ETSI TS 126 234 V9.3.0 (2010-06)

Technical Specification

Universal Mobile Telecommunications System (UMTS);
LTE;
Transparent end-to-end Packet-switched
Streaming Service (PSS);
Protocols and codecs
(3GPP TS 26.234 version 9.3.0 Release 9)

Reference
RTS/TSGS-0426234v930

Keywords
LTE, UMTS

ETSI

650 Route des Lucioles
F-06921 Sophia Antipolis Cedex - FRANCE

Tel.: +33 4 92 94 42 00 Fax: +33 4 93 65 47 16

Siret N° 348 623 562 00017 - NAF 742 C
Association à but non lucratif enregistrée à la
Sous-Préfecture de Grasse (06) N° 7803/88

Important notice

Individual copies of the present document can be downloaded from:
http://www.etsi.org

The present document may be made available in more than one electronic version or in print. In any case of existing or perceived difference in contents between such versions, the reference version is the Portable Document Format (PDF). In case of dispute, the reference shall be the printing on ETSI printers of the PDF version kept on a specific network drive within ETSI Secretariat.

Users of the present document should be aware that the document may be subject to revision or change of status. Information on the current status of this and other ETSI documents is available at
http://portal.etsi.org/tb/status/status.asp

If you find errors in the present document, please send your comment to one of the following services:
http://portal.etsi.org/chaircor/ETSI_support.asp

Intellectual Property Rights

IPRs essential or potentially essential to the present document may have been declared to ETSI. The information pertaining to these essential IPRs, if any, is publicly available for **ETSI members and non-members**, and can be found in ETSI SR 000 314: *"Intellectual Property Rights (IPRs); Essential, or potentially Essential, IPRs notified to ETSI in respect of ETSI standards"*, which is available from the ETSI Secretariat. Latest updates are available on the ETSI Web server (http://webapp.etsi.org/IPR/home.asp).

Pursuant to the ETSI IPR Policy, no investigation, including IPR searches, has been carried out by ETSI. No guarantee can be given as to the existence of other IPRs not referenced in ETSI SR 000 314 (or the updates on the ETSI Web server) which are, or may be, or may become, essential to the present document.

Foreword

This Technical Specification (TS) has been produced by ETSI 3rd Generation Partnership Project (3GPP).

The present document may refer to technical specifications or reports using their 3GPP identities, UMTS identities or GSM identities. These should be interpreted as being references to the corresponding ETSI deliverables.

The cross reference between GSM, UMTS, 3GPP and ETSI identities can be found under http://webapp.etsi.org/key/queryform.asp.

Contents

Foreword

This Technical Specification has been produced by the 3rd Generation Partnership Project (3GPP).

The contents of the present document are subject to continuing work within the TSG and may change following formal TSG approval. Should the TSG modify the contents of the present document, it will be re-released by the TSG with an identifying change of release date and an increase in version number as follows:

Version x.y.z

where:

x the first digit:

1 presented to TSG for information;

2 presented to TSG for approval;

3 or greater indicates TSG approved document under change control.

y the second digit is incremented for all changes of substance, i.e. technical enhancements, corrections, updates, etc.

z the third digit is incremented when editorial only changes have been incorporated in the specification;

The 3GPP transparent end-to-end packet-switched streaming service (PSS) specification consists of six 3GPP TSs: 3GPP TS 22.233 [1], 3GPP TS 26.233 [2], 3GPP TS 26.244 [50], 3GPP TS 26.245 [51], 3GPP TS 26.246 [52] and the present document.

The TS 22.233 contains the service requirements for the PSS. The TS 26.233 provides an overview of the PSS. The TS 26.244 defines the 3GPP file format (3GP) used by the PSS and MMS services. The TS 26.245 defines the Timed text format used by the PSS and MMS services. The TS 26.246 defines the 3GPP SMIL language profile. The present document provides the details of the protocols and codecs used by the PSS.

The TS 26.244, TS 26.245 and TS 26.246 start with Release 6. Earlier releases of the 3GPP file format, the Timed text format and the 3GPP SMIL language profile can be found in TS 26.234.

Introduction

Streaming refers to the ability of an application to play synchronised media streams like audio and video streams in a continuous way while those streams are being transmitted to the client over a data network.

Applications, which can be built on top of streaming services, can be classified into on-demand and live information delivery applications. Examples of the first category are music and news-on-demand applications. Live delivery of radio and television programs are examples of the second category.

The 3GPP PSS provides a framework for Internet Protocol (IP) based streaming applications in 3G networks.

1 Scope

The present document specifies the protocols and codecs for the PSS within the 3GPP system. Protocols for control signalling, capability exchange, media transport, rate adaptation and protection are specified. Codecs for speech, natural and synthetic audio, video, still images, bitmap graphics, vector graphics, timed text and text are specified.

The present document is applicable to IP-based packet-switched networks.

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] (void)

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

[6] IETF RFC 4566: "SDP: Session Description Protocol", Handley M., Jacobson V. and Perkins C., July 2006.

[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 4867: "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., April 2007.

[12] (void)

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

[14] IETF RFC 4629: "RTP Payload Format for the ITU-T Rec. H.263 Video", Ott J. et al., January 2007.

[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 – HTTP/1.1", Fielding R. et al., June 1999.

[18] 3GPP TS 26.071: "Mandatory Speech CODEC speech processing functions; AMR Speech CODEC; General description".

[19] (void)

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

[21] ISO/IEC 14496-3:2005: "Information technology – Coding of audio-visual objects – Part 3: Audio".

[22] ITU-T Recommendation H.263 (01/05): "Video coding for low bit rate communication".

[23] (void)

[24] ISO/IEC 14496-2:2004: "Information technology – Coding of audio-visual objects – Part 2: Visual".

[25] (void)

[26] ITU-T Recommendation T.81 (1992) | ISO/IEC 10918-1:1993: "Information technology – Digital compression and coding of continuous-tone still images – 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 – Universal Multiple-Octet Coded Character Set (UCS) – Part 1: Architecture and Basic Multilingual Plane".

[30] IETF RFC 3629: 'UTF-8, a transformation format of ISO 10646', F. Yergeau, November 2003.

[31] W3C Recommendation: "Synchronized Multimedia Integration Language (SMIL 2.0)-[Second Edition]", http://www.w3.org/TR/2005/REC-SMIL2-20050107/, January 2005.

[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] Open Mobile Alliance: "User Agent Profile Version 2.0", February 2006.

[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 Last Call Working Draft: "Scalable Vector Graphics (SVG) 1.2", http://www.w3.org/TR/2004/WD-SVG12-20041027/, October 2004.

[43] W3C Last Call Working Draft: "Mobile SVG Profile: SVG Tiny, Version 1.2", http://www.w3.org/TR/2004/WD-SVGMobile12-20040813/, August 2004.

[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] (void)

[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 4234: "Augmented BNF for Syntax Specifications: ABNF", Crocker D. and Overell P., October 2005.

[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 RFC 4585: "Extended RTP Profile for Real-time Transport Control Protocol (RTCP)-Based Feedback (RTP/AVPF)", Ott J. et al., July 2006.

[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 3986: "Uniform Resource Identifiers (URI): Generic Syntax", Berners-Lee T., Fielding R. and Masinter L., January 2005.

[61] (void)

[62] (void)

[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 4648: "The Base16, Base32, and Base64 Data Encodings", Josefsson S., October 2006.

[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] IETF RFC 3711: "The Secure Real-time Transport Protocol (SRTP)", Baugher M. et al, March 2004.

[73] Bellovin, S., "Problem Areas for the IP Security Protocols" in Proceedings of the Sixth Usenix Unix Security Symposium, pp. 1-16, San Jose, CA, July 1996

[74] Open Mobile Alliance: "DRM Specification 2.0".

[75] Open Mobile Alliance: "DRM Content Format V 2.0".

[76] IETF RFC 3675: "IPv6 Jumbograms", Borman D., Deering S. and Hinden R., August 1999.

[77] NIST, "Advanced Encryption Standard (AES)", FIPS PUB 197, http://www.nist.gov/aes/.

[78] IETF RFC 3394: "Advanced Encryption Standard (AES) Key Wrap Algorithm", Schaad J. and Housley R., September 2002.

[79] IETF RFC 3839: "MIME Type Registrations for 3rd Generation Partnership Project (3GPP) Multimedia files", Castagno R. and Singer D., July 2004.

[80] IETF RFC 4396: "RTP Payload Format for 3rd Generation Partnership Project (3GPP) Timed Text", Rey J. and Matsui Y., February 2006.

[81] IETF RFC 4588: "RTP Retransmission Payload Format", Rey J. et al., July 2006.

[82] 3GPP TS 26.290: "Extended AMR Wideband codec; Transcoding functions".

[83] 3GPP TS 26.304: "ANSI-C code for the Floating-point; Extended AMR Wideband codec".

[84] 3GPP TS 26.273: "ANSI-C code for the Fixed-point; Extended AMR Wideband codec".

[85] IETF RFC 4352: "RTP Payload Format for the Extended Adaptive Multi-Rate Wideband (AMR-WB+) Audio Codec", Sjoberg J. et al., January 2006.

[86] 3GPP TS 26.401: "General audio codec audio processing functions; Enhanced aacPlus general audio codec; General description".

[87] 3GPP TS 26.410: "General audio codec audio processing functions; Enhanced aacPlus general audio codec; Floating-point ANSI-C code".

[88] 3GPP TS 26.411: "General audio codec audio processing functions; Enhanced aacPlus general audio codec; Fixed-point ANSI-C code".

[89] (void)

[90] ITU-T Recommendation H.264 (03/09): "Advanced video coding for generic audiovisual services" | ISO/IEC 14496-10:2009: "Information technology – Coding of audio-visual objects – Part 10: Advanced Video Coding".

[91] (void)

[92] IETF RFC 3984: "RTP Payload Format for H.264 Video", Wenger S. et al, February 2005.

[93] IETF RFC 3890: "A Transport Independent Bandwidth Modifier for the Session Description Protocol (SDP)", Westerlund M., September 2004.

[94] Standard ECMA-327: "ECMAScript 3rd Edition Compact Profile", June 2001.

[95] 3GPP TR 26.936: "Performance characterization of 3GPP audio codecs".

[96] IETF RFC 4288: "Media Type Specifications and Registration Procedures", Freed N. and Klensin J., December 2005.

[97] IETF RFC 4613: "Media Type Registrations for Downloadable Sounds for Musical Instrument Digital Interface (MIDI)", Frojdh P., Lindgren U. and Westerlund M., September 2006.

[98] 3GPP TS 26.142: "Dynamic and Interactive Multimedia Scene".

[99] OMA-ERELD-DM-V1_2-20070209-A: "Enabler Release Definition for OMA Device Management, Approved Version 1.2".

[100] IETF RFC 1123: "Requirements for Internet Hosts -- Application and Support", Braden R., October 1989.

[101] Internet Streaming Media Alliance (ISMA), ISMA Encryption and Authentication, Version 1.1 release version, September 2006.

[102] Internet Streaming Media Alliance (ISMA), ISMA Encryption and Authentication, Version 2.0 release version, November 2007.

[103] Open Mobile Alliance, Service and Content Protection for Mobile Broadcast Services, Approved Version 1.0, February 2009.

[104] ISO/IEC 14496-12:2005 | 15444-12:2005: "Information technology – Coding of audio-visual objects – Part 12: ISO base media file format" | "Information technology – JPEG 2000 image coding system – Part 12: ISO base media file format".

[105] RFC 2818: HTTP Over TLS, E. Rescorla, May 2000.

[106] RFC 5646: Tags for Identifying Languages, A. Phillips, M. Davis, September 2009.

[107] RFC 4281: The Codecs Parameter for "Bucket" Media Types, R. Gellens, D. Singer, P. Frojdh, November 2005.

[108] Open Mobile Alliance: "DRM Content Format V 2.0".

[109] 3GPP TS 26.430: "Timed Graphics".

3 Definitions and abbreviations

3.1 Definitions

For the purposes of the present document, the following terms and definitions apply:

continuous media: media with an inherent notion of time. In the present document speech, audio, video, timed text and DIMS

discrete media: media that itself does not contain an element of time. In the present document all media not defined as continuous media

device capability description: a description of device capabilities and/or user preferences. Contains a number of capability attributes

device capability profile: same as device capability description

kilobits: 1000 bits

kilobytes: 1024 bytes

presentation description: contains information about one or more media streams within a presentation, such as the set of encodings, network addresses and information about the content

PSS client: client for the 3GPP packet switched streaming service based on the IETF RTSP/SDP and/or HTTP standards, with possible additional 3GPP requirements according to the present document

PSS server: server for the 3GPP packet switched streaming service based on the IETF RTSP/SDP and/or HTTP standards, with possible additional 3GPP requirements according to the present document

scene description: description of the spatial layout and temporal behaviour of a presentation. It can also contain hyperlinks

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
ADU	Application Data Unit
AVC	Advanced Video Coding
CC/PP	Composite Capability / Preference Profiles
DCT	Discrete Cosine Transform
DIMS	Dynamic and Interactive Multimedia Scenes
DLS	Downloadable Sounds
DRM	Digital Rights Management
Enhanced aacPlus	MPEG-4 High Efficiency AAC plus MPEG-4 Parametric Stereo
GIF	Graphics Interchange Format
HTML	Hyper Text Markup Language
ITU-T	International Telecommunications Union – Telecommunications
JFIF	JPEG File Interchange Format
MIDI	Musical Instrument Digital Interface
MIME	Multipurpose Internet Mail Extensions
MMS	Multimedia Messaging Service
NADU	Next Application Data Unit
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
SBR	Spectral Band Replication
SDP	Session Description Protocol
SMIL	Synchronised Multimedia Integration Language
SP-MIDI	Scalable Polyphony MIDI
SRTP	The Secure Real-time Transport Protocol
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

4 System description

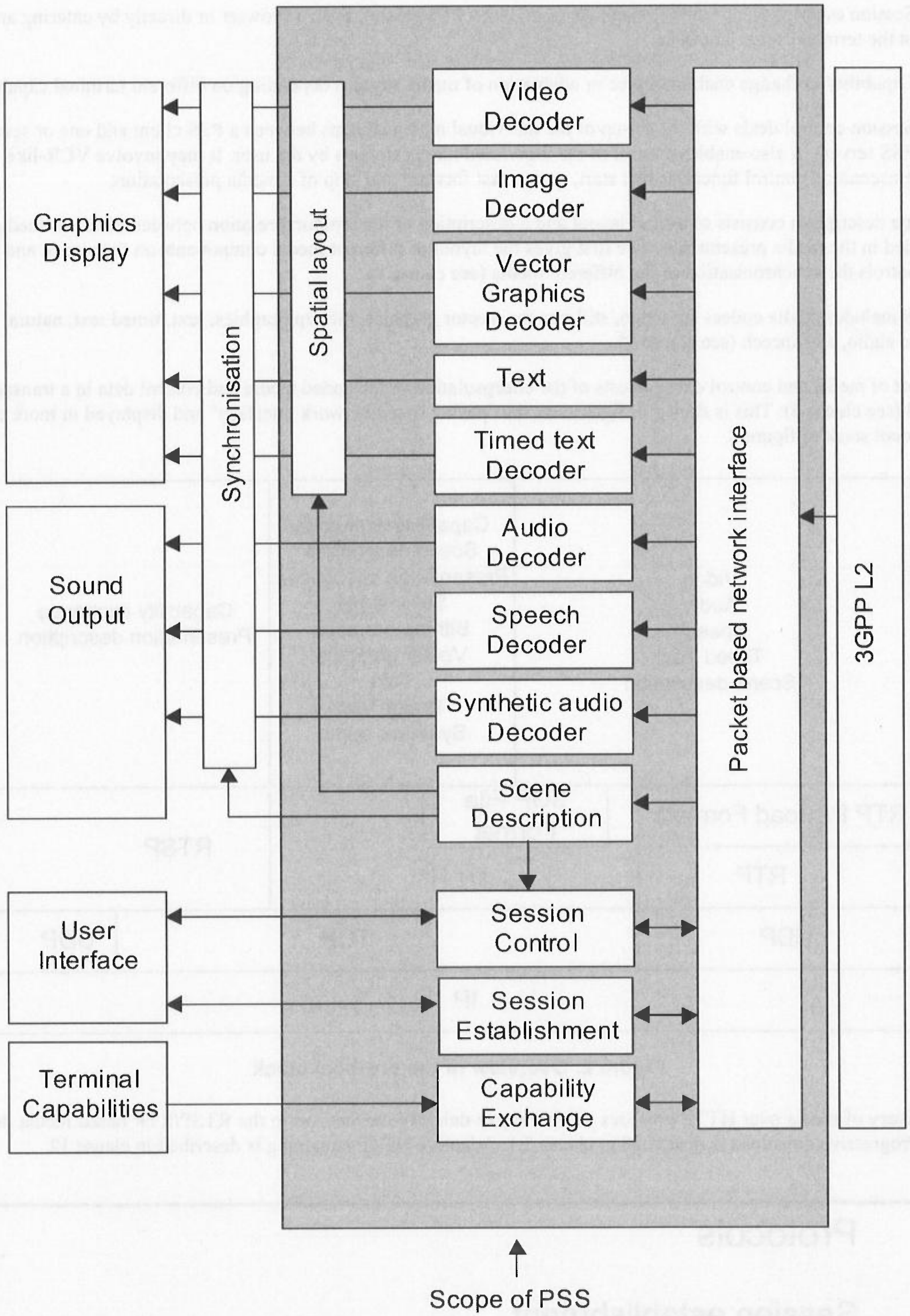


Figure 1: Functional components of a PSS client

Figure 1 shows the functional components of a PSS client. Figure 2 gives an overview of the protocol stack used in a PSS client and also shows a more detailed view of the packet based network interface. The functional components can be divided into control, scene description, media codecs and the transport of media and control data.

The control related elements are session establishment, capability exchange and session control (see clause 5).

- Session establishment refers to methods to invoke a PSS session from a browser or directly by entering an URL in the terminal's user interface.
- Capability exchange enables choice or adaptation of media streams depending on different terminal capabilities.
- Session control deals with the set-up of the individual media streams between a PSS client and one or several PSS servers. It also enables control of the individual media streams by the user. It may involve VCR-like presentation control functions like start, pause, fast forward and stop of a media presentation.

The scene description consists of spatial layout and a description of the temporal relation between different media that is included in the media presentation. The first gives the layout of different media components on the screen and the latter controls the synchronisation of the different media (see clause 8).

The PSS includes media codecs for video, still images, vector graphics, bitmap graphics, text, timed text, natural and synthetic audio, and speech (see clause 7).

Transport of media and control data consists of the encapsulation of the coded media and control data in a transport protocol (see clause 6). This is shown in figure 1 as the "packet based network interface" and displayed in more detail in the protocol stack of figure 2.

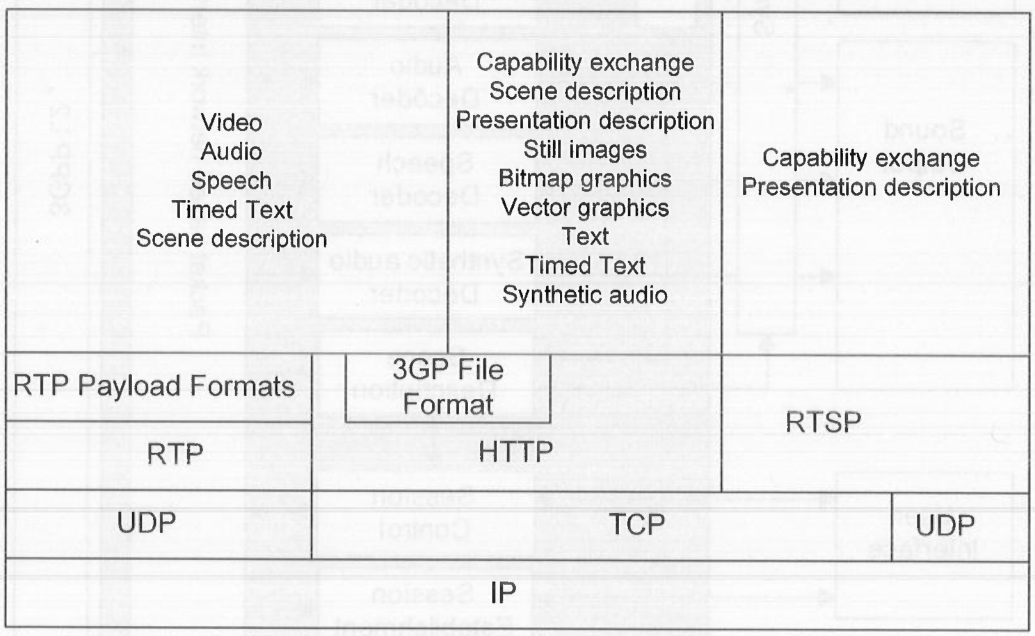


Figure 2: Overview of the protocol stack

The delivery of media over HTTP provides an alternative delivery mechanism to the RTSP/RTP based media delivery. HTTP Progressive download is described in clause 5.1. Adaptive HTTP streaming is described in clause 12.

5 Protocols

5.1 Session establishment

Session establishment refers to the method by which a PSS client obtains the initial session description. The initial session description can e.g. be a presentation description, a scene description or just an URL to the content.

A PSS client shall support initial session descriptions specified in one of the following formats: SMIL, SDP, or plain RTSP URL.

In addition to rtsp:// the PSS client shall support URLs [60] to valid initial session descriptions starting with file:// (for locally stored files) and http:// (for presentation descriptions or scene descriptions delivered via HTTP).

Examples for valid inputs to a PSS client are: file://temp/morning_news.smil, http://example.com/morning_news.sdp, and rtsp://example.com/morning_news.

URLs can be made available to a PSS client in many different ways. It is out of the scope of this specification to mandate any specific mechanism. However, an application using the 3GPP PSS shall at least support URLs of the above type, specified or selected by the user.

The preferred way would be to embed URLs to initial session descriptions within HTML or WML pages. Browser applications that support the HTTP protocol could then download the initial session description and pass the content to the PSS client for further processing. How exactly this is done is an implementation specific issue and out of the scope of this specification.

As an alternative to conventional streaming, a PSS client should also support progressive download of 3GP files [50] delivered via HTTP. A progressive-download session is established with one or more HTTP GET requests. In order to improve playback performance for 3GP files that are not authored for progressive download, a PSS client may issue (multiple pipelined) HTTP GET requests with byte ranges [17]. Example of a valid URL is http://example.com/morning_news.3gp.

5.2 Capability exchange

5.2.1 General

Capability exchange is an important functionality in the PSS. It enables PSS servers to provide a wide range of devices with content suitable for the particular device in question. Another very important task is to provide a smooth transition between different releases of PSS. Therefore, PSS clients and servers should support capability exchange.

The specification of capability exchange for PSS is divided into two parts. The normative part contained in clause 5.2 and an informative part in clause A.4 in Annex A of the present document. The normative part gives all the necessary requirements that a client or server shall conform to when implementing capability exchange in the PSS. The informative part provides additional important information for understanding the concept and usage of the functionality. It is recommended to read clause A.4 in Annex A before continuing with clauses 5.2.2-5.2.7.

5.2.2 The device capability profile structure

A device capability profile is an RDF [41] document that follows the structure of the CC/PP framework [39] and the CC/PP application UAProf [40]. Attributes are used to specify device capabilities and preferences. A set of attribute names, permissible values and semantics constitute a CC/PP vocabulary, which is defined by an RDF schema. For PSS, the UAProf vocabulary is reused and an additional PSS specific vocabulary is defined. The details can be found in clause 5.2.3. The syntax of the attributes is defined in the vocabulary schema, but also, to some extent, the semantics. A PSS device capability profile is an instance of the schema (UAProf and/or the PSS specific schema) and shall follow the rules governing the formation of a profile given in the CC/PP specification [39]. The profile schema shall also be governed by the rules defined in UAProf [40] chapter 7, 7.1, 7.3 and 7.4.

5.2.3 Vocabularies for PSS

5.2.3.1 General

Clause 5.2.3 specifies the attribute vocabularies to be used by the PSS capability exchange.

PSS servers supporting capability exchange shall support the attributes in the four PSS components of the PSS base vocabulary. PSS servers should also support the recommended attributes from the UAProf vocabulary [40]. A server may additionally support other UAProf attributes.

5.2.3.2 PSS base vocabulary

The PSS base vocabulary contains four components called "PssCommon", "Streaming", "ThreeGPFileFormat" and "PssSmil". The division of the vocabulary into these components is motivated by the fact that the PSS contains three different base applications:

- pure RTSP/RTP-based streaming (described by the Streaming component);
- 3GP file download or progressive download (described by the ThreeGPFileFormat component);
- SMIL presentation (described by the PssSmil component).

The last application can consist of downloadable images, text, etc., as well as RTSP/RTP streaming and downloadable 3GP files. Capabilities that are common to all PSS applications are described by the PssCommon component. The three base applications are distinguished from each other by the source of synchronization: for pure streaming it is RTP, for 3GP files it is inherit in the 3GP file format, and for SMIL presentations timing is provided by the SMIL file.

NOTE: DIMS presentations can be streamed (RTSP/RTP) or downloaded (as 3GP files). Capabilities for such presentations are described by the Streaming component and the ThreeGPFileFormat component, respectively. See in particular the StreamingAccept and ThreeGPAccept attributes below.

The PSS base vocabulary is an extension to UAProf and is defined as an RDF schema in Annex F. Together with the description of the attributes in the present clause, it defines the vocabulary. The vocabulary is associated with an XML namespace, which combines a base URI with a local XML element name to yield an URI. Annex F provides the details.

The PSS specific components contain a number of attributes expressing capabilities. The following subclauses list all attributes for each component.

5.2.3.2.1 PssCommon component

Attribute name: **AudioChannels**

Attribute definition: This attribute describes the stereophonic capability of the natural audio device.

Component: PssCommon

Type: Literal

Legal values: 'Mono', 'Stereo'

Resolution rule: Locked

EXAMPLE 1: `<AudioChannels>Mono</AudioChannels>`

Attribute name: **MaxPolyphony**

Attribute definition: The MaxPolyphony attribute refers to the maximal polyphony that the synthetic audio device supports as defined in [44].

NOTE: The MaxPolyphony attribute is used to signal the maximum polyphony capabilities supported by the PSS client. This is a complementary mechanism for the delivery of compatible SP-MIDI content and thus by setting the MaxPolyphony attribute the PSS client is required to support Scalable Polyphony MIDI i.e. Channel Masking defined in [44].

