12] (void)

[13] IETF RFC 3016: "RTP Payload Format for MPEG-4 Audio/Visual Streams", Kikuchi Y. et al., November 2000.

[14] IETF RFC 2429: "RTP Payload Format for the 1998 Version of ITU-T Rec. H.263 Video (H.263+)", Bormann C. et al., October 1998.

[15] IETF RFC 2046: "Multipurpose Internet Mail Extensions (MIME) Part Two: Media Types", Freed N. and Borenstein N., November 1996.

[16] IETF RFC 3236: "The 'application/xhtml+xml' Media Type", Baker M. and Stark P., January 2002.

[17] IETF RFC 2616: "Hypertext Transfer Protocol ñ HTTP/1.1", Fielding R. et al., June 1999.

[18] (void)

[19] 3GPP TS 26.101: "Mandatory Speech Codec speech processing functions; Adaptive Multi-Rate (AMR) speech codec frame structure".

[20] 3GPP TS 26.171: "AMR Wideband Speech Codec; General Description".

[21] ISO/IEC 14496-3:2001: "Information technology ñ Coding of audio-visual objects ñ Part 3: Audio".

[22] ITU-T Recommendation H.263 (1998): "Video coding for low bit rate communication".

[23] ITU-T Recommendation H.263 ñ Annex X (2001): "Annex X: Profiles and levels definition".

[24] ISO/IEC 14496-2:2001: "Information technology ñ Coding of audio-visual objects ñ Part 2: Visual".

[25] ISO/IEC 14496-2:2001/Amd 2:2002: "Streaming video profile".

[26] ITU-T Recommendation T.81 (1992) | ISO/IEC 10918-1:1993: "Information technology ñ Digital compression and coding of continuous-tone still images ñ Requirements and guidelines".

[27] C-Cube Microsystems: "JPEG File Interchange Format", Version 1.02, September 1, 1992.

[28] W3C Recommendation: "XHTML Basic", http://www.w3.org/TR/2000/REC-xhtml-basic-20001219, December 2000.

[29] ISO/IEC 10646-1:2000: "Information technology ñ Universal Multiple-Octet Coded Character Set (UCS) ñ Part 1: Architecture and Basic Multilingual Plane".

[30] The Unicode Consortium: "The Unicode Standard", Version 3.0 Reading, MA, Addison-Wesley Developers Press, 2000, ISBN 0-201-61633-5.

[31] W3C Recommendation: "Synchronized Multimedia Integration Language (SMIL 2.0)", http://www.w3.org/TR/2001/REC-smil20-20010807/, August 2001.

[32] CompuServe Incorporated: "GIF Graphics Interchange Format: A Standard defining a mechanism for the storage and transmission of raster-based graphics information", Columbus, OH, USA, 1987.

[33] CompuServe Incorporated: "Graphics Interchange Format: Version 89a", Columbus, OH, USA, 1990.

[34] (void)

[35] (void)

[36] (void)

[37] (void)

[38] IETF RFC 2083: "PNG (Portable Networks Graphics) Specification Version 1.0", Boutell T., et al., March 1997.

[39] W3C Recommendation: "Composite Capability/Preference Profiles (CC/PP): Structure and Vocabularies 1.0", http://www.w3.org/TR/2004/REC-CCPP-struct-vocab-20040115/, January 2004.

[40] WAP UAProf Specification, http://www1.wapforum.org/tech/terms.asp?doc=WAP-248-UAProf-20011020-a.pdf, October 2001.

[41] W3C Recommendation: "RDF Vocabulary Description Language 1.0: RDF Schema", http://www.w3.org/TR/2004/REC-rdf-schema-20040210/, February 2004.

[42] W3C Recommendation: "Scalable Vector Graphics (SVG) 1.1 Specification", http://www.w3.org/TR/2003/REC-SVG11-20030114/, January 2003.

[43] W3C Recommendation: "Mobile SVG Profiles: SVG Tiny and SVG Basic", http://www.w3.org/TR/2003/REC-SVGMobile-20030114/, January 2003.

[44] Scalable Polyphony MIDI Specification Version 1.0, RP-34, MIDI Manufacturers Association, Los Angeles, CA, February 2002.

[45] Scalable Polyphony MIDI Device 5-to-24 Note Profile for 3GPP Version 1.0, RP-35, MIDI Manufacturers Association, Los Angeles, CA, February 2002.

[46] "Standard MIDI Files 1.0", RP-001, in "The Complete MIDI 1.0 Detailed Specification, Document Version 96.1", The MIDI Manufacturers Association, Los Angeles, CA, USA, February 1996.

[47] WAP Forum Specification: "XHTML Mobile Profile", http://www1.wapforum.org/tech/terms.asp?doc=WAP-277-XHTMLMP-20011029-a.pdf, October 2001.

[48] (void)

[49] IETF RFC 3266: "Support for IPv6 in Session Description Protocol (SDP)", Olson S., Camarillo G. and Roach A. B., June 2002.

[50] 3GPP TS 26.244: "Transparent end-to-end packet switched streaming service (PSS); 3GPP file format (3GP)".

[51] 3GPP TS 26.245: "Transparent end-to-end packet switched streaming service (PSS); Timed text format".

[52] 3GPP TS 26.246: "Transparent end-to-end packet switched streaming service (PSS); 3GPP SMIL Language Profile".

[53] IETF RFC 2234: "Augmented BNF for Syntax Specifications: ABNF", Crocker D. and Overell P., November 1997.

[54] IETF RFC 3066: "Tags for Identification of Languages", Alvestrand H., January 2001.

[55] IETF RFC 3556: "Session Description Protocol (SDP) Bandwidth Modifiers for RTP Control Protocol (RTCP) Bandwidth", Casner S., July 2003.

[56] 3GPP TS 23.107: "Quality of Service (QoS) concept and architecture".

[57] IETF Internet Draft: "Extended RTP Profile for RTCP-based Feedback (RTP/AVPF)", Ott J. et al, http://www.ietf.org/internet-drafts/draft-ietf-avt-rtcp-feedback-08.txt, January 2004.

[58] IETF RFC 3611: "RTP Control Protocol Extended Reports (RTCP XR)", Friedman T., Caceres R. and Clark A., November 2003.

[59] IETF RFC 1952: "GZIP file format specification version 4.3", Deutsch P., May 1996.

[60] IETF RFC 2396: "Uniform Resource Identifiers (URI): Generic Syntax", Berners-Lee T., Fielding R., Irvine U.C. and Masinter L., August 1998.

[61] IETF RFC 2732: "Format for Literal IPv6 Addresses in URL's", Hinden R., Carpenter B. and Masinter L., December 1999.

[62] IETF RFC 3555: "MIME Type Registration of RTP Payload Formats", Casner S. and Hoschka P., July 2003.

[63] 3GPP TS 26.090: "Mandatory Speech Codec speech processing functions; Adaptive Multi-Rate (AMR) speech codec; Transcoding functions".

[64] 3GPP TS 26.073: "ANSI-C code for the Adaptive Multi Rate (AMR) speech codec".

[65] 3GPP TS 26.104: "ANSI-C code for the floating-point Adaptive Multi Rate (AMR) speech codec".

[66] 3GPP TS 26.190: "Speech Codec speech processing functions; AMR Wideband speech codec; Transcoding functions".

[67] 3GPP TS 26.173: "ANCI-C code for the Adaptive Multi Rate - Wideband (AMR-WB) speech codec".

[68] 3GPP TS 26.204: "ANSI-C code for the Floating-point Adaptive Multi-Rate Wideband (AMR-WB) speech codec".

[69] IETF RFC 3548: "The Base16, Base32, and Base64 Data Encodings", Josefsson S., Ed., July 2003.

[70] Mobile DLS, MMA specification v1.0. RP-41 Los Angeles, CA, USA. 2004.

[71] Mobile XMF Content Format Specification, MMA specification v1.0., RP-42, Los Angeles, CA, USA. 2004.

[72] ITU-T Recommendation H.264 (2003): "Advanced video coding for generic audiovisual services" | ISO/IEC 14496-10:2003: "Information technology ñ Coding of audio-visual objects ñ Part 10: Advanced Video Coding".

[73] ISO/IEC 14496-10/FDAM1: "AVC Fidelity Range Extensions".