Component: PssCommon

Type: Number

Legal values: Integer between 5 and 24

Resolution rule: Locked

EXAMPLE 2: `<MaxPolyphony>8</MaxPolyphony>`

Attribute name: **NumOfGM1Voices**

Attribute definition: The NumOfGM1Voices attribute refers to the maximum number of simultaneous GM1 voices that the synthetic audio engine supports.

Component: PssCommon

Type: Number

Legal values: Integer greater or equal than 5

Resolution rule: Locked

EXAMPLE 3: `<NumOfGMIVoices>24</NumOfGMIVoices>`

Attribute name: **NumOfMobileDLSVoicesWithoutOptionalBlocks**

Attribute definition: The NumOfMobileDLSVoicesWithoutOptionalBlocks attribute refers to the maximum number of simultaneous Mobile DLS [70] voices without optional group of processing blocks that the synthetic audio engine supports.

Component: PssCommon

Type: Number

Legal values: Integer greater or equal than 5

Resolution rule: Locked

EXAMPLE 4:
```
<NumOfMobileDLSVoicesWithoutOptionalBlocks>24
</NumOfMobileDLSVoicesWithoutOptionalBlocks>
```

Attribute name: **NumOfMobileDLSVoicesWithOptionalBlocks**

Attribute definition: The NumOfMobileDLSVoicesWithOptionalBlocks attribute refers to the maximum number of simultaneous Mobile DLS voices with optional group of processing blocks that the synthetic audio engine supports. This attribute is set to zero for devices that do not support the optional group of processing blocks.

Component: PssCommon

Type: Number

Legal values: Integer greater than or equal to 0

Resolution rule: Locked

EXAMPLE 5:
```
<NumOfMobileDLSVoicesWithOptionalBlocks>24
</NumOfMobileDLSVoicesWithOptionalBlocks>
```

Attribute name: **PssVersion**

Attribute definition: Latest PSS version supported by the client.

Component: PssCommon

Type: Literal

Legal values: "3GPP-R4", "3GPP-R5", "3GPP-R6", "3GPP-R7" and so forth.

Resolution rule: Locked

EXAMPLE 6: `<PssVersion>3GPP-R6</PssVersion>`

Attribute name: **RenderingScreenSize**

Attribute definition: The rendering size of the device"s screen in unit of pixels available for PSS media presentation. The horizontal size is given followed by the vertical size.

Component: PssCommon

Type: Dimension

Legal values: Two integer values equal or greater than zero. A value equal '0x0'means that there exists no possibility to render visual PSS presentations.

Resolution rule: Locked

EXAMPLE 7: `<RenderingScreenSize>70x15</RenderingScreenSize>`

5.2.3.2.2 Streaming component

Attribute name: **StreamingAccept**

Attribute definition: List of content types (MIME types) relevant for streaming over RTP supported by the PSS application. Content types listed shall be possible to stream over RTP. For each content type a set of MIME parameters can be specified to signal receiver capabilities. A content type that supports multiple parameter sets may occur several times in the list.

Component: Streaming

Type: Literal (Bag)

Legal values: List of MIME types with related parameters.

Resolution rule: Append

EXAMPLE 1:

```
<StreamingAccept>
  <rdf:Bag>
    <rdf:li>audio/AMR-WB; octet-alignment=1</rdf:li>
    <rdf:li>video/H263-2000; profile=0; level=45</rdf:li>
  </rdf:Bag>
</StreamingAccept>
```

EXAMPLE 1b:

```
<StreamingAccept>
  <rdf:Bag>
    <rdf:li>audio/AMR-WB+</rdf:li>
    <rdf:li>video/H264; profile-level-id=42e00a</rdf:li>
    <rdf:li>video/richmedia+xml; Version-profile=10</rdf:li>
  </rdf:Bag>
</StreamingAccept>
```

Attribute name: **StreamingAccept-Subset**

Attribute definition: List of content types for which the PSS application supports a subset. MIME types can in most cases effectively be used to express variations in support for different media types. Many MIME types, e.g. AMR-WB have several parameters that can be used for this purpose. There may exist content types for which the PSS application only supports a subset and this subset cannot be expressed with MIME-type parameters. In these cases the attribute StreamingAccept-Subset is used to describe support for a subset of a specific content type. If a subset of a specific content type is declared in StreamingAccept-Subset, this means that StreamingAccept-Subset has precedence over StreamingAccept. StreamingAccept shall always include the corresponding content types for which StreamingAccept-Subset specifies subsets of.

Subset identifiers and corresponding semantics shall only be defined by the TSG responsible for the present document.

Component: Streaming

Type: Literal (Bag)

Legal values: No subsets defined.

Resolution rule: Append

Attribute name: **ThreeGPPLinkChar**

Attribute definition: Indicates whether the device supports the 3GPP-Link-Char header according to clause 10.2.1.1.

Component: Streaming

Type: Literal

Legal values: "Yes", "No"

Resolution rule: Override

EXAMPLE 2: `<ThreeGPPLinkChar>Yes</ThreeGPPLinkChar>`

Attribute name: **AdaptationSupport**

Attribute definition: Indicates whether the device supports client buffer feedback signaling according to clause 10.2.3.

Component: Streaming

Type: Literal

Legal values: "Yes", "No"

Resolution rule: Locked

EXAMPLE 3: `<AdaptationSupport>Yes</AdaptationSupport>`

Attribute name: **QoESupport**

Attribute definition: Indicates whether the device supports QoE signaling according to clauses 5.3.2.3, 5.3.3.6, and 11.

Component: Streaming

Type: Literal

Legal values: "Yes", "No"

Resolution rule: Locked

EXAMPLE 3: `<QoESupport>Yes</QoESupport>`

Attribute name: **ExtendedRtcpReports**

Attribute definition: Indicates whether the device supports extended RTCP reports according to clause 6.2.3.1.

Component: Streaming

Type: Literal

Legal values: "Yes", "No"

Resolution rule: Locked

EXAMPLE 4: `<ExtendedRtcpReports>Yes</ExtendedRtcpReports>`

Attribute name: **RtpRetransmission**

Attribute definition: Indicates whether the device supports RTP retransmission according to clause 6.2.3.3.

Component: Streaming

Type: Literal

Legal values: "Yes", "No"

Resolution rule: Locked

EXAMPLE 5: `<RtpRetransmission>Yes</RtpRetransmission>`

Attribute name: **MediaAlternatives**

Attribute definition: Indicates whether the device interprets the SDP attributes "alt", "alt-default-id", and "alt-group", defined in clauses 5.3.3.3 and 5.3.3.4.

Component: Streaming

Type: Literal

Legal values: "Yes", "No"

Resolution rule: Override

EXAMPLE 6: `<MediaAlternatives>Yes</MediaAlternatives>`

Attribute name: **RtpProfiles**

Attribute definition: List of supported RTP profiles.

Component: Streaming

Type: Literal (Bag)

Legal values: Profile names registered through the Internet Assigned Numbers Authority (IANA), www.iana.org.

Resolution rule: Append

EXAMPLE 7:

```
<RtpProfiles>
  <rdf:Bag>
    <rdf:li>RTP/AVP</rdf:li>
    <rdf:li>RTP/AVPF</rdf:li>
  </rdf:Bag>
</RtpProfiles>
```

Attribute name: **StreamingOmaDrm**

Attribute definition: Indicates whether the device supports streamed OMA DRM protected content as defined by OMA and Annex K.

Component: Streaming

Type: Literal (Bag)

Legal values: OMA Version numbers supported as a floating number. 0.0 indicates no support.

Resolution rule: Locked

EXAMPLE 8:
```
<StreamingOmaDrm>
  <rdf:Bag>
    <rdf:li>2.0</rdf:li>
  </rdf:Bag>
</StreamingOmaDrm>
```

Attribute name: **PSSIntegrity**

Attribute definition: Indicates whether the device supports integrity protection for streamed content as defined by Annex K.2.

Component: Streaming

Type: Literal

Legal values: "Yes", "No"

Resolution rule: Locked

EXAMPLE 9: `<PSSIntegrity>Yes</PSSIntegrity>`

Attribute name: **VideoDecodingByteRate**

Attribute definition: If Annex G is not supported, the attribute has no meaning. If Annex G is supported, this attribute defines the peak decoding byte rate the PSS client is able to support. In other words, the PSS client fulfils the requirements given in Annex G with the signalled peak decoding byte rate. The values are given in bytes per second and shall be greater than or equal to 16000. According to Annex G, 16000 is the default peak decoding byte rate for the mandatory video codec profile and level (H.263 Profile 0 Level 45).

Component: Streaming

Type: Number

Legal values: Integer value greater than or equal to 16000.

Resolution rule: Locked

EXAMPLE 10: `<VideoDecodingByteRate>16000</VideoDecodingByteRate>`

Attribute name: **VideoInitialPostDecoderBufferingPeriod**

Attribute definition: If Annex G is not supported, the attribute has no meaning. If Annex G is supported, this attribute defines the maximum initial post-decoder buffering period of video. Values are interpreted as clock ticks of a 90-kHz clock. In other words, the value is incremented by one for each 1/90 000 seconds. For example, the value 9000 corresponds to 1/10 of a second initial post-decoder buffering.

Component: Streaming

Type: Number

Legal values: Integer value equal to or greater than zero.

Resolution rule: Locked

EXAMPLE 11:
```
<VideoInitialPostDecoderBufferingPeriod>9000
</VideoInitialPostDecoderBufferingPeriod>
```

Attribute name: **VideoPreDecoderBufferSize**

Attribute definition: This attribute signals if the optional video buffering requirements defined in Annex G are supported. It also defines the size of the hypothetical pre-decoder buffer defined in Annex G. A value equal to zero means that Annex G is not supported. A value equal to one means that Annex G is supported. In this case the size of the buffer is the default size defined in Annex G. A value equal to or greater than the default buffer size defined in Annex G means that Annex G is supported and sets the buffer size to the given number of octets.

Component: Streaming

Type: Number

Legal values: Integer value equal to or greater than zero. Values greater than one but less than the default buffer size defined in Annex G are not allowed.

Resolution rule: Locked

EXAMPLE 12:
```
<VideoPreDecoderBufferSize>30720</VideoPreDecoderBufferSize>
```

5.2.3.2.3 ThreeGPFileFormat component

Attribute name: **Brands**

Attribute definition: List of supported 3GP profiles identified by brand.

Component: ThreeGPFileFormat

Type: Literal (Bag)

Legal values: Brand identifiers according to 5.3.4 and 5.4 in [50].

Resolution rule: Append

EXAMPLE 1:
```
<Brands>
  <rdf:Bag>
    <rdf:li>3gp4</rdf:li>
    <rdf:li>3gp5</rdf:li>
    <rdf:li>3gp6</rdf:li>
    <rdf:li>3gr6</rdf:li>
    <rfd:li>3gp7</rdf:li>
    <rfd:li>3gr7</rdf:li>
    <rfd:li>3ge7</rdf:li>
  </rdf:Bag>
</Brands>
```

Attribute name: **ThreeGPAccept**

Attribute definition: List of content types (MIME types) that can be included in a 3GP file and handled by the PSS application. The content types included in this attribute can be rendered in a 3GP file or a presentation contained therein. If the identifier "Streaming-Media" is included, streaming media can be included within a contained presentation (e.g. in DIMS). Details on the streaming support can then be found in the Streaming component. For each content type a set of supported parameters can be given. A content type that supports multiple parameter sets may occur several times in the list.

Component: ThreeGPFileFormat

Type: Literal (Bag)

Legal values: List of MIME types with related parameters and the "Streaming-Media" identifier.

Resolution rule: Append

EXAMPLE 2:
```
<ThreeGPAccept>
  <rdf:Bag>
    <rdf:li>video/H263-2000; profile=0; level=45</rdf:li>
    <rdf:li>audio/AMR</rdf:li>
  </rdf:Bag>
</ThreeGPAccept>
```

EXAMPLE 2b:
```
<ThreeGPAccept>
  <rdf:Bag>
    <rdf:li>audio/AMR</rdf:li>
    <rdf:li>audio/AMR-WB+</rdf:li>
    <rdf:li>video/H263-2000; profile=0; level=45</rdf:li>
    <rdf:li>video/H264; profile-level-id=42e00a</rdf:li>
    <rdf:li>image/jpeg</rdf:li>
    <rdf:li>video/richmedia+xml; Version-profile=10</rdf:li>
    <rdf:li>Streaming-Media</rdf:li>
  </rdf:Bag>
</ThreeGPAccept>
```

Attribute name: **ThreeGPAccept-Subset**

Attribute definition: List of content types for which the PSS application supports a subset. MIME types can in most cases effectively be used to express variations in support for different media types. Many MIME types have several parameters that can be used for this purpose. There may exist content types for which the PSS application only supports a subset and this subset cannot be expressed with MIME-type parameters. In these cases the attribute ThreeGPAccept-Subset is used to describe support for a subset of a specific content type. If a subset of a specific content type is declared in ThreeGPAccept-Subset, this means that ThreeGPAccept-Subset has precedence over ThreeGPAccept. ThreeGPAccept shall always include the corresponding content types for which ThreeGPAccept-Subset specifies subsets of.

Subset identifiers and corresponding semantics shall only be defined by the TSG responsible for the present document.

Component: ThreeGPFileFormat

Type: Literal (Bag)

Legal values: No subsets defined.

Resolution rule: Append

Attribute name: **ThreeGPOmaDrm**

Attribute definition: List of the OMA DRM versions that is supported to be used for DRM protection of content present in the 3GP file format.

Component: ThreeGPFileFormat

Type: Literal (Bag)

Legal values: OMA DRM version numbers as floating point values. 0.0 indicates no support.

Resolution rule: Locked

EXAMPLE 3:
```
<3gpOMADRM>
  <rdf:Bag>
    <rdf:li>2.0 </rdf:li>
```

```
        </rdf:Bag>
      </3gpOMADRM>
```

5.2.3.2.4 PssSmil component

Attribute name: **SmilAccept**

Attribute definition: List of content types (MIME types) that can be part of a SMIL presentation. The content types included in this attribute can be rendered in a SMIL presentation. If video/3gpp (or audio/3gpp) is included, downloaded 3GP files can be included in a SMIL presentation. Details on the 3GP file support can then be found in the ThreeGPFileFormat component. If the identifier "Streaming-Media" is included, streaming media can be included in the SMIL presentation. Details on the streaming support can then be found in the Streaming component. For each content type a set of supported parameters can be given. A content type that supports multiple parameter sets may occur several times in the list.

Component: PssSmil

Type: Literal (Bag)

Legal values: List of MIME types with related parameters and the "Streaming-Media" identifier.

Resolution rule: Append

EXAMPLE 1:
```
<SmilAccept>
  <rdf:Bag>
    <rdf:li>image/gif</rdf:li>
    <rdf:li>image/jpeg</rdf:li>
    <rdf:li>Streaming-Media</rdf:li>
  </rdf:Bag>
</SmilAccept>
```

Attribute name: **SmilAccept-Subset**

Attribute definition: List of content types for which the PSS application supports a subset. MIME types can in most cases effectively be used to express variations in support for different media types. Many MIME types have several parameters that can be used for this purpose. There may exist content types for which the PSS application only supports a subset and this subset cannot be expressed with MIME-type parameters. In these cases the attribute SmilAccept-Subset is used to describe support for a subset of a specific content type. If a subset of a specific content type is declared in SmilAccept-Subset, this means that SmilAccept-Subset has precedence over SmilAccept. SmilAccept shall always include the corresponding content types for which SmilAccept-Subset specifies subsets of.

The following values are defined:

- "JPEG-PSS": Only the two JPEG modes described in clause 7.5 of the present document are supported.
- "SVG-Tiny"
- "SVG-Basic"

Subset identifiers and corresponding semantics shall only be defined by the TSG responsible for the present document.

Component: PssSmil

Type: Literal (Bag)

Legal values: "JPEG-PSS", "SVG-Tiny", "SVG-Basic"

Resolution rule: Append

EXAMPLE 2:
```
<SmilAccept-Subset>
  <rdf:Bag>
    <rdf:li>JPEG-PSS</rdf:li>
    <rdf:li>SVG-Tiny</rdf:li>
  </rdf:Bag>
</SmilAccept-Subset>
```

Attribute name: **SmilBaseSet**

Attribute definition: Indicates a base set of SMIL 2.0 modules that the client supports.

Component: Streaming

Type: Literal

Legal values: Pre-defined identifiers. "SMIL-3GPP-R4" and "SMIL-3GPP-R5" indicate all SMIL 2.0 modules required for SMIL scene description support according to clause 8 of Release 4 and Release 5, respectively, of TS 26.234. "SMIL-3GPP-R6"and "SMIL-3GPP-R7" indicate all SMIL 2.0 modules required for SMIL scene-description support according to Release 6 and Release 7, respectively, of clause 8 of the present document (TS 26.234) and of TS 26.246 [52].

Resolution rule: Locked

EXAMPLE 3:
```
<SmilBaseSet>SMIL-3GPP-R6</SmilBaseSet>
```

Attribute name: **SmilModules**

Attribute definition: This attribute defines a list of SMIL 2.0 modules supported by the client. If the SmilBaseSet is used those modules do not need to be explicitly listed here. In that case only additional module support needs to be listed.

Component: Streaming

Type: Literal (Bag)

Legal values: SMIL 2.0 module names defined in the SMIL 2.0 recommendation [31], section 2.3.3, table 2.

Resolution rule: Append

EXAMPLE 4:
```
<SmilModules>
  <rdf:Bag>
    <rdf:li>BasicTransitions</rdf:li>
    <rdf:li>MulitArcTiming</rdf:li>
  </rdf:Bag>
</SmilModules>
```

5.2.3.3 Attributes from UAProf

In the UAProf vocabulary [40] there are several attributes that are of interest for the PSS. The formal definition of these attributes is given in [40]. The following list of attributes is recommended for PSS applications:

Attribute name: **BitsPerPixel**

Component: HardwarePlatform

Attribute description: The number of bits of colour or greyscale information per pixel

EXAMPLE 1:
```
<BitsPerPixel>8</BitsPerPixel>
```

Attribute name: **ColorCapable**

Component: HardwarePlatform

Attribute description: Whether the device display supports colour or not.

EXAMPLE 2: `<ColorCapable>Yes</ColorCapable>`

Attribute name: **PixelAspectRatio**

Component: HardwarePlatform

Attribute description: Ratio of pixel width to pixel height

EXAMPLE 3: `<PixelAspectRatio>1x2</PixelAspectRatio>`

Attribute name: **PointingResolution**

Component: HardwarePlatform

Attribute description: Type of resolution of the pointing accessory supported by the device.

EXAMPLE 4: `<PointingResolution>Pixel</PointingResolution>`

Attribute name: **Model**

Component: HardwarePlatform

Attribute description: Model number assigned to the terminal device by the vendor or manufacturer

EXAMPLE 5: `<Model>Model B</Model>`

Attribute name: **Vendor**

Component: HardwarePlatform

Attribute description: Name of the vendor manufacturing the terminal device

EXAMPLE 6: `<Vendor>TerminalManufacturer A</Vendor>`

Attribute name: **CcppAccept-Charset**

Component: SoftwarePlatform

Attribute description: List of character sets the device supports

EXAMPLE 7:

```
<CcppAccept-Charset>
  <rdf:Bag>
    <rdf:li>UTF-8</rdf:li>
  </rdf:Bag>
</CcppAccept-Charset>
```

Attribute name: **CcppAccept-Encoding**

Component: SoftwarePlatform

Attribute description: List of transfer encodings the device supports

EXAMPLE 8:
```
<CcppAccept-Encoding>
  <rdf:Bag>
    <rdf:li>base64</rdf:li>
  </rdf:Bag>
</CcppAccept-Encoding>
```

Attribute name: **CcppAccept-Language**

Component: SoftwarePlatform

Attribute description: List of preferred document languages

EXAMPLE 9:
```
<CcppAccept-Language>
  <rdf:Seq>
    <rdf:li>en</rdf:li>
    <rdf:li>se</rdf:li>
  </rdf:Seq>
</CcppAccept-Language>
```

5.2.4 Extensions to the PSS schema/vocabulary

5.2.4.1 Vocabulary definitions

The use of RDF enables an extensibility mechanism for CC/PP-based schemas that addresses the evolution of new types of devices and applications. The Release-6 PSS profile schema specification has been updated from Release 5 and has thus been assigned a unique RDF schema. The same is true for the Release-7 PSS profile schema specification. The following URIs uniquely identify the RDF schemas for Release 5, Release 6 and Release 7:

PSS Release 5 URI: http://www.3gpp.org/profiles/PSS/ccppschema-PSS5#

PSS Release 6 URI: http://www.3gpp.org/profiles/PSS/ccppschema-PSS6#

PSS Release 7 URI: http://www.3gpp.org/profiles/PSS/ccppschema-PSS7#

In the future new usage scenarios might have need for expressing new attributes. If the base vocabulary is further updated, a new unique namespace will be assigned to the updated schema. The base vocabulary shall only be changed by the TSG responsible for the present document. All extensions to the profile schema shall be governed by the rules defined in [40] clause 7.7.

5.2.4.2 Backward compatibility

An important issue when introducing a new vocabulary is to ensure backward compatibility. PSS Release-6 clients should seamlessly work together with PSS Release-5 servers and vice versa. To obtain backward compatibility, a Release-6 client should provide servers with multiple device-capability profiles using PSS Release-5 and Release-6 vocabularies, respectively. This can be done by providing two URIs referring to two separate profiles or one URI referring to one combined profile that uses both the Relase-5 and the Release-6 namespaces. PSS Release-6 servers should handle both namespaces, whereas PSS Release-5 servers will ignore profiles with unknown namespaces.

5.2.5 Signalling of profile information between client and server

When a PSS client or server support capability exchange it shall support the profile information transport over both HTTP and RTSP between client and server as defined in clause 9.1 (including its subsections) of the WAP 2.0 UAProf specification [40] with the following amendments:

- The "x-wap-profile" and "x-wap-profile-diff" headers shall be present in at least one HTTP or RTSP request per session. That is, the requirement to send this header in all requests has been relaxed.
- The defined headers may be applied to both RTSP and HTTP.
- The "x-wap-profile-diff" header is only valid for the current request. The reason is that PSS does not have the WSP session concept of WAP.
- Push is not relevant for the PSS.

The following guidelines concern how and when profile information is sent between client and server:

- PSS content servers supporting capability exchange shall be able to receive profile information in all HTTP and RTSP requests.
- The terminal should not send the "x-wap-profile-diff" header over the air-interface since there is no compression scheme defined.
- RTSP: the client should send profile information in the DESCRIBE message. It may send it in any other request.

If the terminal has some prior knowledge about the file type it is about to retrieve, e.g. file extensions, the following apply:

- HTTP and SDP: when retrieving an SDP with HTTP the client should include profile information in the GET request. This way the HTTP server can deliver an optimised SDP to the client.
- HTTP and SMIL: When retrieving a SMIL file with HTTP the client should include profile information in the GET request. This way the HTTP server can deliver an optimised SMIL presentation to the client. A SMIL presentation can include links to static media. The server should optimise the SMIL file so that links to the referenced static media are adapted to the requesting client. When the "x-wap-profile-warning" indicates that content selection has been applied (201-203) the PSS client should assume that no more capability exchange has to be performed for the static media components. In this case it should not send any profile information when retrieving static media to be included in the SMIL presentation. This will minimise the HTTP header overhead.

5.2.6 Merging device capability profiles

Profiles need to be merged whenever the PSS server receives multiple device capability profiles. Multiple occurrences of attributes and default values make it necessary to resolve the profiles according to a resolution process.

The resolution process shall be the same as defined in UAProf [40] clause 6.4.1.

- Resolve all indirect references by retrieving URI references contained within the profile.
- Resolve each profile and profile-diff document by first applying attribute values contained in the default URI references and by second applying overriding attribute values contained within the category blocks of that profile or profile-diff.
- Determine the final value of the attributes by applying the resolved attribute values from each profile and profile-diff in order, with the attribute values determined by the resolution rules provided in the schema. Where no resolution rules are provided for a particular attribute in the schema, values provided in profiles or profile-diffs are assumed to override values provided in previous profiles or profile-diffs.

When several URLs are defined in the "x-wap-profile" header and there exists any attribute that occurs more than once in these profiles the rule is that the attribute value in the second URL overrides, or is overridden by, or is appended to the attribute value from the first URL (according to the resolution rule) and so forth. This is what is meant with "Determine the final value of the attributes by applying the resolved attribute values from each profile and profile-diff in order, with…" in the third bullet above. If the profile is completely or partly inaccessible or otherwise corrupted the server should still provide content to the client. The server is responsible for delivering content optimised for the client based on the received profile in a best effort manner.

NOTE: For the reasons explained in Annex A clause A.4.3 the usage of indirect references in profiles (using the CC/PP defaults element) is not recommended.

5.2.7 Profile transfer between the PSS server and the device profile server

The device capability profiles are stored on a device profile server and referenced with URLs. According to the profile resolution process in clause 5.2.6 of the present document, the PSS server ends up with a number of URLs referring to profiles and these shall be retrieved.

- The device profile server shall support HTTP 1.1 for the transfer of device capability profiles to the PSS server.

- If the PSS server supports capability exchange it shall support HTTP 1.1 for transfer of device capability profiles from the device profile server. A URL shall be used to identify a device capability profile.

- Normal content caching provisions as defined by HTTP apply.

5.3 Session set-up and control

5.3.1 General

Continuous media is media that has an intrinsic time line. Discrete media on the other hand does not itself contain an element of time. In this specification speech, audio, video, timed text and DIMS belong to the first category and still images and text to the latter one.

Streaming of continuous media using RTP/UDP/IP (see clause 6.2) requires a session control protocol to set-up and control the individual media streams. For the transport of discrete media (images and text), vector graphics, timed text and synthetic audio this specification adopts the use of HTTP/TCP/IP (see clause 6.3). In this case there is no need for a separate session set-up and control protocol since this is built into HTTP. This clause describes session set-up and control of continuous media.

5.3.2 RTSP

RTSP [5] shall be used for session set-up and session control. PSS clients and servers shall follow the rules for minimal on-demand playback RTSP implementations in appendix D of [5]. In addition to this:

- PSS servers and clients shall implement the DESCRIBE method (see clause 10.2 in [5]);
- PSS servers and clients shall implement the Range header field (see clause 12.29 in [5]);
- PSS servers shall include the Range header field in all PLAY responses;
- PSS servers and clients should implement the SET_PARAMETER method (see clause 10.9 in [5]);
- PSS servers and clients should implement the Bandwidth header field (see clause 12.6 in [5].

Further additions to RTSP are specified in the following subclauses.

5.3.2.1 The 3GPP-Link-Char header

PSS servers and clients should implement the 3GPP-Link-Char header field.

To enable PSS clients to report the link characteristics of the radio interface to the PSS server, the "3GPP-Link-Char" RTSP header is defined. The header takes one or more arguments. The reported information should be taken from a QoS reservation (i.e. the QoS profile as defined in [56]). Note that this information is only valid for the wireless link and does not apply end-to-end. However, the parameters do provide constraints that can be used.

Three parameters are defined that can be included in the header, and future extensions are possible to define. Any unknown parameter shall be ignored. The three parameters are:

- "GBW": the link's guaranteed bit-rate in kilobits per second as defined by [56];
- "MBW": the link's maximum bit-rate in kilobits per second as defined by [56];
- "MTD": the link's maximum transfer delay, as defined by [56] in milliseconds.

The "3GPP-Link-Char" header syntax is defined below using ABNF [53]:

3gpplinkheader = "3GPP-Link-Char" ":" link-char-spec *("," 0*1SP link-char-spec) CRLF

link-char-spec = char-link-url *(";" 0*1SP link-parameters)

char-link-url = "url" "=" <">url<">

link-parameters = Guaranteed-BW / Max-BW / Max-Transfer-delay / extension-type

```
Guaranteed-BW        = "GBW" "=" 1*DIGIT ; kbps

Max-BW               = "MBW" "=" 1*DIGIT ; kbps

Max-Transfer-delay   = "MTD" "=" 1*DIGIT  ; ms

extension-type       = token "=" (token / quoted-string)

DIGIT                = as defined in RFC 2326 [5]

token                = as defined in RFC 2326 [5]

quoted-string        = as defined in RFC 2326 [5]
url                  = as defined in RFC 2326 [5]
```

The "3GPP-Link-Char" header can be included in a request using any of the following RTSP methods: SETUP, PLAY, OPTIONS, and SET_PARAMETER. The header shall not be included in any response. The header can contain one or more characteristics specifications. Each specification contains a URI that can either be an absolute or a relative, any relative URI use the RTSP request URI as base. The URI points out the media component that the given parameters apply to. This can either be an individual media stream or a session aggregate.

If a QoS reservation (PDP context) is shared by several media components in a session the 3GPP-Link-Char header shall not be sent prior to the RTSP PLAY request. In this case the URI to use is the aggregated RTSP URI. If the QoS reservation is not shared (one PDP context per media) the media stream URI must be used in the 3GPP-Link-Char specification. If one QoS reservation (PDP context) per media component is used, the specification parameters shall be sent per media component.

The "3GPP-Link-Char" header should be included in a SETUP or PLAY request by the client, to give the initial values for the link characteristics. A SET_PARAMETER or OPTIONS request can be used to update the 3GPP-Link-Char values in a session currently playing. It is strongly recommended that SET_PARAMETER is used, as this has the correct semantics for the operation and also requires less overhead both in bandwidth and server processing. When performing updates of the parameters, all of the previous signalled values are undefined and only the given ones in the update are defined. This means that even if a parameter has not changed, it must be included in the update.

Example:

```
3GPP-LinkChar: url="rtsp://server.example.com/media.3gp"; GBW=32; MBW=128; MTD=2000
```

In the above example the header tells the server that its radio link has a QoS setting with a guaranteed bit-rate of 32 kbps, a maximum bit-rate of 128 kbps, and a maximum transfer delay of 2.0 seconds. These parameters are valid for the aggregate of all media components, as the URI is an aggregated RTSP URI.

5.3.2.2 The 3GPP-Adaptation header

PSS servers and clients should implement the 3GPP-Adaptation header field.

To enable PSS clients to set bit-rate adaptation parameters, a new RTSP request and response header is defined. The header can be used in the methods SETUP, PLAY, OPTIONS, and SET_PARAMETER. The header defined in ABNF [53] has the following syntax:

```
3GPP-adaptation-def = "3GPP-Adaptation" ":" adaptation-spec 0*("," adaptation-spec)

adaptation-spec      = url-def *adapt-params

adapt-params         = ";" buffer-size-def

                     / ";" target-time-def

url-def              = "url" "=" <"> url <">

buffer-size-def      = "size" "=" 1*9DIGIT ; bytes

target-time-def      = "target-time" "=" 1*9DIGIT; ms

url                  = ( absoluteURI / relativeURI )
```

absoluteURI and relativeURI are defined in RFC 3986 [60]. The base URI for any relative URI is the RTSP request URI.

The "3GPP-Adaptation" header shall be sent in responses to requests containing this header. The PSS server shall not change the values in the response header. The presence of the header in the response indicates to the client that the server acknowledges the request.

The buffer size signalled in the "3GPP-Adaptation" header shall correspond to reception, de-jittering, and, if used, de-interleaving buffer(s) that have this given amount of space for complete application data units (ADU), including the following RTP header and RTP payload header fields: RTP timestamp, and sequence numbers or decoding order numbers. The specified buffer size shall also include any Annex G pre-decoder buffer space used for this media, as the two buffers cannot be separated.

The target protection time signalled in the "target-time" parameter is the targeted minimum buffer level in milliseconds; that is, the minimum amount of playback time the client perceives necessary for interrupt-free playback. This value must be chosen such that the client is never in a buffering state if all media streams have reached or exceeded their target-time in buffered data and playout delay. Once this desired level of target protection is achieved, the server may utilize any additional resources to increase the quality of the media or to increase the buffer duration beyond that required by the target-time, or it may continue sending at the media rate in order to maintain a steady buffer state.

5.3.2.3 The Quality of Experience headers

5.3.2.3.1 Protocol initiation and termination

A new RTSP header is defined to enable the PSS client and server to negotiate which Quality of Experience (QoE) metrics the PSS client should send, how often they should be sent and how to turn the metrics transmission off. This header can be sent in requests and responses of RTSP methods SETUP, SET_PARAMETER, OPTIONS (with Session ID) and PLAY. The exact usage of this header is defined in clause 11. The header is defined in ABNF [53] as follows (see [53] for specifiers not defined here):

```
QoE-Header         = "3GPP-QoE-Metrics" ":" ("Off" / Measure-Spec *("," Measure-Spec)) CRLF

Measure-Spec       = Stream-URL";" ((Metrics ";" Sending-rate [";" Measure-Range]
                     [";" Measure-Resolution] *([";" Metrics-Server]) *([";" Parameter-Ext])) /  'Off')

Stream-URL         = "url" "="  <">Rtsp-URL<">

Metrics            = "metrics" "=" "{"Metrics-Name *("|" Metrics-Name) " }"

Metrics-Name       =  1*((0x21..0x2b) / (0x2d..0x3a) / (0x3c..0x7a) / 0x7e) ;VCHAR except ';', ',', '{' or '}'

Sending-Rate       = 'rate" "=" 1*DIGIT / "End"

Measure-Resolution = "resolution" "=" 1*DIGIT ; in seconds

Metrics-Server     = "server" "=" "{" Server-Name *("|" Server-Name) "}"

Server-Name        = as defined in RFC 1123 [100]

Measure-Range      = "range" ":" Ranges-Specifier

Parameter-Ext      = 'On'/'Off'/ (1*DIGIT ['.' 1*DIGIT]) / (1*((0x21..0x2b) / (0x2d..0x3a) / (0x3c..0x7a) / 0x7c /
                     0x7e))

Ranges-Specifier   = as defined in RFC 2326 [5]

Rtsp-URL           = as defined in RFC 2326 [5]
```

There are two ways to use this header:

- Using only the "Off" parameter is an indication that either server or client wants to cancel the metrics reporting.
- Using other parameters indicates a request to start the metrics transmission.

If "Stream-URL" is an RTSP Session Control URL, then "Metrics" applies to the RTSP session. If "Stream-URL" is an RTSP Media Control URL, then "Metrics" apply only to the indicated media component of the session.

QoE metrics with the same "Stream-URL", "Sending-rate" and "Measure-Range" shall be aggregated within a single "Measure-Spec" declaration. Otherwise, multiple "Stream-URL" declarations shall be used.

The "Metrics" field contains the list of names that describes the metrics/measurements that are required to be reported in a PSS session. The names that are not included in the "Metrics" field shall not be reported during the session.

The "Sending-Rate" shall be set, and it expresses the maximum time period in seconds between two successive QoE reports. If the "Sending-Rate" value is 0, then the client shall decide the sending time of the reports depending on the events occurred in the client. Values ≥ 1 indicate a precise reporting interval. The shortest interval is one second and the longest interval is undefined. The reporting interval can be different for different media, but it is recommended to maintain a degree of synchronization in order to avoid extra traffic in the uplink direction. The value "End" indicates that only one report is sent at the end of the session.

A default QoE reporting is done for each metric. The optional "Measure-Resolution" field, if present, indicates that XML QoE reporting shall be done instead. In this case the "Measure-Resolution" field splits the session duration into a number of equally sized periods where each period is of the length specified by the "Measure-Resolution" field. QoE metrics are calculated for each period and stored in the terminal, and all the stored metrics are then sent together according to the "Sending-Rate" field. This allows long reporting intervals (to save bandwidth) without losing good metric measurement resolution. It is recommended that the Sending-Rate is set to an integer multiple of the Measure-Resolution, or to "End".

Note that both "Sending-Rate" and "Measure-Resolution" shall be evaluated according to a real-time clock. This implies that the real-time intervals for measurements and reporting are not affected by changes in playback rate, for instance due to buffering.

The optional "Metrics-Server" field, if present, specifies that instead of the default RTSP reporting back to the streaming server, the QoE reports should be sent to a separate HTTP server. If more than one server is specified, the terminal shall randomly select one of them to be used during the session. The Metrics-Server parameter can only be used for XML reporting, that is, together with the Measure-Resolution parameter. The formatting of the HTTP reports is specified in sub-clause 5.3.2.3.3. If the PSS client does not support HTTP reporting it shall reject the "Metrics-Server" field during the QoE negotiation phase.

The optional "Measure-Range" field, if used, shall define the time range in the stream for which the QoE metrics will be reported. There shall be only one range per measurement specification. The range format shall be any of the formats allowed by the media. If the "Measure-Range" field is not present, the corresponding (media or session level) range attribute in SDP shall be used. If SDP information is not present, the metrics range shall be the whole session duration.

There shall be only one "3GPP-QoE-Metrics" header in one RTSP request or response.

5.3.2.3.2 Metrics feedback

The QoE metrics feedback can be conveyed in requests to the PSS server using the SET_PARAMETER, PAUSE or TEARDOWN methods by the "3GPP-QoE-Feedback" header. The header is defined in ABNF [53] as follows (see [53] for specifiers not defined here):

```
Feedbackheader = "3GPP-QoE-Feedback" ":" Feedback-Spec *("," Feedback-Spec) CRLF

Feedback-Spec  = Stream-URL 1*(";" Parameters) [";" Measure-Range]

Stream-URL     = as specified in clause 5.3.2.3.1

Parameters     = Metrics-Name "=" "{" SP / (Measure *("|" Measure)) "}"

Metrics-Name   = as defined in clause 5.3.2.3.1

Measure        = Value [SP Timestamp]

Measure-Range  = as defined in clause 5.3.2.3.1

Value          = (["-"]1*DIGIT ["." *DIGIT]) /  1*((0x21..0x2b) / (0x2d..0x3a) / (0x3c..0x7a)  / 0x7e) ;VCHAR
    except ';', ',', '{' or '}'
```

Timestamp = NPT-Time

NPT-Time = as defined in RFC 2326 [5]

"Stream-URL" is the RTSP session or media control URL that identifies the media the feedback parameter applies to.

The "Metrics-Name" field in the "Parameters" definition contains the name of the metrics/measurements and uses the same identifiers as the "3GPP-QoE-Metrics" header in clause 5.3.2.3.1.

The "Value" field indicates the results. There is the possibility that the same event occurs more than once during a monitoring period. In that case the metrics value may occur more than once indicating the number of events to the server. For the XML reporting format only one value is reported for each measurement resolution period.

The optional "Timestamp" (defined in NPT time) indicates the time when the event occurred or when the metric was calculated. If no events have occurred, it shall be reported with an empty set (only containing a space). The "Timestamp" feedback shall not be used for the XML reporting format.

The optional "Measure-Range" indicates the actual measurement period, for which this report is valid.

QoE metrics reporting should be done by the PSS client by using the SET_PARAMETER method. However, for more efficiency, RTSP PAUSE and TEARDOWN methods may also be used in particular cases, such as:

CASE 1: When sending the very last QoE report, the client should embed the QoE information into a TEARDOWN message.

CASE 2: When the client wants to pause the streaming flow, QoE information should be embedded into a PAUSE method. The PSS client should not send any QoE reports to the PSS server when the system is paused, since there is no media flow.

5.3.2.3.3 Metrics feedback over HTTP

If a specific metrics server has been configured the client should send QoE reports using the HTTP (RFC 2616 [73]) POST request carrying XML formatted metadata. Each QoE report is formatted in XML according the following XML schema. An informative example of a single reception report XML object is also given.

5.3.2.3.3.1 XML Syntax for a QoE Report

Below is the formal XML syntax of QoE report instances.

```
<?xml version="1.0" encoding="UTF-8"?>
<xs:schema xmlns:xs="http://www.w3.org/2001/XMLSchema"
targetNamespace="urn:3gpp:metadata:2009:PSS:receptionreport"
xmlns="urn:3gpp:metadata:2009:PSS:receptionreport"
elementFormDefault="qualified">

<xs:element name="receptionReport" type="receptionReportType"/>

<xs:complexType name="receptionReportType">
    <xs:choice>
        <xs:element name="statisticalReport" type="starType"
        minOccurs="0" maxOccurs="unbounded"/>
        <xs:any namespace="##other" processContents="skip" minOccurs="0" maxOccurs="unbounded"/>
    </xs:choice>
</xs:complexType>

<xs:complexType name="starType">
    <xs:sequence>
        <xs:element name="fileURI" type="xs:anyURI" minOccurs="0" maxOccurs="unbounded"/>
        <xs:element name="qoeMetrics" type="qoeMetricsType" minOccurs="0" maxOccurs="1"/>
        <xs:any namespace="##other" processContents="skip" minOccurs="0" maxOccurs="unbounded"/>
    </xs:sequence>
    <xs:attribute name="serviceId" type="xs:string" use="optional"/>
    <xs:attribute name="clientId" type="xs:string" use="optional"/>
    <xs:attribute name="serviceURI" type="xs:anyURI" use="optional"/>
    <xs:anyAttribute processContents="skip"/>
</xs:complexType>

<xs:complexType name="qoeMetricsType">
    <xs:sequence>
        <xs:element name="medialevel_qoeMetrics" type="medialevel_qoeMetricsType"
```

```
            minOccurs="0" maxOccurs="unbounded"/>
         <xs:any namespace="##other" processContents="skip" minOccurs="0" maxOccurs="unbounded"/>
      </xs:sequence>
      <xs:attribute name="totalRebufferingDuration" type="xs:doubleVectorType" use="optional"/>
      <xs:attribute name="numberOfRebufferingEvents" type="xs:unsignedLongVectorType"
         use="optional"/>
      <xs:attribute name="initialBufferingDuration" type="xs:double" use="optional"/>
      <xs:attribute name="contentSwitchTime" type="xs:doubleVectorType" use="optional"/>
      <xs:attribute name="sessionStartTime" type="xs:unsignedLong"/>
      <xs:attribute name="sessionStopTime" type="xs:unsignedLong"/>
      <xs:attribute name="bufferDepth" type="xs:doubleVectorType" use="optional"/>
      <xs:attribute name="allContentBuffered" type="xs:boolean" use="optional"/>
      <xs:anyAttribute processContents="skip"/>
</xs:complexType>

<xs:complexType name="medialevel_qoeMetricsType">
      <xs:attribute name="sessionId" type="xs:string"/>
      <xs:attribute name="totalCorruptionDuration" type="xs:unsignedLongVectorType" use="optional"/>
      <xs:attribute name="numberOfCorruptionEvents" type="xs:unsignedLongVectorType" use="optional"/>
      <xs:attribute name="t" type="xs:boolean" use="optional"/>
      <xs:attribute name="totalNumberofSuccessivePacketLoss" type="xs:unsignedLongVectorType"
         use="optional"/>
      <xs:attribute name="numberOfSuccessiveLossEvents" type="xs:unsignedLongVectorType"
         use="optional"/>
      <xs:attribute name="numberOfReceivedPackets" type="xs:unsignedLongVectorType" use="optional"/>
      <xs:attribute name="totalJitterDuration" type="xs:doubleVectorType" use="optional"/>
      <xs:attribute name="numberOfJitterEvents" type="xs:unsignedLongVectorType" use="optional"/>
      <xs:attribute name="framerate" type="xs:doubleVectorType" use="optional"/>
      <xs:attribute name="codecInfo" type="stringVectorType" use="optional"/>
      <xs:attribute name="codecProfileLevel" type="stringVectorType" use="optional"/>
      <xs:attribute name="codecImageSize" type="stringVectorType" use="optional"/>
      <xs:attribute name="averageCodecBitrate" type="doubleVectorType" use="optional"/>
      <xs:anyAttribute processContents="skip"/>
</xs:complexType>

<xs:simpleType name="doubleVectorType"
      <xs:list itemType="xs:double"/>
</xs:simpleType>

<xs:simpleType name="unsignedLongVectorType"
      <xs:list itemType="xs:unsignedLong"/>
</xs:simpleType>

<xs:simpleType name="stringVectorType"
      <xs:list itemType="xs:string"/>
</xs:simpleType>

</xs:schema>
```

5.3.2.3.3.2 Example XML for the QoE Report

The example shows a QoE report for a streaming session.

```
<?xml version="1.0" encoding="UTF-8"?>
<receptionReport xmlns="urn:3gpp:metadata:2009:PSS:receptionreport"
      xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
      xsi:schemaLocation="urn:3gpp:metadata:2009:PSS:receptionreport receptionreport.xsd">
   <statisticalReport
      clientId="clientID"
      serviceURI="xxx.example.com"
      serviceId="serviceID">
      <qoeMetrics
         numberOfRebufferingEvents="0 1 0"
         initialBufferingDuration="3.213"
         totalRebufferingDuration="0 1.23 0"
         contentAccessTime="2.621"
         sessionStartTime="1219322514"
         sessionStopTime="1219322541">
         bufferDepth="3.571 2.123 2.241"
         allContentBuffered="false">
         <medialevel_qoeMetrics
            sessionId="10.50.65.30:5050"
            framerate="15.1 14.8 15.0"
            t="false"
            numberOfSuccessiveLossEvents="5 0 3"
            numberOfCorruptionEvents="6 5 2"
```

```
            numberOfJitterEvents="0 1 0"
            totalCorruptionDuration="152 234 147"
            totalNumberofSuccessivePacketLoss="25 0 6"
            numberOfReceivedPackets="456 500 478"
            codecInfo="H263-2000/90000 = ="
            codecProfileLevel="profile=0;level=45 = ="
            codecImageSize="176x144 = ="
            averageCodecBitRate="124.5 128.0 115.1"
            totalJitterDuration="0 0.346 0"/>
        </qoeMetrics>
    </statisticalReport>
</receptionReport>
```

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 SDP

5.3.3.1 General

RTSP requires a presentation description. SDP shall be used as the format of the presentation description for both PSS clients and servers. PSS servers shall provide and PSS clients interpret the SDP syntax according to the SDP specification [6] and appendix C of [5]. The SDP delivered to the PSS client shall declare the media types to be used in the session using a codec specific MIME media type for each media. MIME media types to be used in the SDP file are described in clause 5.4 of the present document.

The SDP [6] specification requires certain fields to always be included in an SDP file. Apart from this a PSS server shall always include the following fields in the SDP:

- "a=control:" according to clauses C.1.1, C.2 and C.3 in [5];
- "a=range:" according to clause C.1.5 in [5];
- "a=rtpmap:" according to clause 6 in [6];
- "a=fmtp:" according to clause 6 in [6].

When an SDP document is generated for media stored in a 3GP file, each control URL defined at the media-level 'a=control:' field shall include a stream identifier in the last segment of the path component of the URL. The value of the stream id shall be defined by the track-ID field in the track header (tkhd) box associated with the media track. When a PSS server receives a set-up request for a stream, it shall use the stream identifier specified in the URL to map the request to a media track with a matching track-ID field in the 3GP file. Stream identifiers shall be expressed using the following syntax:

streamIdentifier = <stream-id-token>"="<stream-id>

stream-id-token = 1*alpha

stream-id = 1*digit

The bandwidth field in SDP is needed by the client in order to properly set up QoS parameters. Therefore, a PSS server shall include the 'b=AS:' and "b=TIAS:" and "a=maxprate" [93] fields at the media level for each media stream in SDP, and should include "b=TIAS" and "a=maxprate" at session level. A PSS client shall interpret all of these fields. If both bandwidth modifiers are present, "b=TIAS" should be used; however it may be missing in content produced according to earlier releases. When a PSS client receives SDP, it should ignore the session level 'b=AS:' parameter (if present), and instead calculate session bandwidth from the media level bandwidth values of the relevant streams. If "b=TIAS" and "a=maxprate" is present at session level, it should be used in preference over the media level values, as session level can provide a more accurate description of the needed session bandwidth when aggregating several media streams together. A PSS client shall also handle the case where the bandwidth parameters are not present, since this may occur when connecting to a Release-4 server.

Note that for RTP based applications, "b=AS:" gives the RTP "session bandwidth" (including UDP/IP overhead) as defined in section 6.2 of [9].

The bandwidth for RTCP traffic shall be described using the "RS" and "RR" SDP bandwidth modifiers, as specified by [55]. The "RS" SDP bandwidth modifier indicates the RTCP bandwidth allocated to the sender (i.e. PSS server) and "RR" indicates the RTCP bandwidth allocated to the receiver (i.e. PSS client). A PSS server shall include the "b=RS:" and "b=RR:" fields at the media level for each media stream in SDP, and a PSS client shall interpret them. A PSS client shall also handle the case where the bandwidth modifier is not present according to section 3 of [55], since this may occur when connecting to a Release-4 server.

There shall be a limit on the allowed RTCP bandwidth for senders and receivers in a session. This limit is defined as follows:

- 4000 bps for the RS field (at media level);
- 5000 bps for the RR field (at media level).

In Annex A.2.1 an example SDP in which the limit for the total RTCP bandwidth is 5% of the session bandwidth is presented.

Media which has an SDP description that includes an open ended range (format=startvalue-) in any time format in the SDP attribute "a=range", e.g. "a=range: npt=now-", or "a=range: clock=20030825T152300Z-", shall be considered media of unknown length. Such a media shall be considered as non-seekable, unless other attributes override this property.

The "t=", "r=", and "z=" SDP parameters are used to indicate when the described session is active and can be used to filter out obsolete SDP files. PSS clients and servers shall support "t=", "r=", and "z=" as specified in [6]. The "a=etag" parameter may additionally be used to identify SDP validity. PSS clients should support "a=etag" as specified in [5].

When creating an SDP for a streaming session, one should try to come up with the most accurate estimate of time that the session is active. The "t=", "r=", and "z=" SDP parameters are used for this purpose, i.e., to indicate when the described session is active. If the time at which a session is active is known to be only for a limited period, the "t=", "r=", and "z=" attributes should be filled out appropriately (per [6], the "t=" shall be sent and usually contains non-zero values, possibly using the "r=" and "z=" parameters). If the stop-time is set to zero, the session is not bounded, though it will not become active until after the start-time. If the start-time is also zero, the session is regarded as permanent. A session should only be marked as permanent ("t=0 0") if the session is going to be available for a significantly long period of time or if the start and stop times are not known at the time of SDP file creation. Recommendations for what is considered a significant time is present in the SDP specification [6].

IPv6 addresses in SDP descriptions shall be supported according to RFC 4566 [6].

NOTE: The SDP parsers and/or interpreters shall be able to accept NULL values in the 'c=' field (e.g. 0.0.0.0 in IPv4 case). This may happen when the media content does not have a fixed destination address. For more details, see Section C.1.7 of [5] and Section 6 of [6].

5.3.3.2 Additional SDP fields

The following additional 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 [90]. 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 [90]. If "X-predecbufsize" is not present for an H.264 (AVC) stream, the value of "CpbSize" is calculated as specified in [90].

- "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 [90]. 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 clause G.3 of Annex G 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 [90] 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, 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 [90]. If "3gpp-videopostdecbufsize" is not present for an H.264 (AVC) stream, the value of "max_dec_frame_buffering" is inferred as specified in [90].

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 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.

If integrity protection is supported, the following SDP attributes shall be supported by the client and server:

- "a=3GPP-Integrity-Key" according to annex K;
- "a=3GPP-SRTP-Config" according to Annex K;
- "a=3GPP-SDP-Auth" according to Annex K.

If RTP retransmission is supported, the following SDP attribute shall be supported by the client and server:

- "a=rtcp-fb" according to clause 4.2 in [57].

5.3.3.3 The 'alt' and 'alt-default-id' attributes

The client should interpret the following two media level attributes: "alt" and "alt-default-id". A client from earlier releases will ignore these attributes and can safely do so in a correctly formatted SDP. If the attributes are used by the server they shall be used in a way that makes them backward compatible. When interpreted, they define a number of alternatives from which the client can select the most appropriate one.

A non-extended SDP gives only one alternative for each media part (Annex A.1 Example 1). This is the default alternative for each media. The new SDP attributes defined here are used to modify the default attributes or to add new attributes to the default attributes thus creating new alternatives. Each alternative is numerically identified.

The alternative attribute "alt" is used to replace or add an SDP line to the default configuration. If the alternative attribute contains an SDP line, for which the type and the modifier already exist in the default alternative, the default must be replaced with the given line(s). In case there are multiple lines with the same type and modifier in the default alternative, all of the lines must be replaced. Multiple alternative lines can be used to modify the default alternative. The alternative lines that are used to form a certain alternative shall all carry the same numerical identifier (Annex A.1, Examples 2-4).

The alternative identifier is a unique identifier that points out a single alternative in one media declaration. The identifier must be unique between all media descriptions and their alternatives as it is used for creating combinations between different medias with the grouping attribute (see 5.3.3.4).

The default configuration is in itself a valid alternative. Therefore an attribute (alt-default-id) is defined that assigns an alternative identifier to the default alternative. This identifier can then be used with the grouping attribute (see 5.3.3.4) to create combinations of alternatives from different medias.