[74] IETF Internet Draft: "RTP payload Format for H.264 Video", Wenger S. et al, http://www.ietf.org/internet-drafts/draft-ietf-avt-rtp-h264-11.txt, August 2004.

3.2 Abbreviations

For the purposes of the present document, the abbreviations given in 3GPP TR 21.905 [3] and the following apply.

3GP	3GPP file format
AAC	Advanced Audio Coding
AVC	Advanced Video Coding
CC/PP	Composite Capability / Preference Profiles
DCT	Discrete Cosine Transform
DLS	Downloadable Sounds
GIF	Graphics Interchange Format
HTML	Hyper Text Markup Language
ITU-T	International Telecommunications Union ñ Telecommunications
JFIF	JPEG File Interchange Format
MIDI	Musical Instrument Digital Interface
MIME	Multipurpose Internet Mail Extensions
MMS	Multimedia Messaging Service
PNG	Portable Networks Graphics
PSS	Packet-switched Streaming Service
QCIF	Quarter Common Intermediate Format
RDF	Resource Description Framework
RTCP	RTP Control Protocol
RTP	Real-time Transport Protocol
RTSP	Real-Time Streaming Protocol
SDP	Session Description Protocol

SMIL	Synchronised Multimedia Integration Language
SP-MIDI	Scalable Polyphony MIDI
SVG	Scalable Vector Graphics
UAProf	User Agent Profile
UCS-2	Universal Character Set (the two octet form)
UTF-8	Unicode Transformation Format (the 8-bit form)
W3C	WWW Consortium
WML	Wireless Markup Language
XHTML	eXtensible Hyper Text Markup Language
XMF	eXtensible Music Format
XML	eXtensible Markup Language

5.3.2.4 Video buffering headers

The following header fields are specified for the response of an RTSP PLAY request only:

- x-predecbufsize:<size of the pre-decoder buffer>
- x-initpredecbufperiod:<initial pre-decoder buffering period>
- x-initpostdecbufperiod:<initial post-decoder buffering period>
- 3gpp-videopostdecbufsize:<size of the video post-decoder buffer>

The header fields "x-predecbufsize", "x-initpredecbufperiod", "x-initpostdecbufperiod", and "3gpp-postdecbufsize" have the same definitions as the corresponding SDP attributes (see clause 5.3.3.2) "X-predecbufsize", "X-initpredecbufperiod", "X-initpostdecbufperiod", and "3gpp-postdecbufsize", respectively, with the exception that the RTSP video buffering header fields are valid only for the range specified in the RTSP PLAY response.

For H.263 and MPEG-4 Visual, the usage of these header fields is specified in Annex G.

For H.264 (AVC), PSS servers shall include these header fields in an RTSP PLAY response whenever the values are available in the 3GP file used for the streaming session. If the values are not available in the 3GP file, it is optional for the servers to signal the parameter values in RTSP PLAY responses.

5.3.3.2 Additional SDP fields

The following Annex G and H.264 (AVC) -related media level SDP fields are defined for PSS:

- "a=X-predecbufsize:<size of the hypothetical pre-decoder buffer>"
  If the field is an attribute for an H.263 or MPEG-4 Visual stream and rate adaptation (see clause 10.2) is not in use, this gives the suggested size of the Annex G hypothetical pre-decoder buffer in bytes.

  If the field is an attribute for an H.263 or MPEG-4 Visual stream and rate adaptation is in use, this gives the suggested minimum size of a buffer (hereinafter called the pre-decoder buffer) that is used to smooth out transmit time variation (compared to flat-bitrate transmission scheduling) and video bitrate variation.

If the field is an attribute for an H.264 (AVC) stream, the H.264 (AVC) bitstream is constrained by the value of "CpbSize" equal to X-predecbufsize * 8 for NAL HRD parameters, as specified in [72]. For the VCL HRD parameters, the value of "CpbSize" is equal to X-predecbufsize * 40 / 6. The value of "X-predecbufsize" for H.264 (AVC) streams shall be smaller than or equal to 1200 * MaxCPB, in which the value of "MaxCPB" is derived according to the H.264 (AVC) profile and level of the stream, as specified in [72]. If "X-predecbufsize" is not present for an H.264 (AVC) stream, the value of "CpbSize" is calculated as specified in [72].