The alternative attribute is defined below in ABNF from RFC 4234 [53]. The SDP line is any SDP line allowed at media level except "m=".

```
alt             = "a" "=" "alt" ":" alt-id ":" SDP-line CRLF

SDP-line        = <type>=<value> ; See RFC 4566 [6]

alt-id          = 1*DIGIT ; unique identifier for the alternative in whole SDP.
```

To be able to assign an alternative ID to the default alternative, the following identification attribute is defined.

```
alt-default-id  = "a" "=" "alt-default-id" ":" alt-id CRLF
```

5.3.3.4 The session level grouping attribute, 'alt-group'

The client should handle the following attribute: "alt-group". A client from earlier releases will ignore this attribute and can safely do so. When interpreted, it defines a number of grouping alternatives from which the client can select the most appropriate one. The identifiers defined in 5.3.3.3 are used together with the "alt-group" attribute to create combinations consisting of, e.g., one audio and one video alternative.

A grouping attribute is used to recommend certain combinations of media alternatives to the client. There may be more than one grouping attribute at the session level as long as they are for different grouping types and subtypes.

```
alt-group = "a" "=" "alt-group" ":" alt-group-type ":" alt-group-subtype ":" alt-grouping *(";" alt-grouping) CRLF

alt-group-type      = token    ; "token" defined in RFC 4566 [6]

alt-group-subtype   = token

alt-grouping        = grouping-value "="  alt-id *("," alt-id)

grouping-value      = token
```

The alt-group attribute gives one or more combinations of alternatives through their IDs. Each grouping shall be given a grouping value. The grouping value is used to determine if the alternatives within the grouping suit the client. New types and subtypes can be added later.

The following grouping types and subtypes are defined:

- Type: BW, Subtype: All modifiers defined for the SDP "b=" attribute at session and media level. See www.IANA.org for current list of registered attributes.

  Grouping value: The bandwidth value defined for that modifier calculated over all the alternatives grouped together in that grouping. For SDP bandwidth modifiers defined at session level the value shall be calculated according to its rule over the alternative part of the grouping. For media-level-only modifiers, the grouping value shall be calculated as a sum of the media-level values in the grouped alternatives. For TIAS [93] the bandwidth value alone is not sufficient to provide a receiver with sufficient information to make a decision. The SDP attribute "maxprate" is also needed. To provide this information in the grouping-value the following syntax shall be used: <bit-rate>_<maxprate>, where <bit-rate> is the bit-rate value for TIAS and <maxprate> is the maxprate value corresponding to the SDP attribute.

  Grouping recommendations: Each grouping should only contain one alternative from each media type. There is no need to give groupings for all combinations between the media alternatives, rather it is strongly recommended to only give the most suitable combinations (Annex A.1 Example 5). The client can use the bandwidth values of the grouping to estimate the minimum, guaranteed or maximum bandwidth that will be needed for that session.

- Type: LANG Subtype: RFC3066

  Grouping value: A language tag as defined by RFC 3066 [54]. The grouping MUST contain all media alternatives, which support that language tag.

Grouping recommendations: It is recommended that other mechanisms, like user profiles if existing, are primarily used to ensure that the content has language suitable for the user (Annex A.1, Example 6).

Se also Annex A1, Examples 7 through 16. In the examples all three new attributes "alt", "alt-default-id" and "alt-group" are used.

5.3.3.5 The bit-rate adaptation support attribute, '3GPP-Adaptation-Support'

To signal the support of bit-rate adaptation, a media level only SDP attribute is defined in ABNF [53]:

```
sdp-Adaptation-line  = "a" "=" "3GPP-Adaptation-Support" ":" report-frequency CRLF

report-frequency     = NonZeroDIGIT [ DIGIT ]

NonZeroDIGIT         = %x31-39    ;1-9
```

A server implementing rate adaptation shall signal the "3GPP-Adaptation-Support" attribute in its SDP.

A client receiving an SDP description where the SDP attribute "3GPP-Adaptation-Support" is present knows that the server provides rate adaptation. The client, if it supports bit-rate adaptation, shall then in its subsequent RTSP signalling use the '3GPP-Adaptation' header as defined in clause 5.3.2.2, as well as the RTCP Next Application Data Unit (NADU) APP packet for reporting the next unit to be decoded, as defined in clause 6.2.3.2.

The SDP attribute shall only be present at the media level. The report frequency value, which shall be larger than zero, indicates to the client that it shall include a NADU APP packet in a compound RTCP packet no less often than the interval specified by report-frequency, except prior to receipt of RTP media packets, when the client is unable to generate a valid NADU APP packet. For example, if this value is 3, the client shall send the NADU APP packet in at least every 3rd RTCP packet.

5.3.3.6 The Quality of Experience support attribute, "3GPP-QoE-Metrics"

PSS servers using QoE-Metrics in a session shall use SDP to initiate the QoE negotiation. The reason why SDP is needed is to support the use cases where SDP is distributed through other methods than RTSP DESCRIBE, e.g. WAP, HTTP or email. A new SDP attribute, which can be used either at session or media level, is defined below in ABNF [53] based on RFC 4566 [6]:

```
QoE-Metrics-line    = "a" "=" "3GPP-QoE-Metrics:" att-measure-spec *("," att-measure-spec)) CRLF

att-measure-spec    = Metrics ";" Sending-rate [";" Measure-Range] *([";" Parameter-Ext])

Metrics             = as defined in clause 5.3.2.3.1.

Sending-Rate        = as defined in clause 5.3.2.3.1.

Measure-Range       = as defined in clause 5.3.2.3.1.

Parameter-Ext       = as defined in clause 5.3.2.3.1.
```

PSS servers using QoE-Metrics in a session shall use this attribute to indicate that QoE metrics are supported and will be sent. When present at session level, it shall only contain metrics that apply to the complete session. When present at media level, it shall only contain metrics that are applicable to individual media. The URI that is used in the specification of the RTSP header "3GPP-QoE-Metrics:" is implicit by the RTSP control URI (a=control).

5.3.3.7 The asset information attribute, "3GPP-Asset-Information"

This asset information attribute is defined to transmit asset information in SDP. The attribute is defined ABNF [53]:

```
3GPP-Assets-Info    = "a" "=" "3GPP-Asset-Information:" Asset 0*("," Asset) CRLF

Asset               = ("{" "url" "=" <">URL<"> "}") / ("{"AssetName "=" AssetBox "}")

URL                 = as defined in [60]

AssetName           = "Title" / "Description" / "Copyright" / "Performer" / "Author" / "Genre" / "Rating" /
```

"Classification" / "Keywords" / "Location" / "Album" / "RecordingYear" / asset-extension

asset-extension = 1*((0x01..0x09) / 0x0b / 0x0c / (0x0e..0x1f) / (0x21..0x2b) / (0x2d..0x3c) / (0x3e..0x7a) / 0x7c / (0x7d..0xff)) ;any byte except SP, NUL, CR, LF , "=", ",", "{" or "}"

AssetBox = Base64 encoded version [69] of any asset box as defined in Clause 8 of [50].

This SDP attribute can be present at session level, media level or both. Multiple instances of the attribute are allowed.

The resource referenced by the URL can be any pre-formatted data, e.g. an XHTML page or XML file, containing any asset information. It is up to the client"s capability and user"s preference to render the information pointed by the URL.

Example 17 in Clause A.1 shows an SDP file that includes the "3GPP-Asset-Information" attribute.

5.3.3.8 OMA-DM Configuration of QoE Metrics

As an optional alternative to configure the QoE reporting for each session (i.e. via SDP/RTSP), OMA-DM can be used to specify the default QoE configuration. If such a default QoE configuration has been specified, it shall be used by the terminal for all subsequent PSS sessions where no session-specific QoE configuration is received. QoE reporting based on the default OMA configuration shall always be done over HTTP with the XML reporting format. If the PSS client does not support HTTP reporting it shall not send default QoE reports.

Any session-specific QoE configuration received shall always have higher priority, and will in such cases override any default OMA-DM QoE configuration for that session.

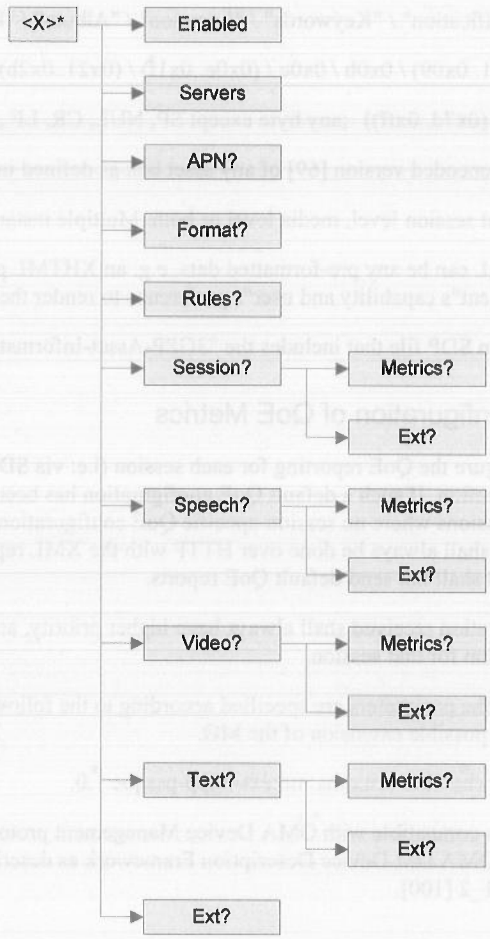
For OMA-DM QoE configuration the parameters are specified according to the following Managed Object (MO). Version numbering is included for possible extension of the MO.

The Management Object Identifier shall be: urn:oma:mo:ext-3gpp-pssqoe:1.0.

Protocol compatibility: The MO is compatible with OMA Device Management protocol specifications, version 1.2 and upwards, and is defined using the OMA DM Device Description Framework as described in the Enabler Release Definition OMA-ERELD _DM-V1_2 [100].

5.3.3.8.1 QoE metrics reporting management object

The following nodes and leaf objects shall be contained under the 3GPP_PSSQOE node if a PSS client supports the feature described in this clause (information of DDF for this MO is given in Annex P):



Node: /<X>

This interior node specifies the unique object id of a PSS QoE metrics management object. The purpose of this interior node is to group together the parameters of a single object.

- Occurrence: ZeroOrOne
- Format: node
- Minimum Access Types: Get

The following interior nodes shall be contained if the PSS client supports the 'PSS QoE metrics Management Object'.

/<X>/Servers

This leaf contains a space-separated list of servers to which the QoE reports are transmitted. It is URI addresses, e.g. http://qoeserver.operator.com. In case of multiple servers, the PSS client randomly selects one of the servers from the list, with uniform distribution.

- Occurrence: One
- Format: chr
- Minimum Access Types: Get

- Values: URI of the servers to receive the QoE report.

/<X>/Enabled

This leaf indicates if QoE reporting is requested by the provider.

- Occurrence: One
- Format: bool
- Minimum Access Types: Get

/<X>/APN

This leaf contains the Access Point Name that should be used for establishing the PDP context on which the QoE metric reports will be transmitted. This may be used to ensure that no costs are charged for QoE metrics reporting. If this leaf is not defined then any QoE reporting is done over the default access point.

- Occurrence: ZeroOrOne
- Format: chr
- Minimum Access Types: Get
- Values: the Access Point Name

/<X>/Format

This leaf specifies the format of the report and if compression (Gzip XML) [xx] is used.

- Occurrence: ZeroOrOne
- Format: chr
- Minimum Access Types: Get
- Values: 'XML', 'GZIPXML'

/<X>/Rules

This leaf provides in textual format the rules used to decide whether metrics are to be reported to the QoE metrics report server. The syntax and semantics of this leaf are defined in sub-clause 5.3.3.8.2.

- Occurrence: ZeroOrOne
- Format: chr
- Minimum Access Types: Get
- Values: See clause 5.3.3.8.2.

/<X>/Ext

The Ext node is an interior node where the vendor specific information can be placed (vendor includes application vendor, device vendor etc.). Usually the vendor extension is identified by vendor specific name under the ext node. The tree structure under the vendor identified is not defined and can therefore include one or more un-standardized sub-trees.

- Occurrence: ZeroOrOne
- Format: node
- Minimum Access Types: Get

/<X>/Session

The Session node is the starting point of the session level QoE metrics definitions.

- Occurrence: ZeroOrOne
- Format: node
- Minimum Access Types: Get

/<X>/Session/Metrics

This leaf provides in textual format the QoE metrics that need to be reported, the measurement frequency, the reporting interval and the reporting range. The syntax and semantics of this leaf are defined in clause 8.3.2.1.

- Occurrence: ZeroOrOne
- Format: chr
- Minimum Access Types: Get
- Values: see clause 8.3.2.1.

/<X>/Session/Ext

The Ext node is an interior node where the vendor specific information can be placed (vendor meaning application vendor, device vendor etc.). Usually the vendor extension is identified by vendor specific name under the ext node. The tree structure under the vendor identified is not defined and can therefore include one or more un-standardized sub-trees.

- Occurrence: ZeroOrOne
- Format: node
- Minimum Access Types: Get

/<X>/Speech

The Speech node is the starting point of the speech/audio media level QoE metrics definitions.

- Occurrence: ZeroOrOne
- Format: node
- Minimum Access Types: Get

/<X>/Speech/Metrics

This leaf provides in textual format the QoE metrics that need to be reported, the measurement frequency, the reporting interval and the reporting range. The syntax and semantics of this leaf are defined in clause 8.3.2.1.

- Occurrence: ZeroOrOne
- Format: chr
- Minimum Access Types: Get
- Values: see clause 8.3.2.1.

/<X>/Speech/Ext

The Ext node is an interior node where the vendor specific information can be placed (vendor meaning application vendor, device vendor etc.). Usually the vendor extension is identified by vendor specific name under the ext node. The tree structure under the vendor identified is not defined and can therefore include one or more un-standardized sub-trees.

- Occurrence: ZeroOrOne

- Format: node
- Minimum Access Types: Get

/<X>/Video

The Video node is the starting point of the video media level QoE metrics definitions.

- Occurrence: ZeroOrOne
- Format: node
- Minimum Access Types: Get

/<X>/Video/Metrics

This leaf provides in textual format the QoE metrics that need to be reported, the measurement frequency, the reporting interval and the reporting range. The syntax and semantics of this leaf are defined in clause 8.3.2.1.

- Occurrence: ZeroOrOne
- Format: chr
- Access Types: Get
- Values: see clause 8.3.2.1.

/<X>/Video/Ext

The Ext is an interior node where the vendor specific information can be placed (vendor meaning application vendor, device vendor etc.). Usually the vendor extension is identified by vendor specific name under the Ext node. The tree structure under the vendor identified is not defined and can therefore include one or more un-standardized sub-trees.

- Occurrence: ZeroOrOne
- Format: node
- Minimum Access Types: Get

/<X>/Text

The Text node is the starting point of the timed-text media level QoE metrics definitions.

- Occurrence: ZeroOrOne
- Format: node
- Minimum Access Types: Get
- Values: see clause 8.3.2.1.

/<X>/Text/Metrics

This leaf provides in textual format the QoE metrics that need to be reported, the measurement frequency, the reporting interval and the reporting range. The syntax and semantics of this leaf are defined in clause 8.3.2.1.

- Occurrence: ZeroOrOne
- Format: chr
- Minimum Access Types: Get
- Values: see clause 8.3.2.1.

/<X>/Text/Ext

The Ext is an interior node where the vendor specific information can be placed (vendor meaning application vendor, device vendor etc.). Usually the vendor extension is identified by vendor specific name under the ext node. The tree structure under the vendor identified is not defined and can therefore include one or more un-standardized sub-trees.

- Occurrence: ZeroOrOne
- Format: node
- Minimum Access Types: Get

5.3.3.8.2 QoE reporting rule definition

This clause defines the syntax and semantics of a set of rules which are used to reduce the amount of reporting to the QoE metrics report server. The syntax of the metrics reporting rules is defined below:

- QoE-Rule = "3GPP-QoE-Rule" ":" rule-spec *("," rule-spec)
- rule-spec = rule-name [";" parameters]
- rule-name = "LimitSessionInterval" / "SamplePercentage"
- parameters = parameter *(";" parameter)
- parameter = Param-Name ["=" Param-Value]
- Param-Name = 1*((0x21..0x2b) / (0x2d..0x3a) / (0x3c..0x7a) / 0x7e) ;VCHAR except ";", ",", "{" or "}"
- Param-Value = (1*DIGIT ["." 1*DIGIT]) / (1*((0x21..0x2b) / (0x2d..0x3a) / (0x3c..0x7a) / 0x7c / 0x7e))

The semantics of the rules and the syntax of its parameters are defined below:

The *SamplePercentage* rule can be used to set a percentage sample of calls which should report reception. This can be useful for statistical data analysis of large populations while increasing scalability due to reduced total uplink signalling. The *sample_percentage* parameter takes on a value between 0 and 100, including the use of decimals. It is recommended that no more than 3 digits follow a decimal point (e.g. 67.323 is sufficient precision).

When the *SamplePercentage* rule is not present or its *sample_percentage* parameter value is 100 each PSS client shall send metric report(s). If the *sample_percentage* value is less than 100, the UE generates a random number which is uniformly distributed in the range of 0 to 100. The UE sends the reception report when the generated random number is of a lower value than the *sample_percentage* value.

The *LimitSessionInterval* rule is used to limit the time interval between consecutive sessions that report metrics. The *min_interval* parameter for this rule indicates the minimum time distance, in seconds, between the start of two sessions that are allowed to report metrics. When this rule is absent there is no limitation on the minimum time interval.

In case multiple rules are defined in the Management Object, the PSS client should only report metrics when all individual rules evaluate to true (i.e. the rules are logically ANDed). In case no rules are present the PSS client should always report metrics (see also clause xx for metrics reporting procedures).

An example for a QoE metric reporting rule is shown below:

3GPP-QoE-Rule:SamplePercentage;sample_percentage=10.0,LimitSessionInterval;min_interval=300

This example rule defines that only 10% of the sessions shall report, with the minimum time interval between the start times of two consecutive sessions that report metrics to be 5 minutes.

5.4 MIME media types

For continuous media the following MIME media types shall be used:

- AMR narrow-band speech codec (see sub-clause 7.2) MIME media type as defined in [11];
- AMR wideband speech codec (see sub-clause 7.2) MIME media type as defined in [11];
- Extended AMR-WB codec (see sub-clause 7.3) MIME media type as defined in [85];

- Enhanced aacPlus and MPEG-4 AAC audio codecs (see clause 7.3) MIME media type as defined in RFC 3016 [13].
  The following applies to servers when this MIME type is used in SDP:
  1. Configuration information is exclusively carried out-of-band in the SDP 'config' parameter; this shall be signaled by sending 'cpresent=0'.
  2. A PSS server serving implicitly signaled Enhanced aacPlus content shall include 'SBR-enabled=1' in the 'a=fmtp' line; it shall include 'SBR-enabled=0' if it serves plain AAC content.
  3. A PSS server serving explicitly signaled content is recommended not to include the 'SBR-enabled' parameter in the 'a=fmtp' line.

  Therefore, the following applies to terminals:
  1. The rtpmap rate parameter should not be considered definitive of the sampling rate (though it is, of course, definitive of the timescale of the RTP timestamps).
  2. If explicit signaling is in use, the StreamMuxConfig contains both the core AAC sampling rate and the SBR sampling rate. The appropriate output sampling rate may be chosen dependant on Enhanced aacPlus support.
  3. If explicit signalling is not in use and no SBR-enabled parameter is present, the StreamMuxConfig contains the AAC sampling rate and the appropriate output sampling rate may be set to this indicated rate.
  4. If explicit signalling is not in use and the SBR-enabled parameter is present, terminals supporting Enhanced aacPlus should set the output sampling rate to either the core AAC sampling rate as indicated in the StreamMuxConfig [21] (where 'SBR-enabled' is set to '0') or twice the indicated rate (where 'SBR-enabled' is set to '1');
- MPEG-4 Visual video codec (see sub-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 sub-clause 7.4) MIME media type as defined in clause 8.1.2 of [14]. In order to guarantee backward compatibility with earlier Releases (before Release 7), MIME parameters other than 'profile' and 'level' should not be used;
- H.264 (AVC) [90] video codec (see sub-clause 7.4) MIME media type as defined in [92];
- 3GPP timed text format [51] MIME media type as defined in sub-clause 7.1 of [80];
- OMA DRM protected streaming media MIME media type as defined in clause K.1.4 in Annex K;
- RTP retransmission payload format MIME media types as defined in clause 8 of [81];
- DIMS MIME media type as defined in [98].

MIME media types for JPEG, GIF, PNG, SP-MIDI, Mobile DLS, Mobile XMF, SVG, timed text, 3GP and XHTML can be used in the "Content-type" field in HTTP, "content_type" field in the item information box of 3GP files, and in the "type" attribute in SMIL 2.0, SVG Tiny 1.2 and DIMS. The following MIME media types shall be used for these media:

- JPEG (see sub-clause 7.5) MIME media type as defined in [15];
- GIF (see sub-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 [97];

- 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 sub-clause 7.8) MIME media type as defined in [16];
- Timed text (see sub-clause 7.9) MIME media type as defined in [79];
- 3GP files (see sub-clause 7.10) MIME media type as defined in [79].

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

NOTE: The 3GP MIME media type [79] is used for all 3GP files, including 3GP files carrying timed text, DIMS, images, etc.

5.5 Extension for Fast Content Switching and Start-up

5.5.1 Introduction

Applications which are built on top of packet switched streaming (PSS) services are classified into on-demand and live information delivery applications. This clause defines procedures to allow faster start up and switching of content for both on-demand and live applications by reducing the client/server interactions to a minimum. Additionally, clients are enabled to reuse the existing RTSP control session and RTP resources while switching to new content.

5.5.2 Extensions to RTSP 1.0

5.5.2.1 Introduction

Various general RTSP extensions are required for support of fast content start-up and switching. These extensions must be implemented by PSS clients and servers wishing to support any of these features.

The following new RTSP feature tags are defined:

- '3gpp-pipelined' feature-tag, section 5.5.3
- '3gpp-switch' feature-tag, section 5.5.4.3
- '3gpp-switch-req-sdp' feature-tag, section 5.5.4.4
- '3gpp-switch-stream' feature-tag, section 5.5.4.5

In addition the following new RTSP header fields are defined:

- 'Switch-Stream' header, section 5.5.4.2
- 'SDP-Requested' header, section 5.5.4.4
- 'Pipelined-Requests' header, section 5.5.3

5.5.2.2 Capability Handling

5.5.2.2.1 Introduction

The PSS client shall determine the PSS server''s capabilities and indicate its own capabilities to the server using the 'Supported' header field (see clause 5.5.2.2.2) as early as possible (e.g. with the DESCRIBE request).

The 'Require' header field is used as defined in RFC 2326 [5] to ensure that a certain feature is supported by the PSS server. An unsupported feature shall cause the containing request to fail with a 551 reply code and 'Option not supported' as the reason. The 551 reply shall also include the unsupported features in the 'Unsupported' header field.

The 'Require' header field should not be used for probing support for features but rather to make sure that a specific request is executed correctly with the specified features.

5.5.2.2.2 Definition of the 'Supported' RTSP Header Field

PSS clients and servers must support the 'Supported' RTSP header field.

The 'Supported' header field enumerates all the extensions supported by the client or server using feature tags. The header carries the extensions supported by the message sending entity. The 'Supported' header may be included in any request. When present in a request, the server shall respond with its corresponding 'Supported' header.

If an RTSP request includes a 'Supported' header but the corresponding response does not include this header field, then the PSS client shall assume that the PSS server does not support any of the indicated features

Note, the 'Supported' header must be included in error as well as success responses.

The Supported header field contains a list of feature-tags, described in Section 5.5.2.1, that are understood by the client or server.

Example:

```
C->S:   OPTIONS rtsp://3gpp.org/ RTSP/1.0
        CSeq: 1
        Supported: 3gpp-pipelined

S->C:   RTSP/1.0 200 OK
        CSeq: 1
        Public: OPTIONS, DESCRIBE, SETUP, PLAY, PAUSE, TEARDOWN
        Supported: 3gpp-pipelined, 3gpp-switch, 3gpp-switch-req-sdp, 3gpp-switch-stream
```

5.5.2.3 SSRC in the 'RTP-Info' RTSP Header Field

The "RTP-Info" response header field is used to set RTP-specific parameters in the PLAY response. For streams using RTP as transport protocol the "RTP-Info" header is always part of a 200 response to PLAY (as defined in Annex A.2.1). In addition to the parameters defined in RFC 2326 [5], the "RTP-Info" header may also include a synchronization source (SSRC) parameter.

Note. The SSRC parameter is mandatory for some content switching procedures defined in clause 5.5.4.

The SSRC parameter gives the synchronization source (SSRC) of the RTP flow to which the RTP timestamp and sequence number apply. It is only possible to describe one synchronization source (SSRC) per media resource.

After a fast content switch (FCS) the SSRC source used on a specific RTP session may change. In the event that the SSRC changes, it shall be included in the RTP-Info header. The PSS server shall change the SSRC value for a specific RTP session after a fast content switching operation is performed and

- if the payload type of the old and of the new media stream is the same but the media codec configuration is different, or

- if the mapping of the new media stream is otherwise unknown to the PSS client.

In case the SSRC remains unchanged after a content switch, the RTP sequence number and timestamp should be continuous and shall be monotonically increasing. Otherwise, a random RTP sequence number and timestamp should be used.

Further details on the usage of the SSRC are described in clause 5.5.3 or clause 5.5.4. Note the SSRC may only be included in the "RTP-Info" header, if the client has requested one of the features defined in clause 5.5.3 or clause 5.5.4.The "RTP-Info" header syntax in ABNF [53] is as follows:

```
RTP-Info         = "RTP-Info" ":"rtsp-info-spec *("," rtsp-info-spec) CRLF

rtsp-info-spec   = stream-url 1*parameter

stream-url       = "url" "=" rtsp-url

parameter        = ";" "ssrc" "=" 8HEX
```

/ ";" "seq" "=" 1*DIGIT

/ ";" "rtptime" "=" 1*DIGIT

5.5.2.4 Semantics of RTSP PLAY method

The queued PLAY functionality described in RFC 2326 [5] is removed. If a PLAY request is received for an RTSP session that is in the Playing State, then the server shall immediately execute the new PLAY request and terminate the old.

5.5.3 Start-up

In order to improve start-up times, a client may pipeline all necessary SETUP requests and the PLAY request. This allows streaming to begin with a single RTSP round trip if the client already has the SDP (or other adequate content description), or two round trips if it needs to first perform a DESCRIBE in order to receive the necessary information.

If the client intends to send upstream packets to ensure correctly open firewalls (also called port punching packets), then the client should not send a PLAY request until all SETUP responses are received. Pipelining of SETUP requests is still possible in this case.

If the client uses RTSP DESCRIBE to fetch the SDP from the server, then the client shall probe the server capabilities as described in clause 5.5.2.2 using the feature-tag value "3gpp-pipelined".

The client shall add the RTSP 'Require' header to all but the first pipelined RTSP SETUP request with the value '3gpp-pipelined". Note that the first RTSP SETUP request shall not use a 'Require' header. This will allow the PSS client to interoperate with minimal impact with older servers that do not support this feature.

Since the session does not yet exist when these pipelined messages are sent, a request header is defined which allows the client to inform the server that these messages are to be carried on the same session once it is created. Clients wishing to use pipelined start-up must implement the 'Pipelined-Requests' header in order to signal the session grouping to the server.

The syntax of the 'Pipelined-Requests' header is defined in ABNF [53] as follows:

Pipe-Hdr = "Pipelined-Requests" COLON startup-id

startup-id = 1*8DIGIT

The client should monitor whether the server behaves as declared.

A client unique 'startup-id' is required until the client receives the session ID. The 'startup-id' is unique for a particular TCP connection. Pipelined requests using this header must be sent on the same TCP connection. The method through which this ID is generated is to be decided by the client.

5.5.4 Fast Content Switching

5.5.4.1 Introduction

In most cases, a content switch can be initiated with a single RTSP request. In order to preserve interoperability with RTSP aware intermediate devices such as application layer gateways, PSS clients should ensure that SETUP requests and responses are sent for each RTP/RTCP port pair to be used. Once a port pair has been negotiated, it may be reused for subsequent content upon a switch.

5.5.4.2 'Switch-Stream' RTSP Header Field

The 'Switch-Stream' header field may be used in an RTSP PLAY request or an RTSP PLAY response message. It is used to describe the replacement of media streams after a content switch. The 'Switch-Stream' header field may be used with aggregated control and with media control URLs.

The 'Switch-Stream' header syntax in ABNF [53] is as follows:

Switch-Stream = "Switch-Stream" COLON switch-spec *(COMMA switch-spec) CRLF

```
switch-spec     = old-stream ";" new-stream
old-stream      = "old" "=" (DQ rtsp-url DQ) / (DQ DQ)
new-stream      = "new" "=" (DQ rtsp-url DQ) / (DQ DQ)
rtsp-url        = as defined in RFC 2326 [5]
DQ              = %x22  ; US-ASCII double-quote mark (34)
LWS             = [CRLF] 1*( SP / HT )
SWS             = [LWS] ; sep whitespace
COMMA           = * ( SP / HT ) "," SWS; comma
COLON           = * ( SP / HT ) ":" SWS; colon
```

If both old media stream and new media stream URLs are indicated in the 'Switch-Stream' header field of a PLAY request from a PSS client to a PSS server, then the server shall interpret this as a request to replace the old media stream with the new media stream, hence reusing the transport parameters of the old media stream for the new media stream.

If the 'Switch-Stream' header field is included in a PLAY response from a PSS server to a PSS client, then this header informs the client about the media streams that are currently being streamed to the PSS client. The old media stream may be omitted in this case.

If only the new media stream URL is indicated in the 'Switch-Stream' header field of a PLAY request from a PSS client to a PSS server, then the PSS server shall interpret this as a request to switch to the new media stream. The PSS server decides the mapping. The PSS server shall indicate the SSRC of the new media stream in the RTP-Info of the reply, in order to enable the PSS client to locate the new stream.

If only the old stream URL is indicated in the 'Switch-Stream' header field of a PLAY request from a PSS client to a PSS server, then the PSS server shall interpret this as a request for complete removal of the specified media stream. The client and the server release the resources for this stream without explicit TEARDOWN signalling. The usage of the switch-stream header is defined in clauses 5.5.4.3, 5.5.4.4 and 5.5.4.7.

5.5.4.3 Switching to new content with available SDP

This clause defines all necessary PSS client and PSS server features for fast content switching where the UE already has the SDP for the new content locally available. The UE may have fetched the SDP file using RTSP DESCRIBE or HTTP GET or in any other method. Clients should assume that the herein defined fast content switching procedure is supported for all content items offered by this server. This PSS feature reduces the switching to new content to a single client-server interaction.

The feature-tag indicating this feature is '3gpp-switch'. The client should probe the server capabilities as early as possible in the communication using the '3gpp-switch' in the 'Supported' header as defined in 5.5.2.2.2. The client shall use the 'Require' header with this feature tag value, when requesting this behaviour from the server. The server shall use the PLAY method as defined in 5.5.2.4 with the '3gpp-switch' feature tag in the 'Require' header when the client requests this feature. Thus, the server replaces the current RTSP PLAY request by the new request resulting in a switch of streamed content.

When the PSS client wants to change the content of the RTSP session, the PSS client sends a PLAY request with the aggregated control URI of the new content to the PSS server. Note, the aggregated control URI is defined in the SDP file by the session level 'a=control:' attribute.

The PSS client shall add the media control URIs of the new streams in the 'Switch-Stream' header field to the RTSP PLAY method request. Whenever possible, the PSS client shall map the media control URIs of the same media type (e.g. audio or video) in the old content to the same media type of the new session. Note, this is only applicable for media types, which are present in the old and new content. The server includes always the 'Switch-Stream' header in the response. The 'Switch-Stream' header field is defined in 5.5.2.4.

If the SSRC have changed, then the server shall indicate the new SSRC values of the new media streams within the 'RTP-Info' header in the response. The SSRC entry for the 'RTP-Info' is defined in clause 5.5.2.3.

Note, if the new SDP contains more media components than the current session, the client may switch according to this section, describing the desired components in the 'Switch-Stream' header, and add the missing components using the method defined in 5.5.4.6.

If less media components are described in the new SDP than currently in use, the client and the server remove the component as defined in 5.5.4.7.

The entity-tag attribute ("a=etag" as specified in [5]) may be used in order to ensure that SDP information is used only if it is valid. When a PSS client requests a switch to content for which an entity-tag is available, it should include the 'If-Match' header and the etag value in the PLAY request. A PSS server shall validate the entity-tag in the If-Match headers prior to accepting the request. If the entity-tag is not valid, the server shall return either 412 (Precondition Failed) or if the client has previously communicated support for the "3gpp-switch-req-sdp" feature, the server may respond with a current SDP in an appropriate success response, as defined in Section 5.5.4.4.

5.5.4.4 Switching to new content without SDP

Clients should assume, that the here defined fast content switching procedure is supported for all content items offered by this server. The client uses the URL of the SDP file as content URL to describe the new content item.

Without an SDP or other adequate content description, the client is unable to specify the streams to which it wishes to subscribe. In order to initiate a content switch within a single RTSP round trip, the client may perform a PLAY request to initiate a switch via content URL without specifying individual streams. This allows the client to request that the server return the SDP, initiate a new session, setup all relevant media streams (or make an appropriate stream selection), and begin playback. The content URL used in the PLAY request is the same content URL used in a DESCRIBE. In order to signal that it wishes to receive the description and make a switch, the client shall include the 'SDP-Requested' header as defined below. This header is defined as follows:

SDP-Requested-Header = " SDP-Requested" COLON "1"

If a server receives a PLAY request and completes all actions successfully, the server responds with the SDP, Session-ID, RTP-Info, and a 'Switch-Stream' descriptor and begins streaming immediately. Whenever possible, the PSS server shall map the media control URIs of the same media type (e.g. audio or video) in the old content to the same media type of the new session. Note, this is only applicable for media types, which are present in the old and new content. The RTP-Info in the PLAY response must contain the SSRC for each stream as defined in 5.5.2.3. The server may issue a new session ID in the response, or it may re-use the existing session ID. The client must be prepared for either case.

If the server is not yet able to begin streaming, it responds with a 202 (Accepted) success code and with the SDP. The client may then perform a switch as described in 5.5.4.3 specifying the streams it would like to receive. This condition can occur if the server requires further client input regarding stream setup prior to beginning playback - for instance if the content requested contains multiple language switch groups and the server does not have the information necessary to choose a language.

If the server is not yet able to begin transmitting all the media streams, it can begin a subset of the streams and respond with a 206 (Partial Data) success code and the SDP. The 'Switch-Stream' header and the 'RTP-Info' header will indicate which streams have been selected for playback. The client may then add additional media components as described in 5.5.4.6.

If fewer media components are described in the new SDP than currently in use, then the server responds with a 200 (OK). The terminal shall remove the 'unused' media components as defined in clause 5.5.4.7.

The client and the server shall release the resources for the unused streams without explicit TEARDOWN signalling.

The feature tag '3gpp-switch-req-sdp' is defined to describe support for this feature. The client should probe the server capabilities as early as possible in the communication using the 'Supported' header as defined in 5.5.2.2 and shall use the 'Require' header with this feature tag value when requesting this behaviour from the server. The server shall use the PLAY method semantics defined in 5.5.2.4 when the client requests this feature.

5.5.4.5 Switching Media described in one SDP

Some content may be available for streaming in different representations. An example of such a use case is the live streaming of a sport event with multiple camera views. The SDP available at the receiver describes multiple options for one or several media types (e.g. video, audio, or subtitles). Upon initial setup of the session, the player (or the user) selects the preferred combination of the presentation to be consumed and sets up the corresponding media streams. At a later point, the user may trigger a switch to a different media stream carrying an alternative representation of the media.

The PLAY request is sent with the 'Switch-Stream' header field as defined in clause 5.5.4.2 indicating the URLs of both the old media stream and the new replacement stream. Upon receiving a PLAY request with a 'Switch-Stream' header field for an active session, a PSS server that supports this feature switches to the new media stream using the same transport parameters described in the initial SETUP request for the old media stream. After successfully processing the request, the PSS server shall reply with an 'RTP-Info' header indicating all active media streams in the changed session. The 'RTP-Info' header may include the SSRCs for each active media stream. The response may also include the 'Switch-Stream' header, indicating the stream switches that were successful. If the 'Switch-Stream' header field is not present in a successful response and the PSS server was identified to support the media switching functionality, the receiver should assume that all requested switches were successful.

The feature tag '3gpp-switch-stream' is defined to describe support for this feature. This feature tag is different than the feature tag '3gpp-switch' feature described in 5.5.4.3 indicating the support for content (aggregated stream) switch. The client should use the 'Require' header with this feature tag value when requesting this behaviour from the server. The server shall use the PLAY method semantics as defined in 5.5.2.4 when the client requests this feature. Note that several media streams of a presentation may be switched at the same time in a single PLAY request.

5.5.4.6 Adding Media Components to an ongoing session

It may happen that the new content stream consists of more media components than the ongoing content stream. In such a case, the client is recommended to switch to the new content with the already established resources and add further components afterwards.

The client should pipeline the setup requests for the new components after the content switching request (see clause 5.5.4). The client shall issue a PLAY request to start all addition media components without interrupting the existing. If the client and server support the "3gpp-pipelined" feature (see clause 5.5.3), then the client shall pipeline the PLAY request with the SETUP requests. The 'RTP-Info' header contains the synchronization information for all media components.

The session id value of the already established session shall be part of the SETUP request header to indicate the relation of the media component to the already established components.

5.5.4.7 Removing Media Components from an ongoing session

A PSS client wishing to terminate the streaming of a specific media stream shall send a PLAY request with a 'Switch-Stream' header indicating the URL of the media stream to be torn down as the old media stream. No URL for the new media stream should be specified.

Upon receiving a PLAY request with 'Switch-Stream' header field indicating that one or more media streams are to be terminated, the server shall stop streaming the indicated media streams and release the used UDP ports for this media component and free the associated resources. However, the other media streams should not be interrupted.

After successfully processing the request, the server shall reply with a success response message and a 'Session' header field, even if the session contains no more media streams.

The PSS client shall only use TEARDOWN to completely tear down the whole session.

5.6 Extension for Time-Shifting Support

5.6.1 Introduction

Time shifting functionality is designed to enhance the access to live streaming sessions. For this reason, the PSS server maintains a time-shift buffer for each live feed. The server side timeshift buffer allows the PSS client to pause live sessions and even navigate (rewind, fast forward) in the offered time-shift buffer range. A timeshift supporting PSS client, which is connected to a timeshift supporting PSS server is able to perform some or all of the following operations on timeshifted streaming sessions:

- Pause and resume the playout at a later point in time

- Start playout from (or seek to) a position in the stream that corresponds to a past time instant in the live streaming session
- Perform operations such as Fast and Slow Forward or Rewind (i.e. Trick Mode).

NOTE: the live streaming feed is not necessarily offered by the same PSS server as the time-shifted session. For example, the PSS server may be handling time-shifting services for a live feed accessible as an MBMS streaming service.

5.6.2 General Description

The PSS timeshift functionality requires the availability of one or more timeshift buffers on the server side. The size of the timeshift buffer is determined by the server. This specification does not limit the size of the timeshift buffer.

The timeshift buffer is defined by an upper and a lower range. New live data is added at the upper range to the timeshift buffer. If the timeshift buffer progresses as sliding window, then the server removes and discards data from the lower range of the timeshift buffer. The upper range of the timeshift buffer is also referred to as current recording time.

The PSS server announces the availability of the timeshift feature and describes the current state of the time-shift buffer in the RTSP SETUP response. The PSS server provides the current recording time and the current time-shift buffer ranges with each RTSP PLAY and PAUSE response messages. The server may report updated timeshift buffer ranges throughout the session via SET_PARAMETER messages and the client may request this information via GET_PARAMETER requests.

In the following, two PSS timeshift use-cases are described to clarify the operations of PSS timeshift.

Usecase A: A user starts a live session using 'range: npt=now-'. The user enters timeshift operation at a later point by pressing 'pause'. The PSS client stores the value of the 'current recording time' header, which is received with the pause response message to resume the session. The PSS client is aware of the server side timeshift buffer ranges.

Usecase B: A user may directly start a live session in a timeshift operation, for instance when changing from broadcast to reception of the same content. The PSS client may use absolute time representation (clock) to continue consuming the content at the same media position.

5.6.2a Extensions to RTSP 1.0

At least one PSS timeshift buffer is provided by the PSS server for at least one live stream. The PSS server may also provide individual timeshift buffers per RTSP session. The PSS server announces the availability of the timeshift feature and describes the current range of the time-shift buffer in the RTSP SETUP response.

The PSS client should always calculate current valid boundaries of the timeshift buffer.

The PSS timeshift feature may be used with either NPT or UTC time ranges. PSS clients that support time shifting shall support also UTC time in addition to NPT.

If the PSS server supports PSS timeshifting as defined in this section, then the new RTSP headers 3GPP-TS-CurrentRecording-Time and 3GPP-TS-Buffer shall be present in all server to client messages.

The Accept-Ranges header field may be used to check for the support of UTC time. The PSS server indicates the preferred media range format for time shifting in the 3GPP-TS-CurrentRecording-Time and 3GPP-TS-Buffer headers in the PLAY response. PLAY requests may contain the 'npt=now-' range indication to seek to the upper range of the time shift buffer. Other time shifting operations shall consistently use the same media range format that is indicated by the PSS server.

5.6.3 Accept-Ranges

The Accept-Ranges request and response-header field allows indication of the format supported in the Range header. The PSS client shall include the header in SETUP requests to indicate which formats it supports in PLAY and PAUSE responses and REDIRECT requests. The server shall include the header in the SETUP response and in any error response caused by an unaccepted range format, to indicate the formats supported for the resource indicated by the request URI.

This header has the following ABNF syntax:

Accept-Ranges = "Accept-Ranges" HCOLON acceptable-ranges CRLF

acceptable-ranges = range-unit *(COMMA range-unit)

range-unit = "npt" / "utc"

5.6.4 Signalling Time Shifting Ranges

In order to allow the PSS server to provide the PSS client with the current ranges of the time shift buffer, two new RTSP headers are defined.

The PSS client may start a PSS session already in 'timeshift mode'.
The PSS server provides the upper bound of the timeshift buffer with the '3GPP-TS-CurrentRecording-Time' header. The PSS client should not request playback beginning beyond the current recording time (i.e. no future playback times). If a PSS server receives a PLAY request outside of the time shift buffer range, the PSS server should handle the request as a request for the appropriate buffer boundary time. In case of a range request exceeding the range end time, this will be the end time; in case of a range request beginning earlier than the buffer start time, this will be the start time. The actual range streamed will then be reported in the 200 OK response message.

The PSS server describes the current available range for PSS time shifting with the '3GPP-TS-Buffer' header. This header may use one of three descriptive formats, depending on the current state of the time-shift buffer. The client is responsible for keeping track of the available timeshift buffer boundaries.

The 'buffer-depth' parameter indicates the depth of the timeshift buffer in seconds. When this format is indicated, the timeshift buffer is constantly filled to the specified depth; the timeshift buffer is progressed as a sliding window. The PSS client calculates together with the information from the '3GPP-TS-CurrentRecording-Time' header the lower and upper range of the timeshift buffer. The PSS client shall continuously update the upper and lower boundary of the timeshift buffer.

The 'interval' parameter can either indicate a closed range (two value) or an open range (one value) indicating absolute times for the timeshift buffer ranges. No timeshift data is available earlier than the left range value of the 3GPP-TS-Buffer parameter and, if the upper range is present, time shifting is not available beyond that time. If this value indicates an open range (one absolute start time for the timeshift recording), then the PSS server has not given any end-time for timeshift recording.

When both parameters are present, this indicates that the timeshift buffer is still being established. Recording has started at the lower bound of the interval and will progress until the specified depth is reached. Once the buffer depth is reached, then the timeshift buffering continues in a sliding window as described above and the PSS server uses the buffer header parameter for timeshift buffer reporting.

The ABNF syntax for the new headers is as follows:

```
3GPP-TS-CurrentRecording-Time='3GPP-TS-CurrentRecording-Time' COLON (npt-time-indication / utc-time-indication) CRLF

npt-time-indication = 'npt=' (npt-sec / npt-hhmmss)

utc-time-indication = 'clock=' utc-time

3GPP-TS-Buffer = '3GPP-TS-Buffer' COLON (depth / interval / interval_depth) [SEMIParameter-Ext] CRLF

depth = 'buffer-depth=' 1*DIGIT

interval = utc-range / npt-range

interval_depth = interval SEMI depth

Parameter-Ext = (1*DIGIT ['.' 1*DIGIT]) / (1*((0x21..0x2b) / (0x2d..0x3a) / (0x3c..0x7a) / 0x7c / 0x7e))

COLON = < as defined in clause 5.5.4.2 >

SEMI = SWS ";" SWS ; semicolon;

SWS = < as defined in clause 5.5.4.2>
```

Note: utc-range, npt-range, utc-time, npt-sec, and npt-hhmmss are defined in [5]RFC2326.

5.6.5 Timeshift buffer status updates

The GET_PARAMETER and the SET_PARAMETER methods allow the client and the server to synchronize timeshift buffer information.

The PSS client may query the PSS server at any time for the current timeshift buffer fill states. For this procedure, the PSS Client issues the RTSP GET_PARAMETER request message with the 3GPP-TS-Buffer and/or 3GPP-TS-CurrentRecording-Time tags in the message body. The MIME Type 'text/plain' is used for the RTSP message body.

The GET_PARAMETER response shall carry the 3GPP-TS-Buffer and the 3GPP-TS-CurrentRecording headers in the header fields and also in the RTSP response message body.

The server may update the current timshift information using the RTSP SET_PARAMETER method (S->C). The SET_PARAMETER request message shall carry the 3GPP-TS-Buffer and the 3GPP-TS-CurrentRecording headers in the header fields and also in the RTSP request message body. The PSS client shall confirm the reception with a '200 OK' RTSP response message or, if it client does not understand the parameters, it shall return '451 Invalid Parameters' and the PSS server shall not attempt further updates.

5.7 Support for Trick Mode Operations

PSS client and server may support trick mode operations such as fast/slow forward and rewind. The level of scale support may also depend on the selected streaming content.

A PSS Client may use the 'play.scale' feature tag to query for support of scaled playout. The 'play.scale' feature tag applies only to PSS servers and indicates the support for scale operations for media streaming.

If the server supports the scale value requested, and the content is capable of the scale value requested, the server shall serve the scaled stream to the client. If the server does not support the scale value requested or the content does not support the scale value requested then the server shall decide what level of scaled playout to serve to the client. When trick mode operations are supported, the 'Scale' header field of RTSP [5] shall be used for requesting trick mode operations.

Streaming with a scale value other than 1 should not change the streaming bitrate of the corresponding media stream.

The PSS Client is made aware of the scale values supported for the content via the 'X-Scale' media level attribute, or by using the 'scales' parameter in the GET_PARAMETER and SET_PARAMETER. The 'scales' parameter takes precedence in the case of conflicting values. The PSS server may temporarily omit the transmission of media data for one or more media streams on which playout with the selected scale is not possible. For example, this may mean that the video is scaled at a rate of 2, but the audio is omitted for the duration of the scaled playout as it is unable to be scaled.

A PLAY request with a 'Scale' indication that is not supported by any of the media streams of the streaming session may be ignored by the PSS server. The PSS server should use the closest supported scale value instead and it shall indicate the chosen value in the response, unless the selected scale value is 1.

The 'X-Scale' SDP attribute and the 'scales' parameter are defined according to the following ABNF syntax:

```
Scale="X-Scale:" payload_type SP scale_value *(";" scale_value) CRLF

scales="scales:" scale-spec *("," scale-spec) CRLF

scale-spec= stream-url "=" scale *(";" scale)

scale_value=["-"] 1*DIGIT [ "." *DIGIT]
```

In the SDP, the payload_type indicates the payload type of the media to which the scale values apply. The stream-url indicates the URL of the media stream to which the indicated scales are applicable. The scale_value is a decimal number that indicates the possible scale value that may be requested for the specified media stream. For a single media stream, multiple scale values are possible.

When timeshifting is used, the following applies:

- in the case that the PSS Client knows that, due to the PSS Server"s buffer, a particular request for scaled playback is impossible to complete, the PSS Client shall not request this scale value.

- in the case that the PSS Server is providing scaled playback and a buffer limit is reached, the PSS Server shall return the scale value of the stream to 1.

- the PSS client should take into account the playout scale when calculating its current position in the time shifting buffer, and take the time shifting buffer into account when requesting a playout scale.

6 Data transport

6.1 Packet based network interface

PSS clients and servers shall support an IP-based network interface for the transport of session control and media data. Control and media data are sent using TCP/IP [8] and UDP/IP [7]. An overview of the protocol stack can be found in figure 2 of the present document.

6.2 RTP over UDP/IP

6.2.1 General

The IETF RTP [9] provides means for sending real-time or streaming data over UDP (see [7]). The encoded media is encapsulated in the RTP packets with media specific RTP payload formats. RTP payload formats are defined by IETF. RTP also provides a protocol called RTCP (see clause 6 in [9]) for feedback about the transmission quality.

RTP/UDP/IP transport of speech, audio and video shall be supported. RTP/UDP/IP transport of timed text should be supported. Sending of RTCP shall be performed according to the used RTP profile, indicated RTCP bandwidth, and other RTCP related parameters. The transmission times of RTCP shall be controlled by algorithms performing as the ones specified in the RTP specification [9], and if AVPF is used according to [57]. For information on how the RTCP transmission interval depends on different values of the RTCP parameters, see Annex A.3.2.3.

6.2.2 RTP profiles

For RTP/UDP/IP transport of continuous media the following RTP profile shall be supported:

- RTP Profile for Audio and Video Conferences with Minimal Control [10], also called RTP/AVP;

For RTP/UDP/IP transport of continuous media the following RTP profile should be supported:

- Extended RTP Profile for RTCP-based Feedback (RTP/AVPF) [57], also called RTP/AVPF. A PSS client or server shall support the generic NACK message specified in section 6.2.1 of [57] if RTP retransmission is supported. A PSS client or server is not required to support the other feedback formats specified in section 6 of [57].

Clause A.3.2.3 in Annex A of the present document provides more information about the minimum RTCP transmission interval.

For integrity protected RTP/UDP/IP transport of continuous media, the following RTP profile should be supported:

- The Secure Real-time Transport Protocol (SRTP) [72], also called RTP/SAVP.

6.2.3 RTP and RTCP extensions

6.2.3.1 RTCP extended reports

A PSS client should implement the framework and SDP signalling of the RTP Control Protocol Extended Reports [58]. A PSS client should further implement the following report formats:

- Loss RLE Report Block defined in section 4.1of [58].

A PSS client should send the report block(s) indicated by SDP signalling from the PSS server. A PSS server may limit the report blocks size using SDP signalling. For best utility the client should report in every packet and provide redundancy by reporting also on past RTCP intervals. In cases where size restrictions prevent the client from both reporting on all the RTP packets and providing redundancy, the client shall stop the redundant reporting to address this restriction. If this action is still not enough to reduce the reports to satisfactory sizes, the client may then choose not to send the report in every packet.

6.2.3.2 RTCP App packet for client buffer feedback (NADU APP packet)

A PSS client supporting Signalling for Client Buffer Feedback (see clause 10.2.3) shall report the next application data unit to be decoded for buffer status reporting and rate adaptation by sending the RTCP APP packet. A NADU APP packet shall be sent only after the client has received at least one RTP packet on the media stream and shall be accompanied by a complementary RR packet. The RR and NADU packets shall contain information that represents a single simultaneous 'snapshot' of the media stream. The format of a generic RTCP APP packet is shown in Figure 3 below:

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|V=2|P| subtype |   PT=APP=204  |             length            |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                           SSRC/CSRC                           |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                          name (ASCII)                         |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                   application-dependent data                ...
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 3: Generic Format of an RTCP APP packet.