- "a=X-initpredecbufperiod:<initial pre-decoder buffering period>"
  If the field is an attribute for an H.263 or MPEG-4 Visual stream and rate adaptation is not in use, this gives the required initial pre-decoder buffering period specified according to Annex G. Values are interpreted as clock ticks of a 90-kHz clock. That is, the value is incremented by one for each 1/90 000 seconds. For example, value 180 000 corresponds to a two second initial pre-decoder buffering.

  If the field is an attribute for an H.263 or MPEG-4 Visual stream and rate adaptation is in use, this gives the suggested minimum greatest difference in RTP timestamps in the pre-decoder buffer after any de-interleaving has been applied. Note that "X-initpredecbufperiod" is expressed as clock ticks of a 90-kHz clock. Hence, conversion may be required if the RTP timestamp clock frequency is not 90 kHz.

  If the field is an attribute for an H.264 (AVC) stream, the H.264 (AVC) bitstream is constrained by the value of the nominal removal time of the first access unit from the coded picture buffer (CPB), $t_{r,n}(0)$, equal to "X-initpredecbufperiod" as specified in [72]. If "X-initpredecbufperiod" is not present for an H.264 (AVC) stream, $t_{r,n}(0)$ shall be equal to the earliest time when the first access unit in decoding order has been completely received.

- "a=X-initpostdecbufperiod:<initial post-decoder buffering period>"
  If the field is an attribute for an H.263 or MPEG-4 Visual stream and rate adaptation is not in use, this gives the required initial post-decoder buffering period specified according to Annex G. Values are interpreted as clock ticks of a 90-kHz clock.

  If the field is an attribute for an H.263 or MPEG-4 Visual stream and rate adaptation is in use, this gives the initial post-decoder buffering period assuming that the hypothetical decoding and post-decoder buffering model given in points 5 to 10 in Annex G clause G.3 would be followed. Note that the operation of the post-decoder buffer is logically independent from rate adaptation and is used to compensate non-instantaneous decoding of pictures.

  If the field is an attribute for an H.264 (AVC) stream, the H.264 (AVC) bitstream is constrained by the value of dpb_output_delay for the first decoded picture in output order equal to "X-initpostdecbufperiod" as specified in [72] assuming that the clock tick variable, t_c, is equal to 1 / 90 000. If "X-initpostdecbufperiod" is not present for an H.264 (AVC) stream, the value of dpb_output_delay for the first decoded picture in output order is inferred to be equal to 0.

- "a=X-decbyterate:<peak decoding byte rate>"
  If the field is an attribute for an H.263 or MPEG-4 Visual stream and rate adaptation is not in use, ~~T~~this gives the peak decoding byte rate that was used to verify the compatibility of the stream with Annex G. Values are given in bytes per second.

  If the field is an attribute for an H.263 or MPEG-4 Visual stream and rate adaptation is in use, "X-decbyterate" has no meaning.

  This field shall not be present for H.264 (AVC) streams.

- "a=3gpp-videopostdecbufsize:<size of the video post-decoder buffer>"
  This attribute may be present for H.264 (AVC) streams and it shall not be present for other types of streams. If the attribute is present, the H.264 (AVC) bitstream is constrained by the value of "max_dec_frame_buffering" equal to Min(16, Floor(3gpp-videopostdecbufsize / (PicWidthInMbs * FrameHeightInMbs * 256 * ChromaFormatFactor))) as specified in [72]. If "3gpp-videopostdecbufsize" is not present for an H.264 (AVC) stream, the value of "max_dec_frame_buffering" is inferred as specified in [72].

If none of the attributes "a=X-predecbufsize:", "a=X-initpredecbufperiod:", "a=X-initpostdecbufperiod:", and "a=x-decbyterate:" is present for an H.263 or MPEG-4 Visual stream, clients should not expect a packet stream according to Annex G. If at least one of the listed attributes is present for an H.263 or MPEG-4 Visual stream, and if the client does not choose the usage of bit-rate adaptation via RTSP as described in clause 5.3.2.2, the transmitted video packet stream shall conform to Annex G. If at least one of the listed attributes is present for an H.263 or MPEG-4 Visual stream, but some of the listed attributes are missing in an SDP description, clients should expect a default value for the missing attributes according to Annex G.