For rate adaptation the name and subtype fields must be set to the following values:

name: The NADU APP data format is detected through the name "PSS0", i.e. 0x50535330 and the subtype.
subtype: This field shall be set to 0 for the NADU format.
length: The number of 32 bit words –1, as defined in RFC 3550 [9]. This means that the field will be 2+3*N, where N is the number of sources reported on. The length field will typically be 5, i.e. 24 bytes packets.
application-dependent data: One or more of the following data format blocks (as described in Figure 4) can be included in the application-dependent data location of the APP packet. The APP packets length field is used to detect how many blocks of data are present. The block shall be sent for the SSRCs for which there are a report block as part of either a Receiver Report or a Sender Report, included in the RTCP compound packet. A NADU APP packet shall not contain any other data format than the one described in figure 4 below.

```
0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                             SSRC                              |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|      Playout Delay            |              NSN              |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|  Reserved           |   NUN   |    Free Buffer Space (FBS)    |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 4: Data format block for NADU reporting

SSRC: The SSRC of the media stream the buffered packets belong to.

Playout delay (16 bits): The difference in milliseconds between the scheduled playout time of the next ADU to be decoded, (whose sequence number is indicated in the NSN field) and the current time when generating the RTCP packet that contains the NADU APP block, both measured on the media playout clock. The client shall always indicate this value, unless it is not well defined, when it may use the reserved value (0xFFFF). When the buffer is empty (the client has not yet received the packet with sequence number NSN), the playout delay is not well defined and the client should use the reserved value 0xFFFF for this field. When the media clock is not advancing (e.g. while paused or re-buffering), the playout delay corresponds to the difference between the playout time of the next ADU and the media time at which playout will resume.

The point at which the media playout clock is measured should be chosen such that, if the only packet in the buffer is that with sequence number NSN, the playout delay indicates the time remaining until the media playout will 'starve' and this stream might need re-buffering. In the calculations of playout delay above, this point is used to determine the playout point of a media packet even though actual playout may occur later in the decoding chain. The target buffer time (see clause 5.3.2.2) must be measured from the same point.
The playout delay allows the server to have a more precise value of the amount of time before the client will underflow. The playout delay shall be computed until the actual media playout (i.e., audio playback or video display).

NSN (16 bits): The RTP sequence number of the next ADU to be decoded for the SSRC reported on. In the case where the buffer does not contain any packets for this SSRC, the next not yet received sequence number shall be reported, i.e. an NSN value that is one larger than the least significant 16 bits of the RTCP SR or RR report block's "extended highest sequence number received".

NUN (5 bits): The unit number (within the RTP packet) of the next ADU to be decoded. The first unit in a packet has a unit number equal to zero. The unit number is incremented by one for each ADU in an RTP packet. In the case of an audio codec, an ADU is defined as an audio frame. In the case of H.264 (AVC), an ADU is defined as a NAL unit. In the case of H.263 and MPEG4 Visual Simple Profile, an ADU is defined as a whole or a part of an H.263 video picture or MPEG4 VOP that is included in a RTP packet. In the specific case of H.263, each packet carries a single ADU and the NUN field shall be thus set to zero. Future additions of media encoding or transports capable of having more than one ADU in each RTP payload shall define what shall be counted as an ADU for this format.

FBS (16 bit): The amount of free buffer space available in the client at the time of reporting. The reported free buffer space shall be less than or equal to the buffer space that has been reported as available for adaptation by the 3GPP-Adaptation RTSP header, see clause 5.3.2.2. The amount of free buffer space are reported in number of complete 64 byte blocks, thus allowing for up to 4194304 bytes to be reported as free. If more is available, it shall be reported as the maximal amount available, i.e. 4194304 with a field value 0xffff.

Reserved (11 bits): These bits are not used and shall be set to 0 and shall be ignored by the receiver.

6.2.3.3 RTP retransmission

6.2.3.3.1 General

A PSS client should implement RTP retransmission. A PSS server should implement RTP retransmission. A PSS client or server implementing RTP retransmission shall implement the payload format, SDP signalling and mechanisms of the

RTP retransmission payload format [81]. In addition to the specifications and recommendations in [81], a PSS client and server supporting RTP retransmission shall follow the definitions in the following clauses.

6.2.3.3.2 Multiplexing scheme

The RTP retransmission payload format [81] provides two different schemes for multiplexing the original and the retransmission stream, i.e. session-multiplexing and SSRC-multiplexing. PSS servers shall use SSRC-multiplexing and shall not use session-multiplexing.

6.2.3.3.3 RTCP retransmission request

PSS clients shall use the NACK feedback message format defined in the "Extended RTP Profile for RTCP-based Feedback (RTP/AVPF)" [57] for requesting the retransmission of RTP packets.

Before requesting the retransmission of RTP packets the client should assess whether a requested packet can be decoded in time by checking the latest receiver buffer status. If the client sends RTCP APP packets for client buffer feedback, as defined in section 6.2.3.2, the same assessment should be performed by the server, according to the latest RTCP APP packet it has received.

6.2.3.3.4 Congestion control and usage with rate adaptation

To avoid network congestion due to the additional bandwidth required for the retransmission of lost packets, a PSS server or client implementing RTP retransmission shall estimate the available link rate and adapt the total transmission rate of the RTP session, including retransmissions, to the available link rate. The actual algorithms providing link-rate estimation and transmission-rate adaptation are implementation specific. Rules and information sources for the estimation of the available link rate are described in clause 10.2.1 of the present document. To adapt the total transmission rate including retransmissions, a PSS server can e.g. skip retransmissions, use the transmission rate adaptation described in clause 10.2.2 of the present document or use any other suitable method.

If the server uses multiple streams for rate adaptation, the server may receive retransmission requests for a stream that is different from the one it is currently using. The server should thus not flush its retransmission buffer after switching streams.

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;
- Extended AMR-WB codec (see clause 7.3) RTP payload format according to [85];
- Enhanced aacPlus and MPEG-4 AAC codec (see clause 7.3) RTP payload format according to [13]; the size of audioMuxElements shall be limited to the maximum size of one audio frame, which is 6144 bits per AAC channel; moreover multiplexing of multiple audio frames into one audioMuxElement should be avoided if this would lead to fragmentation across RTP packets;
- MPEG-4 Visual 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 4629 [14];
- H.264 (AVC) video codec (see clause 7.4) RTP payload format according to [92]. 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 [90]. 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;
- 3GPP timed text format (see clause 7.9) RTP payload format according to [80];

- DIMS (see subclause 8.3) RTP payload format according to [98];
- DRM encrypted RTP payload format according to clause K.1 in Annex K;
- RTP retransmission payload format according to [81].

NOTE: The payload format RFC 3016 for enhanced aacPlus and 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 enhanced aacPlus and/or AAC shall follow the changes to the LATM format included in [21]. It should be noted further that the enhanced aacPlus signalling mode 'backwards compatible explicit signalling' (as defined in [21]) can not be used with LATM.

6.3 HTTP over TCP/IP

The IETF TCP provides reliable transport of data over IP networks, but with no delay guarantees. It is the preferred way for sending the scene description, text, bitmap graphics and still images. There is also need for an application protocol to control the transfer. The IETF HTTP [17] provides this functionality.

HTTP/TCP/IP transport shall be supported for:

- still images (see clause 7.5);
- bitmap graphics (see clause 7.6);
- synthetic audio (see clause 7.3A);
- vector graphics (see clause 7.7);
- text (see clause 7.8);
- timed text (see clause 7.9);
- SMIL scene description (see clause 8);
- presentation description (see clause 5.3.3).

HTTP/TCP/IP transport should be supported for:

- 3GP files for download and progressive download (see clause 7.10).

NOTE: DIMS scene description can be provided either over RTP/UDP/IP (see clauses 5.4 and 6.2.4) or contained in a 3GP file downloaded over HTTP/TCP/IP (see clauses 5.4 and this clause).

6.4 Transport of RTSP

Transport of RTSP shall be supported according to RFC 2326 [5].

7 Codecs

7.1 General

For PSS clients supporting a particular media type, corresponding media decoders are specified in the following clauses.

7.2 Speech

If speech is supported, the AMR decoder shall be supported for narrow-band speech [18][63][64][65]. The AMR wideband speech decoder, [20][66][67][68], shall be supported when wideband speech working at 16 kHz sampling frequency is supported.

7.3 Audio

If audio is supported, then one or both of the following two audio decoders should be supported:

- Enhanced aacPlus [86] [87] [88]
- Extended AMR-WB [82] [83] [84]

Specifically, based on the audio codec selection test results Extended AMR-WB is strong for the scenarios marked with blue, Enhanced aacPlus is strong for the scenarios marked with orange, and both are strong for the scenarios marked with green colour in the table below:

Content type Bit rate	Music	Speech over Music	Speech between Music	Speech
14 kbps mono				
18 kbps stereo				
24 kbps stereo				
24 kbps mono				
32 kbps stereo				
48 kbps stereo				

More recent information on the performance of the codecs based on more recent versions of the codecs can be found in TR 26.936 [95].

Enhanced aacPlus decoder is also able to decode AAC-LC content.

Extended AMR-WB decoder is also able to decode AMR-WB content.

In addition, MPEG-4 AAC Low Complexity (AAC-LC) and MPEG-4 AAC Long Term Prediction (AAC-LTP) object type decoders [21] may be supported. The maximum sampling rate to be supported by the decoder is 48 kHz. The channel configurations to be supported are mono (1/0) and stereo (2/0).

When a server offers an AAC-LC or AAC-LTP stream with the specified restrictions, it shall include the 'profile-level-id' and 'object' MIME parameters in the SDP 'a=fmtp' line. The following values shall be used:

Object Type	profile-level-id	object
AAC-LC	15	2
AAC-LTP	15	4

7.3a Synthetic audio

If a PSS client supports synthetic audio both the Scalable Polyphony MIDI (SP-MIDI) content format defined in Scalable Polyphony MIDI Specification [44] and the device requirements defined in Scalable Polyphony MIDI Device 5-to-24 Note Profile for 3GPP [45] should be supported.

SP-MIDI content is delivered in the structure specified in Standard MIDI Files 1.0 [46], either in format 0 or format 1.

In addition a PSS client supporting synthetic audio should also support both the Mobile DLS instrument format defined in [70] and the Mobile XMF content format defined in [71].

A PSS client supporting Mobile DLS shall meet the minimum device requirements defined in [70] in section 1.3 and the requirements for the common part of the synthesizer voice as defined in [70] in sections 1.2.1.2. If Mobile DLS is supported, wavetables encoded with the G.711 A-law codec (wFormatTag value 0x0006, as defined in [70]) shall also be supported. The optional group of processing blocks as defined in [70] may be supported. Mobile DLS resources are delivered either in the file format defined in [70], or within Mobile XMF as defined in [71]. For Mobile DLS files delivered outside of Mobile XMF, the loading application should unload Mobile DLS instruments so that the sound bank required by the SP-MIDI profile [45] is not persistently altered by temporary loadings of Mobile DLS files.

Content that pairs Mobile DLS and SP-MIDI resources is delivered in the structure specified in Mobile XMF [71]. As defined in [71], a Mobile XMF file shall contain one SP-MIDI SMF file and no more than one Mobile DLS file. PSS clients supporting Mobile XMF must not support any other resource types in the Mobile XMF file. Media handling behaviours for the SP-MIDI SMF and Mobile DLS resources contained within Mobile XMF are defined in [71].

7.4 Video

If a PSS client supports video, ITU-T Recommendation H.263 Profile 0 Level 45 decoder [22] shall be supported. In addition, a PSS client should support:

- H.263 Profile 3 Level 45 decoder [22];
- MPEG-4 Visual Simple Profile Level 3 decoder [24] with the following constraints:
    - The number of Visual Objects supported shall be limited to 1.
    - The maximum frame rate shall be 30 frames per second;
    - The maximum f_code shall be 2;
    - The intra_dc_vlc_threshold shall be 0;
    - The maximum horizontal luminance pixel resolution shall be 352 pels/line;
    - The maximum vertical luminance pixel resolution shall be 288 pels/VOP;
    - If AC prediction is used, the following restriction applies: QP value shall not be changed within a VOP (or within a video packet if video packets are used in a VOP). If AC prediction is not used, there are no restrictions to changing QP value;
- H.264 (AVC) Constrained Baseline Profile Level 1.3 decoder [90] without requirements on output timing conformance (Annex C of [90]).

If H.264 (AVC) High Profile is supported by a PSS client, the decoder shall support decoding any stream compliant to H.264 (AVC) High Profile Level 3.0 [43] with frame_mbs_only_flag=1, without requirements on output timing conformance (Annex C of ITU-T Recommendation H.264 [43]).

NOTE: H.264 (AVC) Main Profile is a subset of H.264 High Profile, and a High Profile decoder is required to be able to decode Main Profile streams.

The video buffer model given in Annex G of the present document should be supported if H.263 or MPEG-4 Visual 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 5.3.3.2, and in RTSP, as specified in clause 5.3.2.4. A PSS client supporting H.264 (AVC) shall also use the RTP timestamp or NALU-time (as specified in [92]) 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 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 Visual (short header), as MPEG-4 Visual is not mandated by PSS.

7.5 Still images

If a PSS client supports still images, ISO/IEC JPEG [26] together with JFIF [27] decoders shall be supported. The support requirement for ISO/IEC JPEG only applies to the following two modes:

- baseline DCT, non-differential, Huffman coding, as defined in table B.1, symbol 'SOF0' in [26];
- progressive DCT, non-differential, Huffman coding, as defined in table B.1, symbol 'SOF2' [26].

7.6 Bitmap graphics

If a PSS client supports bitmap graphics, the following bitmap graphics decoders should be supported:

- GIF87a, [32];
- GIF89a, [33];
- PNG, [38].

7.7 Vector graphics

If a PSS client supports vector graphics, SVG Tiny 1.2 [42] [43] and ECMAScript [94] shall be supported.

NOTE 1: The compression format for SVG content is GZIP [59], in accordance with the SVG specification [42].

NOTE 2: Only codecs and MIME media types supported by PSS, as specified in clause 7 and in subclause 5.4, respectively, shall be used. In particular, PSS clients are not required to support the Ogg Vorbis format.

NOTE 3: Content creators of SVG Tiny 1.2 are strongly recommended to follow the content creation guidelines provided in Annex L.

NOTE 4: A DIMS client is capable of processing SVG Tiny 1.2 data.

7.8 Text

The text decoder is intended to enable formatted text in a SMIL presentation.

If a PSS client supports text it shall support

- text formatted according to XHTML Mobile Profile [47];
- rendering a SMIL presentation where text is referenced with the SMIL 2.0 "text" element together with the SMIL 2.0 "src" attribute.

If text is supported, the following character coding formats shall be supported:

- UTF-8, [30];
- UCS-2, [29].

NOTE: Since both SMIL and XHTML are XML based languages it would be possible to define a SMIL plus XHTML profile. In contrast to the presently defined SMIL Language Profile that only contain SMIL modules, such a profile would also contain XHTML modules. No combined SMIL and XHTML profile is specified for PSS. Rendering of such documents is out of the scope of the present document.

7.9 Timed text

PSS clients supporting timed text shall support [51]. Timed text may be transported over RTP or downloaded contained in 3GP files using Basic profile.

NOTE: A PSS client supporting timed text shall receive and parse 3GP files containing the text streams. This does not imply a requirement on PSS clients to be able to render other continuous media types contained in 3GP files, e.g. AMR and H.263, if such media types are included in a presentation together with timed text. Audio and video are instead streamed to the client using RTSP/RTP (see clause 6.2).

7.10 3GPP file format

3GP files [50] can be used by both PSS clients and PSS servers. The following profiles are used:

- Basic profile shall be supported by PSS clients if timed text is supported;
- Basic profile, Extended-presentation profile and Progressive-download profile should be supported by PSS clients;
- Streaming server profile should be supported by PSS servers.

7.11 Timed graphics

PSS clients supporting timed graphics shall support 3GPP TS 26.430 [109].

8 Scene description

8.1 General

There are several options for scene description in PSS:

- SMIL presentation, where the SMIL file is provided on its own or included as a primary item in a 3GP file (Extended-presentation profile);
- DIMS, where DIMS is included as a track in a 3GP file (Basic profile) or as a primary item (possibly in combination with a track) in a 3GP file (Extended-presentation profile).

The usage of SMIL and DIMS in 3GP files is defined in [50]. For pure RTSP/RTP-based streaming and 3GP files containing continuous media only, no separate scene description is required.

8.2 Synchronised Multimedia Integration Language

The 3GPP PSS uses a subset of SMIL 2.0 [31] as format of the scene description. PSS clients and servers with support for scene descriptions shall support the 3GPP SMIL Language Profile defined in [52]. This profile is a subset of the SMIL 2.0 Language Profile, but a superset of the SMIL 2.0 Basic Language Profile. Document [52] also includes an informative Annex A that provides guidelines for SMIL content authors.

NOTE: The interpretation of this is not that all streaming sessions are required to use SMIL. For some types of sessions, e.g. consisting of one single continuous media or two media synchronised by using RTP timestamps, SMIL may not be needed.

8.3 Dynamic and Interactive Multimedia Scenes

The 3GPP PSS uses DIMS [98] as a format of the scene description. PSS clients and servers with support for scene description shall support DIMS Mobile Profile at level 10 [98].

9 3GPP file format (interchange format for MMS)

The 3GPP file format is defined in [50].

10 Adaptation of continuous media

10.1 General

The PSS includes a number of protocols and functionalities that can be utilized to allow the PSS session to adapt transmission and content rates to the available network resources. The goal of this is of course to achieve highest possible quality of experience for the end-user with the available resources, while maintaining interrupt-free playback of the media. This requires that the available network resources are estimated and that transmission rates are adapted to the available network link rates. This can prevent overflowing network buffers and thereby avoid packet losses. The real-time properties of the transmitted media must be considered so that media does not arrive too late to be useful. This will require that media content rate is adapted to the transmission rate.

To avoid buffer overflows, resulting in that the client must discard useful data, while still allowing the server to deliver as much data as possible into the client buffer, a functionality for client buffer feedback is defined. This allows the server to closely monitor the buffering situation on the client side and to do what it is capable in order to avoid client buffer underflow. The client specifies how much buffer space the server can utilize and the minimum target level of protection the client perceives necessary to provide interrupt-free playback. Once this desired level of target protection is achieved, the server may utilize any resources beyond what is needed to maintain that protection level to increase the quality of the media or, the server may choose to leave the transmission rate alone and simply accrue additional time in the client buffer at the present rate. The server can also utilize the buffer feedback information to decide if the media quality needs to be lowered in order to avoid a buffer underflow and the resulting play-back interruption.

10.2 Bit-rate adaptation

The bit-rate adaptation for PSS is server centric in the meaning that transmission and content rate are controlled by the server. The server uses RTCP and RTSP as the basic information sources about the state of the client and network. This allows link-rate adaptation also when communicating with PSS clients of earlier releases, as long as they send RTCP receiver reports frequently enough.

10.2.1 Link-rate estimation

The actual algorithm providing the link-rate estimation is implementation specific. However, this chapter describes and gives rules for the different information sources that can be used for link-rate estimation.

10.2.1.1 Initial values

A PSS client should inform the server the quality of service parameters for the used wireless link. The known parameters should be included in the RTSP "3GPP-Link-Char" header (chapter 5.3.2.1) in either the RTSP SETUP or PLAY request. This enables the server to set some basic assumption about the possible bit-rates and link response. If the client has initially reported these parameters and they are changed during the session the client shall update these parameters by including the "3GPP-Link-Char" header in a SET_PARAMETER or OPTIONS request.

A PSS client should inform the server about initial bit-rate available over the link, if known. This reporting shall be done using the RTSP "Bandwidth" header in either the RTSP SETUP or PLAY request. The QoS negotiated guaranteed bit-rate is the best estimate for the bandwidth value.

10.2.1.2 Regular information sources

The basic information source giving regular reports useful for bit-rate estimations is the RTCP receiver reports as defined by [9]. The RTCP reporting interval is dependent on the RTP profile in use, the bit-rate assigned to RTCP, the average size of RTCP packets, and the number of reporting entities. Most of these parameters can be set or affected by the PSS server through signalling. This allows the server to configure the reporting interval to a desirable working point. See chapter 5.3.3.1 for specification on how the RTCP bandwidth is signalled by the server.

In most PSS RTP sessions the server and the client only have one SSRC each, thus providing the highest possible reporting rate. However some scenarios could result in that the number of used SSRC is larger, thereby possibly lowering the effective reporting interval for client, server or both.

The average size of the RTCP packets cannot be tightly controlled, but a loose control is possible by controlling which RTCP packet types that are used. This will depend on which of the below-listed RTCP extensions are in use.

The PSS server can signal the PSS client in SDP, to request that "Loss RLE Report Block" in RTCP XR (section 6.2.3) are used to report packet loss vectors.

10.2.2 Transmission adaptation

The transmission adaptation is implementation dependent. The 3GPP file format server extensions [50] provide a server the possibility to store alternative encodings useful for stream switching.

A server doing transmission rate adaptation through content rate adaptation shall still deliver content according to the SDP description of the media streams, e.g. a video stream delivered after content rate adaptation must still belong to the SDP announced profile and be consistent with any configuration. This will either put restrictions on the possible alternatives or require declaration of several RTP payload types or media encodings that might not be used.

10.2.3 Signalling for client buffer feedback

The client buffer feedback signalling functionality should be supported by PSS clients and PSS servers. For PSS clients and servers that support the client buffer feedback signalling functionality, the following parts shall be implemented:

- SDP service support, as described in clause 5.3.3.5.
- The size (in bytes) of the buffer the client provides for rate adaptation. It is signalled to the server through RTSP, as described in clause 5.3.2.2
- The target buffer protection time (in milliseconds). It is signalled to the server through RTSP, as described in clause 5.3.2.2.
- The client buffer status feedback information, including free buffer space, next ADU to be decoded and playout delay. It is signalled to the server via RTCP, as described in clause 6.2.3.2.

If a PSS server supports client buffer feedback, it shall include the attribute "3GPP-Adaptation-Support" in the SDP, as described in clause 5.3.3.5. If a PSS client supports client buffer feedback, upon reception of an SDP containing the "3GPP-Adaptation-Support" attribute, it shall include the "3GPP-Adaptation" header in the SETUP for each individual media. Furthermore, upon reception of a successful SETUP response (including "3GPP-Adaptation" header), the PSS client shall send NADU APP packets according to clause 5.3.3.5 and 6.2.3.2.

The "3GPP-Adaptation" header may be included in PLAY, OPTIONS and SET_PARAMETER requests in order to update the target buffer protection time value during a session. However, the target-protection-time is intended to be stable for the entire session with the server there are very few reasons for a client to modify the target buffer protection time once a session is established. The buffer size value shall not be modified during a session.

With the total buffer size, and the reported amount of free buffer space, the server can avoid overflowing the buffer. A server should assume that any sent RTP packet will consume receiver buffer space equal to the complete RTP packet size. For interleaved or aggregated media, the actual buffer space consumption may be slightly larger if buffering is done in the ADU domain. This is because each ADU may save metadata corresponding to the RTP header and payload fields, like timestamp and decoding sequence numbers individually. This should only be a problem if a server tries to fill exactly to the last free memory block.

The server can determine the time to underflow by calculating the amount of media time present in the buffer. This is done using the next ADU numbers, the highest received sequence number, and the playout delay, combined with the server's view of the sent ADUs and their decoding order and playout time. The information about the ADUs for 3GP files that are produced according to the streaming-server profile can be read from the "3gau" box [50]. It is also possible to derive some of the information about the ADUs from the media track, or hint-track, or the actual RTP packets.

A client needs to choose the target-time and the point on the playout timeline from which it will measure PlayoutDelay such that it will never re-buffer when the target-time is fulfilled. A client should typically begin rebuffering only when it has reached 0 ms buffered data. Once rebuffering has begun, the client should resume playback when the target-time has been fulfilled for all synchronized media streams.

The level of protection needed against transmission rate variations over a wireless network can be substantial (throughput variation because of network load, radio conditions, several seconds of interruption because of handovers, possible extra buffering to perform retransmission). In order to minimise the initial buffering delay, the client may choose an initial buffering that is less than the required buffering it has determined would be satisfactory. The client needs to take into account, however, that it may be unsafe to begin playback prior to fulfilling its target time. For this reason, the target buffer protection time indicates the amount of playable media (in time), which the client perceives necessary to have in its buffer. Therefore a server should not perform content adaptation towards higher content rates until the given target time of media units is available in the buffer.

It is important to note that target-protection-time is intended only to guide the server in its attempts to sustain or improve the quality of the media. There are many situations in which the target-protection time may not be respected by the server which will actually result in better media quality for the client (e.g. when the client sends a target-protection-time smaller than the perceived jitter or when the client sends a target-protection-time that is close to or exceeds the client buffer maximum). The only requirement the target-time places on the server is that the server shall not attempt to upshift prior to attaining the target-time.

Furthermore, while it is possible for the client to modify the target protection time in the 3GPP adaptation header with each RTSP request that is sent to the server, the target protection time is intended to be a stable value for the entire session with the server and should only be modified in circumstances where the client has a more accurate understanding of network and transmission jitter and the efficiency of its ability to process the network buffer. In these circumstances, adjusting the target time up could prevent buffer low points which will cause rebuffering or, adjusting the target time down could provide more head room to allow the server to adapt to the most appropriate rate.

10.3 Issues with deriving adaptation information (informative)

This clause attempts to provide some insight into the functions and issues that exist in deriving client''s buffer status in the server. The issues and the complexity of the functions depend on the media format, but can be characterised by media properties, in particular how much flexibility the media formats allows in transmission, decoding, and playout order. As there are three orderings of encoded media data that are possible, there are two re-orderings:

a) Data may be interleaved (i.e. the transmission order of data differs from the decoding order), and it must be de-interleaved before passing to the decoder.

b) There are forward references in the encoding, e.g. in a video stream, then those references are decoded 'early' (out of order) compared to playout order. Thus, the playout order in this case differs from the decode order.

In buffer management, we are trying to ensure

1. that the client's receiver buffer does not get over-filled (this is over-run);
2. that data does not arrive at an operation point after its need. Specifically, this means that ADUs should not be placed into the final playout queue with a timestamp that has already been passed in playout (this is under-run).

The parameters supplied enable a server to deduce at least this much. The server can always protect against buffer over-run by respecting the 'free space' that is periodically signalled by the client. This free-space is totalled over all data held before the decoder (decoder and de-interleave buffers). If the server desires more visibility, it can inspect the ADU that has been reported as 'next to decode'. If there has been no interleaving, the client holds all data between that ADU and the highest sequence number received, and will probably hold up to the last packet the server has sent. If interleaving is used, then there may have been ADUs that were sent after the reported ADU, but which passed out of both the de-interleaving and decoder buffers before that ADU. The server would have to analyze the de-interleaving process to work out which ADUs these are. The hint-track extension "3gau" to the 3GP file format [50] provides extended information about both the decoding and playout order in relation to transmission order of the ADUs. This extension does also provide the size of the ADUs to the server.

Protection against under-run is more subtle. It is in general not possible for the client to know which ADUs that are yet to be decoded (or yet to be received) that have earlier timestamps than ADUs already received and decoded. Therefore the client does not in fact know what is the 'latest playable timestamp', up to which it has received all the ADUs in the sequence to that time.

If the server does not adapt its transmission bit-rate and the transmission path has sufficient bit-rate, the parameters supplied at stream setup (such as the initial buffering delay) are sufficient to protect against under-run. The simple generalization of this is that if the server calculates its average bit-rate since starting the stream, and ensures that the average never falls below the bit-rate that would have been used without rate adaptation, it must be safe. Put in another way, the server may send a packet earlier than it would without rate-adaptation, but it might not be safe to send it later.

A more subtle analysis uses the reported information about the next-to-be-decoded ADU: the sequence number of the packet that contained it, the ADU number within that packet, and the offset (playout delay) of its timestamp (playback time) from the current playback time. Given the first pair of numbers, the server can find the ADU and therefore its timestamp. By subtracting the reported play-out delay from this timestamp, the server can now estimate the current playback time. It can find the earliest timestamp in the ADUs it has yet to transmit, and it can also examine the data that has been sent that will still be in the de-interleaving buffer, for the earliest timestamp still held in the client's de-interleaving buffer. If the earlier of these two timestamps is at, or close to, the current play time, the client has, or is about to, under-run.

Consider now the following cases, in increasing order of complexity:

1. simple data that is neither interleaved nor re-ordered for display (e.g. AMR without interleave, AAC, H.263, MPEG-4 video).
2. data that is interleaved, but not re-ordered (e.g. AMR with interleave).
3. data that is re-ordered, but not interleaved (AVC without interleave).
4. data that is both interleaved and re-ordered (AVC with interleave).

Consider now over-run and under-run protection for these streams. In all cases, the free-space can be used to protect against over-run, and the maintenance of the average rate at or above the static rate protects against under-run.

1. By subtracting the reported free-space from the overall buffer size (reported in stream setup) the buffered data can be calculated. If this is nearly exhausted, the buffer is about to under-run. However for codecs with variable bit-rate encoding, the buffered space may represent different amounts of playout time. In these cases the playout time present in the yet to be decoded part of the buffer can easily be calculated as the RTP timestamp difference between the latest ADU received by the client as reported implicitly by Highest Received Sequence number and the ADU reported by NADU.

2, The server can estimate the playback time as above. However to perform the calculation of the playout time of the buffer before the decoding, the server may need to maintain a list of the ADUs in the decoding order, rather than in transmission order. Also the data present in the de-interleaving buffer is not complete and would have holes in it and should not be considered to be playable. The server can determine, by looking at the decoding order of the different ADUs present in the transmitted packets, how far the client is expected to have a receiver buffer without holes, due to not yet transmitted packets.

3, In this case it may be fairly complicated to estimate the actual playout time of the un-decoded media. The reason is that the present RTP timestamp associated with the ADUs may fluctuate widely in ADUs consecutive in both transmission and decoding order, due to the early decoding of referenced ADUs. Therefore to perform an accurate estimation the server needs to make special consideration of any ADU with early decoding so that it does not skew the measurement. Note that for AVC bitstreams, a bound of the difference between presentation order and decoding order is given by the bitstream restriction parameter num_reorder_frames.

4, As 3 above, but with the further consideration of needing to perform any investigation in decoding order and consider the holes of the de-interleaving buffer.

11 Quality of Experience

11.1 General

The PSS Quality of Experience (QoE) metrics feature is optional for both PSS servers and clients, and shall not disturb the PSS service. A PSS server that supports the QoE metrics feature shall signal the activation and gathering of client QoE metrics when desired. QoE metrics can also be activated by a default setting via OMA-DM. A 3GPP PSS client supporting the feature shall perform the quality measurements in accordance to the measurement definitions, aggregate them into client QoE metrics and report the metrics to a server, which may or may not be the PSS server, using the QoE transport protocol when so requested. The way the QoE metrics are processed and made available is out of the scope of this specification.

11.2 QoE metrics

A PSS client should measure the metrics at the transport layer, but may also do it at the application layer for better accuracy.

The measurement period for the metrics is the period over which a set of metrics is calculated. The maximum value of the measurement period is negotiated via the QoE protocol as in clause 11.3. The measurement period shall not include any voluntary event that impacts the actual play, such as pause or rewind, or any buffering or freezes/gaps caused by them.

The following metrics shall be derived by the PSS client implementing QoE. All the metrics defined below are only applicable to at least one of audio, video, speech and timed text media types, and are not applicable to other media types such as synthetic audio, still images, bitmap graphics, vector graphics, and text. Any unknown metrics shall be ignored by the client and not included in any QoE report. Among the QoE metrics, corruption duration, successive loss of RTP packets, frame-rate deviation and jitter duration are of media level, whereas content switch time, initial buffering duration and rebuffering duration are of session level.

In the case of guaranteed delivery transports, such as HTTP as used in progressive download or HTTP-based streaming, metrics relating to loss or corruption (such as "Corruption duration", "Successive loss of RTP packets" and "Jitter duration") are not relevant and should be omitted from the report or report that no corruption has occurred.

11.2.1 Corruption duration metric

11.2.1.1 Default reporting format

Corruption duration, M, is the time period from the NPT time of the last good frame before the corruption (since the NPT time for the first corrupted frame cannot always be determined) or the start of the measurement period (whichever is later), to the NPT time of the first subsequent good frame or the end of the measurement period (whichever is sooner). A corrupted frame is either an entirely lost frame, or a media frame that has quality degradation and the decoded frame is not the same as in error-free decoding. A good frame is a "completely received" frame X that

- either it is a refresh frame (does not reference any previously decoded frames AND where none of the subsequently decoded frames reference any frames decoded prior to X);
- or does not reference any previously decoded frames;
- or only references previously decoded "good frames".

"Completely received" means that all the bits are received and no bit error has occurred.

Corruption duration, M, in milliseconds can be calculated according to the derivation of good frames as below:

a) A good frame can be derived by the client using the codec layer, in which case the codec layer signals the decoding of a good frame to the client. A good frame could also be derived by error tracking methods, but decoding quality evaluation methods shall not be used. An error tracking method may derive that a frame is a good frame even when it references previously decoded corrupted frames, as long as all the referenced pixels for generating the prediction signal were correctly reconstructed when decoding the reference frames. A decoding quality evaluation method may derive that a frame is a good frame even one or more pixels of the frame have not been correctly reconstructed, as long as the decoding quality is considered by the method as acceptable. Such a frame is not a good frame according to the definition above, which shall be strictly followed.

b) In the absence of information from the codec layer, a good frame is derived according to N, where N is optionally signalled from server to client and represents the maximum duration, in presentation time, between two subsequent refresh frames in milliseconds. After a corrupted frame, if all subsequent frames within N milliseconds in presentation time have been completely received, then the next frame is a good frame.

c) If N is not signalled, then it defaults to infinity (for video) or to one frame duration (for audio).

The optional parameter D is defined to indicate which of options a) and b) is in use. D is signalled from the client to the server. When D is equal to "a", option a) shall be in use, and the optional parameter T shall be present. When D is equal to "b", option b) shall be in use and the optional parameter T shall not be present.

The optional parameter N as defined in point b is used with the "Corruption_Duration" parameter in the "3GPP-QoE-Metrics" header. The optional parameter T is defined to indicate whether the client uses error tracking (when T is equal to "On") or not (when T is equal to "Off"). T is signalled from the client to the server.

The syntax for D, N and T to be included in the "Measure-Spec" (clause 5.3.2.3.1) is as follows:

```
D  = "D" "=" "a" / "b"
N  = "N" "=" 1*DIGIT
T  = "T" "=" "On" / "Off"
```

The syntax for the "Metrics-Name Corruption_Duration" for the QoE-Feedback header is as defined in clause 5.3.2.3.2

The absence of an event is reported using the space (SP).

For the "Metrics-Name Corruption_Duration", the "Value" field in 5.3.2.3.2 indicates the corruption duration. The unit of this metrics is expressed in milliseconds. There is the possibility that corruption occurs more than once during a measurement period. In that case the value can occur more than once indicating the number of corruption events.

The value of "Timestamp" is equal to the NPT time of the last good frame inside the measurement period, in playback order, before the occurrence of the corruption, relative to the starting time of the measurement period. If there is no good frame inside the measurement period and before the corruption, the timestamp is set to the starting time of the measurement period.

11.2.1.2 XML reporting format

All the occurred corruption durations within each resolution period are summed and stored in the vector *TotalCorruptionDuration*. Within each resolution period the number of individual corruption events are summed up and stored in the vector *NumberOfCorruptionEvents*. These two vectors are then reported by the PSS client as Metric-Name "TotalCorruptionDuration" and "NumberOfCorruptionEvents" respectively. The use of error tracking is reported by setting the parameter *t* to "True" or "False".

11.2.2 Rebuffering duration metric

11.2.2.1 Default reporting format

Rebuffering is defined as any stall in playback time due to any involuntary event at the client side.

The syntax for the 'Metrics-Name Rebuffering_Duration" for the QoE-Feedback header is as defined in clause 5.3.2.3.2.

The absence of an event is reported using the space (SP).

For the "Metrics-Name Rebuffering_Duration", the "Value" field in 5.3.2.3.2 indicates the rebuffering duration. The unit of this metrics is expressed in seconds, and can be a fractional value. There is the possibility that rebuffering occurs more than once during a measurement period. In that case the metrics value can occur more than once indicating the number of rebuffering events.

The optional "Timestamp" indicates the time when the rebuffering has occurred since the beginning of the measurement period. The value of the "Timestamp" is equal to the NPT time of the last played frame inside the measurement period and before the occurrence of the rebuffering. If there is no played frame inside the measurement period, the timestamp is set to the starting time of the measurement period.

11.2.2.2 XML reporting format

All the occurred rebuffering durations are summed up over each resolution period of the stream and stored in the vector *TotalRebufferingDuration*. The number of individual rebuffering events for each resolution period are summed up and stored in the vector *NumberOfRebufferingEvents*. These two vectors are then reported by the PSS client as Metric-Name "TotalRebufferingDuration" and "NumberOfRebufferingEvents" respectively.

11.2.3 Initial buffering duration metric

11.2.3.1 Default reporting format

Initial buffering duration is the time from receiving the first media packet until playing starts.

The syntax for the "Metrics-Name Initial_Buffering_Duration" for the QoE-Feedback header is as defined in clause 5.3.2.3.2 with the exception that "Timestamp" in "Measure" is undefined for this metric. If the measurement period is shorter than the "Initial_Buffering_Duration" then the client should send this parameter for each measurement period as long as it observes it. The "Value" field indicates the initial buffering duration occurring during the current measurement period, where the unit of this metrics is expressed in seconds, and can be a fractional value. There can be only one "Measure" and it can only take one "Value". The absence of an event can be reported using the space (SP). "Initial_Buffering_Duration" is a session level parameter.

For instance, if the measurement period is set to one second, and the total initial buffering duration is 2.4 seconds, then the three first initial buffering duration values reported will be 1 second, 1 second and 0.4 seconds.

11.2.3.2 XML reporting format

The XML reporting format is identical to the default reporting format.

11.2.4 Successive loss of RTP packets

11.2.4.1 Default reporting format

This parameter indicates the number of RTP packets lost in succession per media channel.

The syntax for the "Metrics-Name Successive_Loss" for the QoE-Feedback header is as defined in clause 5.3.2.3.2.

The absence of an event can be reported using the space (SP).

For the "Metrics-Name Successive_Loss", the "Value" field indicates the number of RTP packets lost in succession. The unit of this metric is expressed as an integer equal to or larger than 1. There is the possibility that successive loss occurs more than once during a measurement period. In that case the metrics value can occur more than once indicating the number of successive losses.

The optional "Timestamp" indicates the time when the succession of lost packets has occurred. The value of the "Timestamp" is equal to the NPT time of the last received RTP packet inside the measurement period, in playback order, before the occurrence of the succession of lost packets, relative to the starting time of the measurement period. If there is no received RTP packet inside the measurement period and before the succession of loss, the timestamp is set to the starting time of the measurement period.

If a full run length encoding of RTP losses with sequence number information is desired, RTCP XR [RFC 3611] Loss RLE Reporting Blocks should be used instead of the successive loss metric.

11.2.4.2 XML reporting format

All the number of successively lost RTP packets are summed up over each resolution period of the stream and stored in the vector *TotalNumberofSuccessivePacketLoss*. The unit of this metric is expressed as an integer equal to or larger than 0. The number of individual successive packet loss events over each resolution period are summed up and stored in the vector *NumberOfSuccessiveLossEvents.* The number of received packets is also summed up over each resolution period and stored in the vector *NumberOfReceivedPackets*. These three vectors are reported by the PSS client as as Metric-Name "TotalNumberofSuccessivePacketLoss", "NumberOfSuccessiveLossEvents" and "NumberOfReceivedPackets" respectively.

11.2.5 Frame rate deviation

11.2.5.1 Default reporting format

Frame rate deviation indicates the playback frame rate information. Frame rate deviation happens when the actual playback frame rate during a measurement period is deviated from a pre-defined value.

The actual playback frame rate is equal to the number of frames played during the measurement period divided by the time duration, in seconds, of the measurement period.

The parameter FR that denotes the pre-defined frame rate value is used with the "Framerate_Deviation" parameter in the "3GPP-QoE-Metrics" header. The value of FR shall be set by the server. The syntax for FR to be included in the "Measure-Spec" (clause 5.3.2.3.1) is as follows:

FR = "FR" "=" 1*DIGIT "." 1*DIGIT

The syntax for the Metrics-Name "Framerate_Deviation" for the QoE-Feedback header is as defined in clause 5.3.2.3.2 with the exception that "Timestamp" in "Measure" is undefined for this metric. The absence of an event can be reported using the space (SP).

For the Metrics-Name "Framerate_Deviation", "Value" field indicates the frame rate deviation value that is equal to the pre-defined frame rate minus the actual playback frame rate. This metric is expressed in frames per second, and can be a fractional value, and can be negative. The metric value can occur only once for this metric.

11.2.5.2 XML reporting format

In the XML reporting format the frame rate is reported instead of the frame rate deviation. The metric "Framerate" indicates the average actual playback frame rate used during each resolution period. It is expressed in frames per second, and can be a fractional value. The average frame rate for each resolution period is stored in the vector *Framerate* and the vector is reported by the PSS client as Metric-Name "FrameRate".

11.2.6 Jitter duration

11.2.6.1 Default reporting format

Jitter happens when the absolute difference between the actual playback time and the expected playback time is larger than a pre-defined value, which is 100 milliseconds. The expected time of a frame is equal to the actual playback time of the last played frame plus the difference between the NPT time of the frame and the NPT time of the last played frame.

The syntax for the Metrics-Name "Jitter_Duration" for the QoE-Feedback header is as defined in clause 5.3.2.3.2.

The absence of an event can be reported using the space (SP).

For the Metrics-Name "Jitter_Duration", the "Value" field in 5.3.2.3.2 indicates the time duration of the playback jitter. The unit of this metrics is expressed in seconds, and can be a fractional value. There is the possibility that jitter occurs more than once during a measurement period. In that case the metric value can occur more than once indicating the number of jitter events.

The optional "Timestamp" field indicates the time when the jitter has occurred since the beginning of the measurement period. The value of the "Timestamp" is equal to the NPT time of the first played frame in the playback jitter, relative to the starting time of the measurement period.

11.2.6.2 XML reporting format

All the Jitter_Durations are summed up over each resolution period and stored in the vector *TotalJitterDuration*. The number of individual events over the resolution duration are summed up and stored in the vector *NumberOfJitterEvents*. These two vectors are then reported by the PSS client as Metric-Name "TotalJitterDuration" and "NumberOfJitterEvents" respectively.

11.2.7 Content Switch Time

11.2.7.1 Default reporting format

Fast content switching is defined in section 5.5 and allows for improving the switch time between different content accessible via the same RTSP server. Content switch time has a significant impact on the quality of experience for the user. The content switch time metric is used to report the time that elapses between the initiation of the content switch by the user and up to the time of reception of the first media packet from the new content or media stream.

The syntax for the metric 'Content_Switch_Time' for the QoE Feedback header is defined in clause 5.3.2.3.2.

The absence of a content switch event or the impossibility to determine the duration of a content switch can be reported using the space (SP).

For the metric name 'Content_Switch_Time', the 'Value' field in 5.3.2.3.2 indicates the duration of the content switch as defined above. The unit of this metric is expressed in milliseconds.

In case several content switch events have occurred during the measurement period, a list of values is reported each relating to the corresponding old content or media URL.

The optional 'Timestamp' field indicates the time when the content switch event was triggered by the user. The value of the 'Timestamp' is equal to the NPT time of the old content that corresponds to the content switch triggering time.

11.2.7.2 XML reporting format

All the Content_Switch_Times are summed up over each resolution period and stored in the vector *TotalContentSwitchTime*. The number of individual events over the resolution duration are summed up and stored in the vector *NumberOfContentSwitchEvents*. These two vectors are then reported by the PSS client as Metric-Name "TotalContentSwitchTime" and "NumberOfContentSwitchEvents" respectively.

11.2.8 Average Codec Bitrate

11.2.8.1 Default reporting format

The average codec bitrate is the bitrate used for coding 'active' media information during the measurement resolution period.

For audio media 'active' information is defined by frames containing audio. If the audio codec uses silence frames (SID-frames), these frames are not counted as "active", and the SID-frames and the corresponding DTX time periods are excluded from the calculation. Thus for audio media the average codec bitrate can be calculated as the number of audio bits received for 'active' frames , divided by the total time, in seconds, covered by these frames. The total time covered is calculated as the number of 'active' frames times the length of each audio frame.

For non-audio media the average codec bitrate is the total number of media bits played out during the measurement resolution period, divided by the length of the playout period. The playout period length is normally equal to the length of the measurement resolution period, but if rebuffering occurs the playout period will be shorter (i.e. any rebuffering time shall be ignored when calculating the codec bitrate).

The syntax for the metric "Average_Codec_Bitrate" is defined in sub-clause 5.3.2.3.2.

For the metric name 'Average_Codec_Bitrate', the 'Value' field in 5.3.2.3.2 indicates the codec bitrate as defined above. The unit of this metrics is expressed in kbit/s and can be a fractional value.

11.2.8.2 XML reporting format

The average codec bitrate value for each measurement resolution period shall be stored in the vector *AverageCodecBitrate*. The unit of this metrics is expressed in kbit/s and can be a fractional value. The vector is then reported by the PSS client as Metric-Name "AverageCodecBitrate".

11.2.9 Codec Information

11.2.9.1 Default reporting format

The codec information metrics contain details of the media codec used during the measurement period. The unit of this metric is a string value. No "white space" characters are allowed in the string values, and shall be removed if necessary.

For audio media the codec information contains the audio codec type, represented as in an SDP offer, for instance "AMR-WB/16000/1".

For video media, the codec information contains the video codec type, represented as in an SDP offer, for instance 'H263-2000/90000'. Furthermore, the video profile and level used, as well as the image size used shall be reported. For instance "profile=0;level=45" for the profile and level information and '176x144' for the image size. In some cases the profile and level is reported together, for instance "profile-level-id=42e00a". Note that the image size reported for each measurement resolution period shall be the one actually used, not the maximum size allowed by the SDP negotiation.

For timed text media, the codec information contains the text encoding, represented as in an SDP offer, for instance "3gpp-tt/1000".

The syntax for the metric "CodecInfo", 'CodecProfileLevel' and 'CodecImageSize' are defined in sub-clause 5.3.2.3.2.

There is the possibility that the codec information is changed during the measurement period. In that case the metrics can occur more than once indicating the codecs used.

The optional "Timestamp" field indicates the time when codec changes have occurred, relative to the beginning of the measurement period.

11.2.9.2 XML reporting format

The codec info, profile/level and codec image size value for each measurement resolution period shall be stored in the vectors *CodecInfo*, *CodecProfileLevel* and *CodecImageSize* respectively. If the metric values in these vectors for a measurement resolution period are unchanged from the previous values in the respective vector, it is allowed to put the value '=' in the vector to indicate this. These three vectors are reported by the PSS client as as Metric-Name "CodecInfo", "CodecProfileLevel" and "CodecImageSize" respectively.

11.2.10 Buffer Status

11.2.10.1 Default reporting format

The buffer depth is the number of seconds of future media which resides in the buffer. It is calculated as the difference between the latest playout time of the media units in the buffer minus the current playout time. If the calculation result is negative, the buffer depth shall be set to zero. The buffer depth metric shall be calculated close to the end of each measurement period.

The unit of the metric bufferDepth is in seconds and can be a fractional value. If the length of the media is known, and if all remaining media already is buffered, then the boolean metric allContentBuffered shall be set to true.

The syntax for the metric "bufferDepth" and "allContentBuffered" is defined in sub-clause 5.3.2.3.2.

11.2.10.2 XML reporting format

The buffer depth close to the end of each measurement period shall be stored in the vector *bufferDepth*. The vector and the *allContentBuffered* status is then reported at the end of each reporting period.

11.3 The QoE protocol

11.3.1 General

PSS clients and servers supporting QoE Metrics shall support the QoE protocol described below.

The RTSP and SDP based protocol extensions (see clauses 5.3.2.3 and 5.3.3.6) are used for transport and negotiation of the QoE metrics between the PSS client and the PSS server. As an alternative, OMA-DM and HTTP can also be used for QoE configuration and reporting (see clauses 5.3.3.8 and 5.3.2.3.3).

The QoE metrics negotiation starts with the response to the DESCRIBE request, if the metrics information is embedded in the SDP data (as described in example 1 in clause 11.3.2). For the case of locally stored SDP which contains QoE-Metrics attribute, the negotiation starts with client"s SETUP request. If the PSS client supports QoE metrics, then it shall send a SETUP request containing the selected (i.e. accepted by client)/modified (for re-negotiation) QoE metrics for either session level, or the media level, which is being set-up. Such a SETUP request is shown in example 2 in clause 11.3.3.

Upon receiving this SETUP request, the server shall return the RTSP Response with the "accepted" QoE metrics (i.e. metrics and metrics values which are identical to the ones in the client"s request and accepted by the server) and the "re-negotiation" QoE metrics (i.e. metrics and metrics values which are not identical to the ones in the client"s request and modified for re-negotiation by the server) .The echoing of the "accepted" QoE metrics is for re-acknowledging the client. The server may also reject the changes made by the client, i.e. reject the "re-negotiation" QoE metrics. If the server rejects the changes, it shall either set new values or resend the modified metrics back to the client, or it shall ignore the "re-negotiation" metrics and not re-acknowledge them. Any QoE metric that has been acknowledged as "accepted" by the server shall not be re-negotiated, i.e., it shall not be resent in the "3GPP-QoE-Metrics" header in the next RTSP request and shall not be re-acknowledged in the next RTSP response.

If the server does not approve the modifications done by the client, they should continue to re-negotiate. However, negotiations shall not exceed 4 round trips, in order to bound the negotiation process. It must be noted that each time the "QoE-Metrics" header field is sent in an RTSP request, it shall also be present in the response corresponding to that particular request. Otherwise, the receiver of the response shall assume that the other end does NOT support QoE metrics.

If there is no DESCRIBE – RTSP Response pair sending at the beginning of the RTSP signalling (see Figure 11.2), it means that the SDP description is received by other means. If such an SDP contains the "3GPP-QoE-Metrics" attribute, the negotiation happens in the same way as it is described above, i.e. starts with SETUP request containing "3GPP-QoE-Metrics" header. If the SDP does not contain the "3GPP-QoE-Metrics" attribute and the server would still like to check whether the client supports QoE Protocol or not, the server shall include the "3GPP-QoE-Metrics" header containing the initial QoE metrics in the SETUP response. If the PSS client sends the QoE metrics information in the next request (indicating that it supports QoE Protocol), the negotiation shall continue until the mutual agreement is reached or the negotiation limit is reached. If pipelined startup is not in use and the client does not send QoE metrics information in the next request to SETUP response, then the server shall assume that the client does not support QoE metrics. In case pipelined startup is in use, the server may initiate QoE negotiation but it should not expect an answer from the PSS client.

In case of switching without the SDP, the PSS client shall assume that the same QoE metrics as negotiated for the old stream will be used for the new stream. In the PLAY response, the server includes the '3GPP-QoE-Metrics' header to acknowledge the QoE metric mapping to the new media streams or to change them.

During fast content switching with SDP, the client shall indicate the QoE metrics to be used for the new content using the '3GPP-QoE-Metrics' header. The client should use the already negotiated parameters as much as possible to avoid further negotiations. The server shall either acknowledge the proposed QoE metrics or continue negotiation beyond the PLAY response message. It is possible to turn off the metrics during a streaming session. In clause 11.3 an example of messages, where the metrics are set to "Off" is given. The metrics can be set to "Off" at session level or at media level. The request url indicates what level is used. If no url is used, then 'Off' applies to session level. The server should use OPTIONS (with Session ID) or SET_PARAMETER RTSP methods to turn off the QoE feedback.

A client should not send QoE feedback during RTSP ready state. After the ready state is ended (i.e., RTSP state=playing), the periodic feedback and normal operations continue. This reduces the network load in the uplink and downlink directions, and the processing overhead for the PSS client. When an RTSP PLAY request is sent by the PSS client after a PAUSE, the clock for the measurement period (based on the defined "Sending Rate") shall be reset.

If there are multiple non-aggregated sessions, i.e. each media delivery is initiated by a different PLAY request, the QoE metrics are negotiated and reported for each session separately.

All the QoE Metrics in the following examples are fictitious. Clause 11.2 defines the actual QoE Metrics.

11.3.2 Metrics initiation with SDP

QoE metrics initiation with SDP shall be done according to clause 5.3.3.6.

This following example shows the syntax of the SDP attribute for QoE metrics. The session level QoE metrics description (Initial buffering duration and rebufferings) are to be monitored and reported only once at the end of the session. Also video specific description of metrics (corruptions and decoded bytes) is to be monitored and reported every 15 seconds from the beginning of the stream until the time 40s. Finally, audio specific description of metrics (corruptions) is to be monitored and reported every 20 seconds from the beginning until the end of the stream.

EXAMPLE 1:

```
S->C  RTSP/1.0 200 OK
Cseq: 1
Content-Type: application/sdp
Content-Base: rtsp://example.com/foo/bar/baz.3gp/
Content-Length: 800
Server: PSSR7 Server

v=0
o=- 3268077682 433392265 IN IP4 63.108.142.6
s=QoE Enables Session Description Example
e=support@foo.com
c=IN IP4 0.0.0.0
t=0 0
a=range:npt=0-83.660000
a=3GPP-QoE-Metrics:metrics={Initial_Buffering_Duration|Rebuffering_Duration};rate=End
a=control:*
m=video 0 RTP/AVP 96
b=AS:28
a=3GPP-QoE-Metrics:metrics={Corruption_Duration|Decoded_Bytes};rate=15;range:npt=0-40
a=control:trackID=3
a=rtpmap:96 MP4V-ES/1000
a=range:npt=0-83.666000
a=fmtp:96profile-level-id=8;config=000001b008000001b50900012000
m=audio 0 RTP/AVP 98
b=AS:13
a=3GPP-QoE-Metrics:metrics={Corruption_Duration};rate=20
a=control:trackID=5
a=rtpmap:98 AMR/8000
a=range:npt=0-83.660000
a=fmtp:98 octet-align=1
a=maxptime:200
```

11.3.3 Metrics initiation/termination with RTSP

QoE Metrics initiation with RTSP can be done according to clause 5.3.2.3.1

In the following example it is shown how to negotiate QoE metrics during RTSP session setup.

EXAMPLE 1 (QoE metrics negotiation):

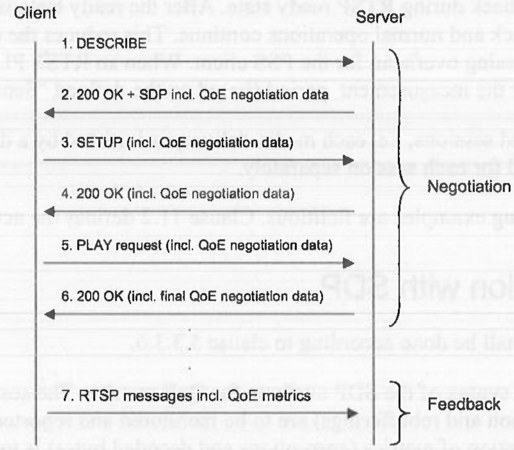


Figure 11.1: QoE metrics negotiation

```
C->S    SETUP rtsp://example.com/foo/bar/baz.3gp/trackID=3 RTSP/1.0
        Cseq: 2
        3GPP-QoE-Metrics:url='rtsp://example.com/foo/bar/baz.3gp/trackID=3';
        metrics={Corruption_Duration|Decoded_Bytes};rate=10; Range:npt=0-40,
        url='rtsp://example.com/foo/bar/baz.3gp';
        metrics={Initial_Buffering_Duration|Rebuffering_Duration};rate=End
```

In the above SETUP request, the client modifies the sending rate of the QoE metrics for the control URL 'rtsp://example.com/foo/bar/baz.3gp/trackID=3' from 15 to 10 (compared to the initial SDP description).

Assuming that the server acknowledged the changes, the server will send back a SETUP response as follows:

```
S->C    RTSP/1.0 200 OK
        Cseq: 2
        Session: 17903320
        Transport: RTP/AVP;unicast;client_port=7000-7001;server_port= 6970-6971
        3GPP-QoE-Metrics:url='rtsp://example.com/foo/bar/baz.3gp/trackID=3';
        metrics={Corruption_Duration|Decoded_Bytes};rate=10;Range:npt=0-40,
        url='rtsp://example.com/foo/bar/baz.3gp';
        metrics={Initial_Buffering_Duration|Rebuffering_Duration};rate=End
```

EXAMPLE 2 (QoE metrics negotiation – no DESCRIBE – 200/OK):

An example is shown in Figure 11.2 and can make use of the same RTSP header defined in clause 5.3.2.3.

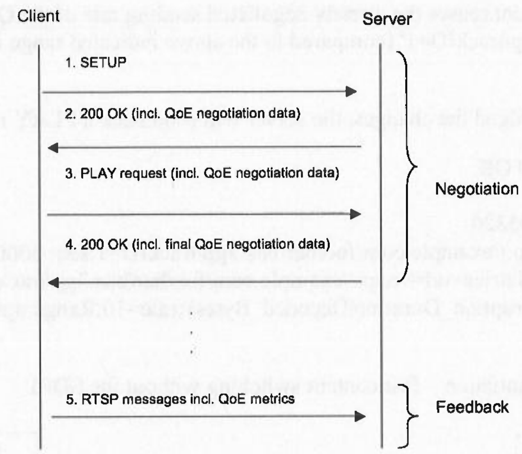


Figure 11.2: QoE metrics negotiation (no DESCRIBE-200/OK)

EXAMPLE 3 (QoE metrics negotiation – fast content switching with the SDP)

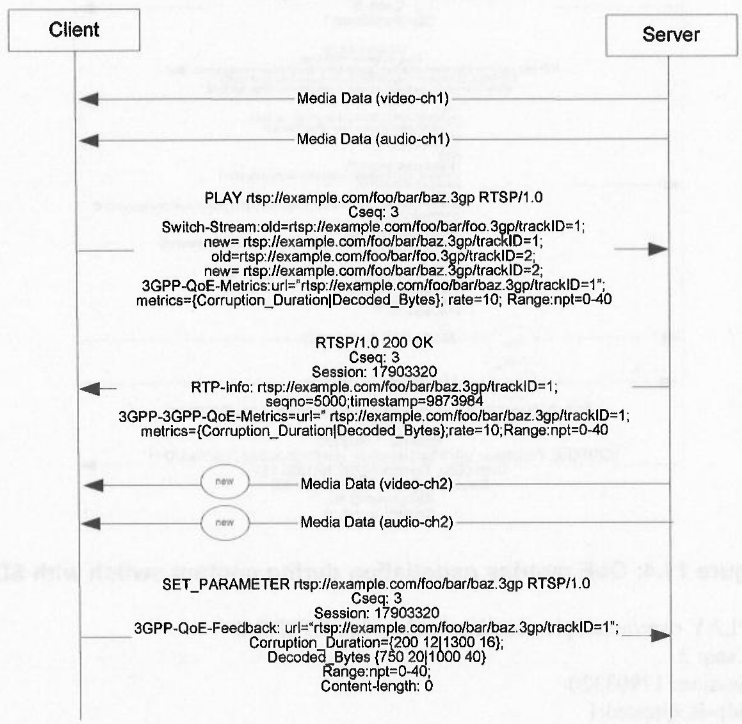


Figure 11.3: QoE metrics negotiation during content switch with SDP

```
C->S    PLAY rtsp://example.com/foo/bar/baz.3gp RTSP/1.0
        Cseq: 3
        Session: 17903320
        Switch-Stream:old='rtsp://example.com/foo/bar/foo.3gp/trackID=1';new='
        rtsp://example.com/foo/bar/baz.3gp/trackID=1';
        old='rtsp://example.com/foo/bar/foo.3gp/trackID=2';new='
        rtsp://example.com/foo/bar/baz.3gp/trackID=2';
        3GPP-QoE-Metrics:url='rtsp://example.com/foo/bar/baz.3gp/trackID=1';
        metrics={Corruption_Duration|Decoded_Bytes};rate=10; Range:npt=0-40
```

In the above PLAY request, the client reuses the already negotiated sending rate of the QoE metrics for the control URL 'rtsp://example.com/foo/bar/baz.3gp/trackID=1' (compared to the above indicated range in the SDP description, which is 15).

Assuming that the server acknowledged the changes, the server will send back a PLAY response as follows:

S->C RTSP/1.0 200 OK
Cseq: 3
Session: 17903320
RTP-Info: rtsp://example.com/foo/bar/baz.3gp/trackID=1;seq=5000;rtptime=9873984 3GPP-
3GPP-QoE-Metrics=url=' rtsp://example.com/foo/bar/baz.3gp/trackID=1;
metrics={Corruption_Duration|Decoded_Bytes};rate=10;Range:npt=0-40

EXAMPLE 4 (QoE metrics negotiation – fast content switching without the SDP)

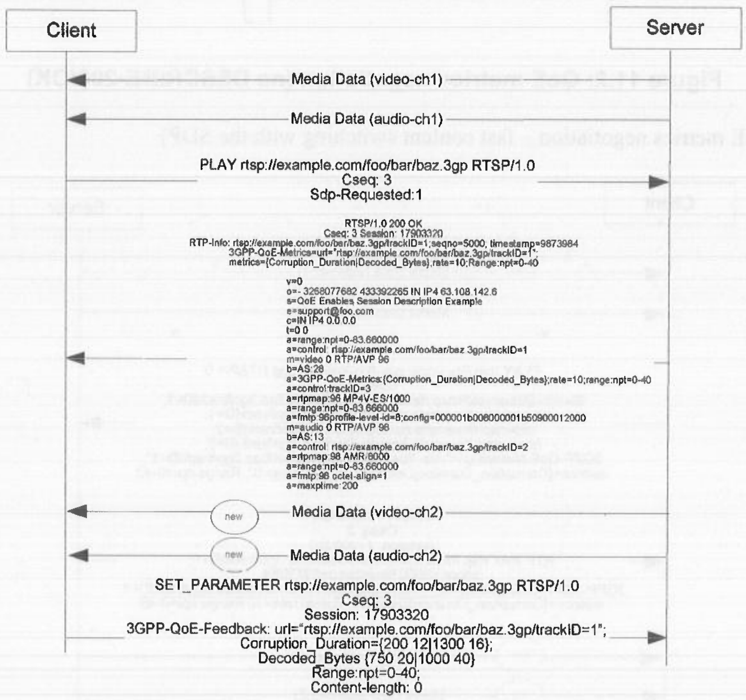


Figure 11.4: QoE metrics negotiation during content switch with SDP

C->S PLAY rtsp://example.com/foo/bar/baz.3gp RTSP/1.0
Cseq: 3
Session: 17903320
Sdp-Requested:1

In the above PLAY request, the client switches to new content without having the SDP. By consequence, the client does not have any information about the requested QoE parameters and cannot indicate those in the PLAY request.

In the case of a successful switch, the server indicates the QoE metrics in the SDP and in the 3GPP-QoE-Metrics header. The server reuses the already negotiated feedback period of 10 seconds.

S->C RTSP/1.0 200 OK
Cseq: 3
Session: 17903320
RTP-Info: rtsp://example.com/foo/bar/baz.3gp/trackID=1;seq=5000;rtptime=9873984 3GPP-
3GPP-QoE-Metrics=url='rtsp://example.com/foo/bar/baz.3gp/trackID=1';
metrics={Corruption_Duration|Decoded_Bytes};rate=10;Range:npt=0-40

```
v=0
o=- 3268077682 433392265 IN IP4 63.108.142.6
s=QoE Enables Session Description Example
e=support@foo.com
c=IN IP4 0.0.0.0
t=0 0
a=range:npt=0-83.660000
a=control: rtsp://example.com/foo/bar/baz.3gp/trackID=1
m=video 0 RTP/AVP 96
b=AS:28
a=3GPP-QoE-Metrics:{Corruption_Duration|Decoded_Bytes};rate=10;range:npt=0-40
a=control:trackID=3
a=rtpmap:96 MP4V-ES/1000
a=range:npt=0-83.666000
a=fmtp:96profile-level-id=8;config=000001b008000001b50900012000
m=audio 0 RTP/AVP 98
b=AS:13
a=control: rtsp://example.com/foo/bar/baz.3gp/trackID=2
a=rtpmap:98 AMR/8000
a=range:npt=0-83.660000
a=fmtp:98 octet-align=1
a=maxptime:200
```

The client may further negotiate the offered QoE metrics using OPTIONS or SET_PARAMETER methods.

EXAMPLE 4 (setting the metrics off):

In this example, the metrics are switched off at session level (for all media).

```
C->S, S->C    SET_PARAMETER rtsp://example.com/foo/bar/baz.3gp RTSP/1.0
              Cseq: 302
              Session: 17903320
              3GPP-QoE-Metrics: Off
              Content-length: 0
```

The response for setting the metrics off would be:

```
S->C, C->S    RTSP/1.0 200 OK
              Cseq: 302
              Session: 17903320
              3GPP-QoE-Metrics: Off
```

11.3.4 Sending the metrics feedback with RTSP

QoE Metric feedback with RTSP can be formatted and sent according to clause 5.3.2.3.2.

The following example shows that during the monitoring time 2 corruption periods have occurred. Each value indicates the duration (in milliseconds) of each corruption period.

EXAMPLE 5 (Feedback):

```
C->S          SET_PARAMETER rtsp://example.com/foo/bar/baz.3gp RTSP/1.0
              Cseq: 302
              Session: 17903320
              3GPP-QoE-Feedback:
              url='rtsp://example.com/foo/bar/baz.3gp/trackID=3';Corruption_Duration={200 1300}
              Content-length: 0
```

The following example shows that during the monitoring time 2 corruption periods have occurred. Each values couple indicates the duration (in milliseconds) of each corruption period and the timestamp of the corruption (for example, the first corruption occurred at second 12 and lasted 200 milliseconds).

EXAMPLE 6 (Feedback with timestamps and range):

```
C->S      SET_PARAMETER rtsp://example.com/foo/bar/baz.3gp RTSP/1.0
          Cseq: 302
          Session: 17903320
          3GPP-QoE-Feedback: url='rtsp://example.com/foo/bar/baz.3gp/trackID=1';
          Corruption_Duration={200 12|1300 16};Range:npt=10-20;Content_Switch_Time={1055 34498}
          Content-length: 0
```

In the following example there are no events to report.

EXAMPLE 7 (Feedback with no events):

```
C->S      SET_PARAMETER rtsp://example.com/foo/bar/baz.3gp RTSP/1.0
          Cseq: 302
          Session: 17903320
          3GPP-QoE-Feedback: url='rtsp://example.com/foo/bar/baz.3gp/trackID=3';Corruption_Duration={
          }
          Content-length: 0
```

12 Adaptive HTTP Streaming

12.1 System Description

The 3GPP Adaptive HTTP-Streaming protocol provides a streaming service. This enables delivering content from standard HTTP servers to an HTTP-Streaming client and enables caching content by standard HTTP caches.

Figure 12.1 shows the architecture for Adaptive HTTP streaming. This specification only deals with the specification of interface 1 between the HTTP-Streaming Client and the HTTP-Streaming Server. All other interfaces are out-of-scope of this specification.

It is assumed that the HTTP-Streaming Client has access to a Media Presentation Description (MPD). An MPD provides sufficient information for the HTTP-Streaming Client to provide a streaming service to the user by sequentially downloading media data from an HTTP server and rendering the included media appropriately.

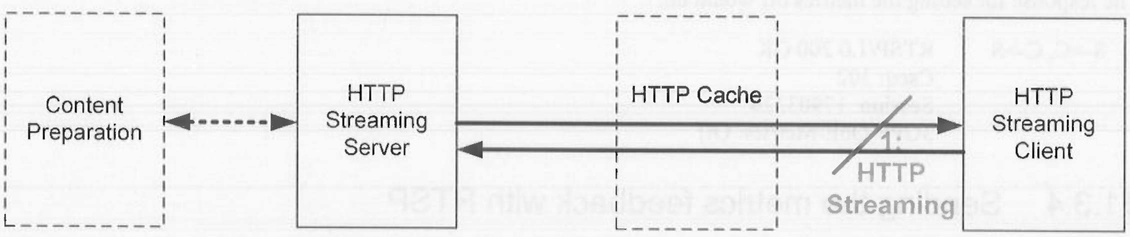


Figure 12.1 System Architecture for Adaptive HTTP Streaming

To initiate the streaming service to the user, the HTTP Streaming Client establishes a Media Presentation by downloading the relevant metadata and subsequently the media data. The Media Presentation is defined in clause 12.2. The protocols used in this specification are specified in clause 12.3. The usage of 3GP file format as media container format is specified in clause 12.4. The codecs are introduced in clause 12.5. Guidelines on the Client Behaviour are presented in clause 12.6. Security-related aspects are addressed in clause 12.7.

12.2 Media Presentation

12.2.1 Introduction

A Media Presentation is a structured collection of data that is accessible to the HTTP-Streaming Client. The HTTP-Streaming client requests and downloads media data information to present the streaming service to the user.

A Media Presentation is described in a Media Presentation Description (MPD) including any possible updates of the MPD. The MPD, including all attributes and elements, is specified in clause 12.2.5.

A Media Presentation fconsists of

- A sequence of one or more Periods.
- Each Period contains one or more Representations from the same media content.
- Each Representation consists of one or more Segments.
- Segments contain media data and/or metadata to decode and present the included media content.

The Media Presentation has a time line that is defined by the concatenation of the timeline of each Period. The timeline in each Period is common to all Representations.

The Media Presentation may be of type OnDemand or Live. The MPD attribute type provides the type of the Media Presentation.

If present, the MPD attribute *availabilityStartTime* gives the earliest time at which the Media Presentation is available at the server.

If present, the MPD attribute *availabilityEndTime* gives the time after which the Media Presentation will no longer be available at the server.

If present, the MPD attribute *mediaPresentationDuration* describes the duration of the entire Media Presentation. If not present, the duration of the Media Presentation is unknown.

12.2.2 Period

A Media Presentation consists of one or more Periods. Periods are defined by *Period* elements in the MPD.

Each Period has an attribute *start*.

For live services, the sum of the *start* attribute of the Period and the MPD attribute *availableStartTime* specifies the availability time of the Period in UTC format, in particular the first Media Segment of each Representation in this Period.

For on-demand services the *start* attribute of the first Period shall be 0. For any other Period the *start* attribute specifies the timeoffset between the start time of the Period relative to the *start* time of the first Period.

Each Period extends until the start of the next Period, or until the end of the Media Presentation in the case of the last Period.

Period start times are precise. They reflect the actual timing resulting from playing the media of all prior Periods.

12.2.3 Representation

Each Period consists of one or more *Representations*.

A Representation is one of the alternative choices of the media content or a subset thereof typically differing by the encoding choice, e.g. by bitrate, resolution, language, codec, etc.

A Representation starts at the *start* of the Period and continues to the end of the Period.

A Representation consists of one or more Segments.

Each Representation either contains an Initialisation Segment or each Media Segment in the Representation is self-initialising.

Each Representation includes one or more media components, where each media component is an encoded version of one individual media type such as audio, video or timed text. Media *components* are time-continuous across boundaries of consecutive Media Segments within one Representation.

Representations are assigned to a group indicated by the *group* attribute. Representations in the same group are alternatives to each other. The media content within one Period is sufficiently represented by:

- either one Representation from group 0, if present,
- or the combination of at most one Representation from each non-zero group.

The timing within each Representation is relative to the start time of the Period that contains this Representation.

12.2.4 Segments

12.2.4.1 Definition

A Segment is defined as a unit that can be uniquely referenced by an http-URL included in the MPD, where an http-URL is defined as an <absolute-URI> according to RFC3986 [60], clause 4.3, with a fixed scheme of 'http://' or https://, possibly restricted by a byte range if the *range* attribute is provided. The byte range is expressed as a *Ranges*-specifier as defined in RFC2616 [17], section 14.35.1 restricted to a single expression identifying a contiguous range of bytes.

The Initialisation Segment contains initialisation information for accessing the Representation. The Initialisation Segment shall not contain any media data.

A Media Segment contains media components that are either described within this Media Segment or described by the Initialisation Segment of this Representation. In addition, a Media Segment

- is assigned a unique MPD URL Element.
- is explicitly or implicitly assigned a start time relative to the start of the Representation provided by the MPD. The client can therefore download the appropriate Segments in regular play-out mode or after seeking. The start time shall be drift-free between the time indicated in the MPD and internal clock of the Media Segments, i.e. the accuracy of the start time documented in the MPD relative to the internal clock does not depend on the position of the Segment in the Representation.
- provides random access information, namely whether this Representation can be randomly accessed within this Segment and if yes, how to randomly access the Media Presentation within this Segment, e.g. exact timing, byte position. There is no requirement that a Media Segment starts with a random access point (RAP), but it is possible to signal in the MPD that all Segments within a Representation start with a RAP. The first Media Segment of a Representation shall always start with a RAP.
- ′ when considered in conjunction with the information and structure of the MPD, contains sufficient information to time-accurately present each contained media component in the Representation without accessing any previous Media Segment in this Representation provided that the Media Segment contains a RAP. The time-accuracy enables seamlessly switching Representations and jointly presenting multiple Representations.
- may contain information for randomly accessing subsets of the Segment by using partial HTTP GET requests.

12.2.4.2 Segment URLs and Media Segment Start Times

12.2.4.2.1 Overview

Each Representation contains exactly one *SegmentInfo* Element that together with a possibly present *SegmentInfoDefault* Element on Period level permits to generate the Segment access information of each Segment within a Representation.

Specifically, the combination of the *SegmentInfo* Element and the *SegmentInfoDefault* Element contains sufficient information to generate a list of Media Segment URLs (possibly restricted by byte ranges) and Media Segment start times relative to the start of the Representation, for example by using the procedures described in Section 12.6.3.

The following rules apply for an MPD:

- URLs within the MPD may be relative or absolute as defined in RFC 3986 [60]. Relative URLs at each level of the MPD are resolved with respect to the *baseURL* attribute specified at that level of the document or the document 'base URI' as defined in RFC3986 Section 5.1 in the case of the *baseURL* attribute at the MPD level.. If only relative URLs are specified and the document base URI cannot be established according to RFC3986 then the MPD cannot be interpreted. *SegmentInfo* elements may contain at most one *InitialisationSegmentURL* element. If no *InitialisationSegmentURL* element is present in a *SegmentInfo* Element element, then each Media Segment within the Representation shall be self-initialising.

- *SegmentInfo* elements may contain at most one *InitialisationSegmentURL* element. If no *InitialisationSegmentURL* element is present in a *SegmentInfo* element, then each Media Segment within the Representation shall be self-initialising.

- The elements *SegmentInfoDefault* and *SegmentInfo* as well as some other timing attributes in the MPD define the URLs (possibly restricted by byte ranges) and approximate time spans within a Period of the available Media Segments. An example segment list generation process as specified in section 12.6.3 generates a valid list of Media Segments.

- Each *SegmentInfo* element shall contain either one *UrlTemplate* element or one or more *Url* elements.
- If a *UrlTemplate* element is present, each *UrlTemplate* element shall contain
    - either exactly one *id* attribute. In this case, the *SegmentInfoDefault* element on Period level shall contain a *sourceUrlTemplatePeriod* attribute with further restrictions as specified in section 12.2.4.2.2. The *duration* attribute shall be present either in the *SegmentInfo* element or in the *SegmentInfoDefault* element.
    - or exactly one *sourceURL* attribute, specifying the URL construction for Media Segments in this Representation. In this case, the *duration* attribute shall be present either in the *SegmentInfo* element or in the *SegmentInfoDefault* element.

- *Url* elements provide a set of explicit URL(s), each of which may contain a *range* attribute, for Media Segments. If more than one *Url* element is provided in the *SegmentInfo* element, the *duration* attribute shall be present either in the *SegmentInfo* element or in the *SegmentInfoDefault* element, where as in case for a single *Url* element, the *duration* attribute may not be present.

- If the duration of the last Media Segment of any Representation in a Period is significantly shorter than the value of the *duration* attribute for this Representation, or if the *duration* attribute for this Representation is not present, then MPD shall either include at least the *start* attribute of the next Period or, in case this is the last Period in the Media Presentation, include the *mediaPresentationDuration* attribute.

12.2.4.2.2 Template-based Media Segment URLs

The *SegmentInfo* Element may contain a *UrlTemplate* element and the *SegmentInfoDefault* element may contain a *sourceUrlTemplatePeriod* attribute. The *sourceURL* attribute of the *UrlTemplate* element or the *sourceUrlTemplatePeriod* attribute represents a string that contains one or more of the identifiers as listed in Table 12.1. The *sourceUrlTemplatePeriod* attribute, when present, shall contain both the *$RepresentationID$* identifier and the *$Index$* identifier. The *sourceURL* attribute, when present, shall contain the *$Index$* identifier and shall not contain the *$RepresentationID$* identifier.

A sub-string "$<Identifier>$" shall name a substitution placeholder matching a mapping key of "<Identifier>". In the request URL, the substitution placeholder shall be replaced by the substitution parameter as defined in Table 12.1. Substitution is performed left to right and identifier matching is case-sensitive. Unrecognized identifiers shall cause the URL formation to fail. In this case the client shall ignore the *Representation* element and processing of the MPD shall continue as if this *Representation* element was not present.

Table 12.1 Identifiers for URL Templates

$<Identifier>$	Substitution parameter
$$	Is an escape sequence, i.e. "$$" is replaced with a single "$"
$RepresentationID$	This identifier is substituted by the attribute *id* of the requested Representation in the MPD.
$Index$	This identifier is substituted by the *Index*, where the Media Segments within a Representation have assigned consecutive Segment indices from *startIndex* to *endIndex*. For an example client using this identifier to construct the list Media Segment URLs, refer to

	section 12.6.3.

12.2.5 Media Presentation Description

12.2.5.1 Introduction

The Media Presentation Description (MPD) contains metadata required by the client to construct appropriate URLs to access Segments and to provide the streaming service to the user.

The Media Presentation may be available in different Representations (different bitrates, languages, resolutions, etc.). The service may be on-demand or live. The MPD contains information that enables the client to build the URLs to access any available Segment (or parts thereof) of the Media Presentation.

The MPD is an XML-document that is formatted according to the XML schema provided in clause 12.2.5.3. The MIME type of the MPD shall be 'video/vnd.3gpp.mpd'. The delivery of the MPD is not in scope of this specification.

If the MPD is delivered over HTTP, then the MPD may be content encoded for transport, as described in [17] using the generic GZip algorithm RFC 1952 [59]. HTTP Streaming clients shall support GZip content decoding of the MPD when delivered over HTTP (GZIP RFC 1952 [59], clause 9).

An adaptive HTTP streaming client shall ignore any XML attributes or elements in a valid XML document formatted according to the XML schema provided in clause 12.2.5.3 that they do not recognize.

12.2.5.2 Media Presentation Description Attributes and Elements

The MPD contains attributes and elements as presented in Table 12.2.

Table 12.2 Semantics of Media Presentation Description (M=Mandatory, O=Optional, OD=Optional with Default Value, CM=Conditionally Mandatory)

Element or Attribute Name	Type (Attribute or Element)	Cardinality	Optionality	Description
MPD	E	1	M	The root element that carries the Media Presentation Description for a Media Presentation.
type	A		OD default: OnDemand	'OnDemand' or 'Live'. Indicates the type of the Media Presentation. Currently, on-demand and live types are defined. If not present, the type of the presentation shall be inferred as OnDemand.
availabilityStartTime	A		CM Must be present for type='Live'	Gives the availability time (in UTC format) of the start of the first Period of the Media Presentation.
availabilityEndTime	A		O	Gives the availability end time (in UTC format). After this time, the Media Presentation described in this MPD is no longer accessible. When not present, the value is unknown.
mediaPresentationDuration	A		O	Specifies the duration of the entire Media Presentation. If the attribute is not present, the duration of the Media Presentation is unknown.
minimumUpdatePeriodMPD	A		O	Provides the minimum period the MPD is updated on the server. If not present the minimum update period is unknown.
minBufferTime	A		M	Provides the minimum amount of initially buffered media that is needed to ensure