If the interleaved packetization mode of H.264 (AVC) is in use, attributes "a=X-predecbufsize:", "a=X-initpredecbufperiod:", "a=X-initpostdecbufperiod:", and "a=3gpp-videopostdecbufsize:" apply to an H.264 (AVC) bitstream when de-interleaving of the stream from transmission order to decoding order has been done.

The following media level SDP field is defined for PSS:

- "a=framesize:<payload type number> <width>-<height>"
  This gives the largest video frame size of H.263 streams.

The frame size field in SDP is needed by the client in order to properly allocate frame buffer memory. For MPEG-4 ~~v~~Visual streams, the frame size shall be extracted from the "config" information in the SDP. For H.264 (AVC) streams, the frame size shall be extracted from the sprop-parameters-sets information in the SDP. For H.263 streams, a PSS server shall include the "a=framesize" field at the media level for each stream in SDP, and a PSS client should interpret this field, if present. Clients should be ready to receive SDP descriptions without this attribute.

If this attribute is present, the frame size parameters shall exactly match the largest frame size defined in the video stream. The width and height values shall be expressed in pixels.

5.4 MIME media types

For continuous media (speech, audio and video) the following MIME media types shall be used:

- AMR narrow-band speech codec (see clause 7.2) MIME media type as defined in [11];
- AMR wideband speech codec (see clause 7.2) MIME media type as defined in [11];
- MPEG-4 AAC audio codec (see clause 7.3) MIME media type as defined in RFC 3016 [13]. When used in SDP the attribute ìcpresentî SHALL be set to ì0î indicating that the configuration information is only carried out of band in the SDP ìconfigî parameter;
- MPEG-4 video codec (see clause 7.4) MIME media type as defined in RFC 3016 [13]. When used in SDP the configuration information shall be carried outband in the "config" SDP parameter and inband (as stated in RFC 3016). As described in RFC 3016, the configuration information sent inband and the config information in the SDP shall be the same except that first_half_vbv_occupancy and latter_half_vbv_occupancy which, if exist, may vary in the configuration information sent inband;
- H.263 [22] video codec (see clause 7.4) MIME media type as defined in clause 4.2.7 of [62];~~.~~
- H.264 (AVC) [72] video codec (see clause 7.4) MIME media type as defined in [74].

MIME media types for JPEG, GIF, PNG, SP-MIDI, Mobile DLS, Mobile XMF, SVG, timed text and XHTML can be used both in the "Content-type" field in HTTP and in the "type" attribute in SMIL 2.0. The following MIME media types shall be used for these media:

- JPEG (see clause 7.5) MIME media type as defined in [15];
- GIF (see clause 7.6) MIME media type as defined in [15];
- PNG (see sub clause 7.6) MIME media type as defined in [38];
- SP-MIDI (see sub clause 7.3A) MIME media type as defined in clause C.2 in Annex C of the present document;

- DLS MIME media type to represent Mobile DLS (see sub clause 7.3A) as defined in clause C.4 in Annex C of the present document;
- Mobile XMF (see sub clause 7.3A) MIME media type as defined in clause C.3 in Annex C of the present document;
- SVG (see sub clause 7.7) MIME media type as defined in [42];
- XHTML (see clause 7.8) MIME media type as defined in [16];
- Timed text (see subclause 7.9) MIME media type as defined in [50].

MIME media type used for SMIL files shall be according to [31] and for SDP files according to [6].

6.2.4 RTP payload formats

For RTP/UDP/IP transport of continuous media the following RTP payload formats shall be used:

- AMR narrow-band speech codec (see clause 7.2) RTP payload format according to [11]. A PSS client is not required to support multi-channel sessions;
- AMR wideband speech codec (see clause 7.2) RTP payload format according to [11]. A PSS client is not required to support multi-channel sessions;
- MPEG-4 AAC audio codec (see clause 7.3) RTP payload format according to RFC 3016 [13];
- MPEG-4 video codec (see clause 7.4) RTP payload format according to RFC 3016 [13];
- H.263 video codec (see clause 7.4) RTP payload format according to RFC 2429 [14];~~.~~
- H.264 (AVC) video codec (see clause 7.4) RTP payload format according to [74]. A PSS client is required to support all three packetization modes: single NAL unit mode, non-interleaved mode and interleaved mode. For the interleaved packetization mode, a PSS client shall support streams for which the value of the "sprop-deint-buf-req" MIME parameter is less than or equal to MaxCPB * 1000 / 8, inclusive, in which "MaxCPB" is the value for VCL parameters of the H.264 (AVC) profile and level in use, as specified in [72]. Parameter sets shall not be transmitted within the RTP payload, i.e., all parameter sets required for a session must be provided in the SDP.

NOTE: The payload format RFC 3016 for MPEG-4 AAC specify that the audio streams shall be formatted by the LATM (Low-overhead MPEG-4 Audio Transport Multiplex) tool [21]. It should be noted that the references for the LATM format in the RFC 3016 [13] point to an older version of the LATM format than included in [21]. In [21] a corrigendum to the LATM tool is included. This corrigendum includes changes to the LATM format making implementations using the corrigendum incompatible with implementations not using it. To avoid future interoperability problems, implementations of PSS client and servers supporting AAC shall follow the changes to the LATM format included in [21].

7.4 Video

If video is supported, ITU-T Recommendation H.263 [22] profile 0 level 10 decoder shall be supported. In addition, a PSS client should support:

- H.263 [23] Profile 3 Level 10 decoder;
- MPEG-4 Visual Simple Profile Level 0 decoder, [24] and [25];~~.~~
- H.264 (AVC) Baseline Profile Level 1b decoder [72][73] with constraint_set1_flag=1 and without requirements on output timing conformance (Annex C of [72]).

The video buffer model given in Annex G of the present document should be supported if H.263 or MPEG-4 Visual ~~video~~ is supported. It shall not be used with H.264 (AVC).

The H.264 (AVC) decoder in a PSS client shall start decoding immediately when it receives data (even if the stream does not start with an IDR access unit) or alternatively no later than it receives the next IDR access unit or the next recovery point SEI message, whichever is earlier in decoding order. Note that when the interleaved packetization mode of H.264 (AVC) is in use, de-interleaving is done normally before starting the decoding process. The decoding process for a stream not starting with an IDR access unit shall be the same as for a valid H.264 (AVC) bitstream. However, the client shall be aware that such a stream may contain references to pictures not available in the decoded picture buffer. The display behaviour of the client is out of scope of this specification.

A PSS client supporting H.264 (AVC) should ignore any VUI HRD parameters, buffering period SEI message, and picture timing SEI message in H.264 (AVC) streams or conveyed in the "sprop-parameter-sets" MIME/SDP parameter. Instead, a PSS client supporting H.264 (AVC) shall follow buffering parameters conveyed in SDP, as specified in clause 5.3.2.2, and in RTSP, as specified in clause 5.3.2.4. A PSS client shall also use the RTP timestamp or NALU-time (as specified in [74]) of a picture as its presentation time, and, when the interleaved RTP packetization mode is in use, follow the "sprop-interleaving-depth", "sprop-deint-buf-req", "sprop-init-buf-time", and "sprop-max-don-diff" MIME/SDP parameters for the de-interleaving process. However, if VUI HRD parameters, buffering period SEI messages, and picture timing SEI messages are present in the bitstream, their contents shall not contradict any of the parameters mentioned in the previous sentence.

NOTE: ITU-T Recommendation H.263 profile 0 has been mandated to ensure that video-enabled PSS supports a minimum baseline video capability. Both H.263 and MPEG-4 ~~v~~Visual decoders can decode an H.263 profile 0 bitstream. It is strongly recommended, though, that an H.263 profile 0 bitstream is transported and stored as H.263 and not as MPEG-4 ~~v~~Visual (short header), as MPEG-4 ~~v~~Visual is not mandated by PSS.

G.2 PSS Buffering Parameters

The behaviour of the PSS buffering model is controlled with the following parameters: the initial pre-decoder buffering period, the initial post-decoder buffering period, the size of the hypothetical pre-decoder buffer, the peak decoding byte rate, and the decoding macroblock rate. The default values of the parameters are defined below.

- The default initial pre-decoder buffering period is 1 second.
- The default initial post-decoder buffering period is zero.
- The default size of the hypothetical pre-decoder buffer is defined according to the maximum video bit-rate according to the table below:

Table G.1: Default size of the hypothetical pre-decoder buffer

Maximum video bit-rate	Default size of the hypothetical pre-decoder buffer
65536 bits per second	20480 bytes
131072 bits per second	40960 bytes
Undefined	51200 bytes

- The maximum video bit-rate can be signalled in the media-level bandwidth attribute of SDP as defined in clause 5.3.3 of this document. If the video-level bandwidth attribute was not present in the presentation description, the maximum video bit-rate is defined according to the video coding profile and level in use.
- The size of the hypothetical post-decoder buffer is an implementation-specific issue. The buffer size can be estimated from the maximum output data rate of the decoders in use and from the initial post-decoder buffering period.
- By default, the peak decoding byte rate is defined according to the video coding profile and level in use. For example, H.263 Level 10 requires support for bit-rates up to 64000 bits per second. Thus, the peak decoding byte rate equals to 8000 bytes per second.
- The default decoding macroblock rate is defined according to the video coding profile and level in use. If MPEG-4 Visual is in use, the default macroblock rate equals to VCV decoder rate. If H.263 is in use, the default macroblock rate equals to (1 / minimum picture interval) multiplied by number of macroblocks in maximum picture format. For example, H.263 Level 10 requires support for picture formats up to QCIF and minimum picture interval down to 2002 / 30000 sec. Thus, the default macroblock rate would be 30000 x 99 / 2002 ≈ 1484 macroblocks per second.

PSS clients may signal their capability of providing larger buffers and faster peak decoding byte rates in the capability exchange process described in clause 5.2 of the present document. The average coded video bit-rate should be smaller than or equal to the bit-rate indicated by the video coding profile and level in use, even if a faster peak decoding byte rate were signalled.

Initial parameter values for each stream can be signalled within the SDP description of the stream. Signalled parameter values override the corresponding default parameter values. The values signalled within the SDP description guarantee pauseless playback from the beginning of the stream until the end of the stream (assuming a constant-delay reliable transmission channel).

PSS servers may update parameter values in the response for an RTSP PLAY request. If an updated parameter value is present, it shall replace the value signalled in the SDP description or the default parameter value in the operation of the PSS buffering model. An updated parameter value is valid only in the indicated playback range, and it has no effect after that. Assuming a constant-delay reliable transmission channel, the updated parameter values guarantee pauseless playback of the actual range indicated in the response for the PLAY request. The indicated pre-decoder buffer size and initial post-decoder buffering period shall be smaller than or equal to the corresponding values in the SDP description or the corresponding default values, whichever ones are valid. The header fields for RTSP are specified in clause 5.3.2.4.~~The following header fields are defined for RTSP:~~

- ~~x-predecbufsize:<size of the hypothetical pre-decoder buffer>~~
  ~~This gives the suggested size of the Annex G hypothetical pre-decoder buffer in bytes.~~
- ~~x-initpredecbufperiod:<initial pre-decoder buffering period>~~
  ~~This gives the required initial pre-decoder buffering period specified according to Annex G. Values are~~

~~interpreted as clock ticks of a 90-kHz clock. That is, the value is incremented by one for each 1/90 000 seconds. For example, value 180 000 corresponds to a two second initial pre-decoder buffering.~~

- ~~x-initpostdecbufperiod:<initial post-decoder buffering period>~~
  ~~This gives the required initial post-decoder buffering period specified according to Annex G. Values are interpreted as clock ticks of a 90-kHz clock.~~

~~These header fields are defined for the response of an RTSP PLAY request only. Their use is optional.~~

The following example plays the whole presentation starting at SMPTE time code 0:10:20 until the end of the clip. The playback is to start at 15:36 on 23 Jan 1997. The suggested initial pre-decoder buffering period is half a second.

```
C->S: PLAY rtsp://audio.example.com/twister.en RTSP/1.0
      CSeq: 833
      Session: 12345678
      Range: smpte=0:10:20-;time=19970123T153600Z
      User-Agent: TheStreamClient/1.1b2

S->C: RTSP/1.0 200 OK
      CSeq: 833
      Date: 23 Jan 1997 15:35:06 GMT
      Range: smpte=0:10:22-;time=19970123T153600Z
      x-initpredecbufperiod: 45000
```