

Source: TSG-SA WG4

Title: CRs TS 26.244 on Storage of AMR-WB+ and / or Enhanced
aacPlus audio in 3GP files (Release 6)

Document for: Discussion / Decision

Agenda Item: 7.4.3

At SA#24 the way forward for the selection of audio codecs was formulated in [TD SP-040481](#). TSG SA WG4 was tasked to draft two change requests to the PSS protocol and codecs specification (TS 26.234), one change request for each of the codecs AMR WB+ and enhanced aacPlus to include them for use as PSS codec. The following CRs to TS 26.244 PSS 3GPP file format (3GP) are therefore presented to TSG SA #25 for Discussion / Decision.

Spec	CR	Rev	Phase	Subject	Cat	Vers	WG	Meeting	S4 doc
26.244	002	1	Rel-6	Storage of AMR-WB+ audio in 3GP files	B	6.0.0	S4	TSG-SA WG4#32	S4-040505
26.244	005	1	Rel-6	Storage of Enhanced aacPlus audio in 3GP files	B	6.0.0	S4	TSG-SA WG4#32	S4-040597

CHANGE REQUEST

26.244 CR 002 # rev 1 # Current version: 6.0.0

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

Proposed change affects: UICC apps# ME Radio Access Network Core Network

Title:	# Storage of AMR-WB+ audio in 3GP files		
Source:	# TSG SA WG4		
Work item code:	# PSSrel6-Stage3	Date:	# 14/09/2004
Category:	# B	Release:	# Rel-6
	Use <u>one</u> of the following categories:		Use <u>one</u> of the following releases:
	F (correction)	2	(GSM Phase 2)
	A (corresponds to a correction in an earlier release)	R96	(Release 1996)
	B (addition of feature),	R97	(Release 1997)
	C (functional modification of feature)	R98	(Release 1998)
	D (editorial modification)	R99	(Release 1999)
	Detailed explanations of the above categories can be found in 3GPP TR 21.900 .	Rel-4	(Release 4)
		Rel-5	(Release 5)
		Rel-6	(Release 6)

Reason for change:	# Enable 3GP files to carry AMR-WB+ audio.		
Summary of change:	# Sample entries and sample storage format for AMR-WB+ are defined.		
Consequences if not approved:	# An MMS cannot contain AMR-WB+ clips. Terminals cannot download video clips with AMR-WB+. PSS servers cannot stream AMR-WB+ from 3GP files.		

Clauses affected:	# 2, 3.2, 5.2.1, 5.2.2, 6.1, 6.2, 6.9, 6.10, 10.2										
Other specs affected:	<table border="1" style="display: inline-table; border-collapse: collapse;"> <tr> <td style="text-align: center;">Y</td> <td style="text-align: center;">N</td> </tr> <tr> <td style="text-align: center;"><input type="checkbox"/></td> <td style="text-align: center;"><input checked="" type="checkbox"/></td> </tr> <tr> <td style="text-align: center;"><input type="checkbox"/></td> <td style="text-align: center;"><input checked="" type="checkbox"/></td> </tr> <tr> <td style="text-align: center;"><input type="checkbox"/></td> <td style="text-align: center;"><input checked="" type="checkbox"/></td> </tr> </table>	Y	N	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	Other core specifications	#
Y	N										
<input type="checkbox"/>	<input checked="" type="checkbox"/>										
<input type="checkbox"/>	<input checked="" type="checkbox"/>										
<input type="checkbox"/>	<input checked="" type="checkbox"/>										
		Test specifications									
		O&M Specifications									
Other comments:	#										

How to create CRs using this form:

Comprehensive information and tips about how to create CRs can be found at <http://www.3gpp.org/specs/CR.htm>.

Below is a brief summary:

- 1) Fill out the above form. The symbols above marked # contain pop-up help information about the field that they are closest to.
- 2) Obtain the latest version for the release of the specification to which the change is proposed. Use the MS Word "revision marks" feature (also known as "track changes") when making the changes. All 3GPP specifications can be downloaded from the 3GPP server under <ftp://ftp.3gpp.org/specs/> For the latest version, look for the directory name with the latest date e.g. 2001-03 contains the specifications resulting from the March 2001 TSG meetings.
- 3) With "track changes" disabled, paste the entire CR form (use CTRL-A to select it) into the specification just in front of the clause containing the first piece of changed text. Delete those parts of the specification which are not relevant to the change request.

2 References

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

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

- [1] 3GPP TS 22.233: "Transparent End-to-End Packet-switched Streaming Service; Stage 1".
- [2] 3GPP TS 26.233: "Transparent end-to-end packet switched streaming service (PSS); General description".
- [3] 3GPP TS 26.234: "Transparent end-to-end packet switched streaming service (PSS); Protocols and codecs".
- [4] 3GPP TS 26.245: "Transparent end-to-end packet switched streaming service (PSS); Timed text format".
- [5] 3GPP TS 26.246: "Transparent end-to-end packet switched streaming service (PSS); 3GPP SMIL Language Profile".
- [6] 3GPP TR 21.905: "Vocabulary for 3GPP Specifications".
- [7] ISO/IEC 14496-12:2003 | 15444-12:2003: "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".
- [8] 3GPP TS 26.140: "Multimedia Messaging Service (MMS); Media formats and codecs".
- [9] ITU-T Recommendation H.263 (1998): "Video coding for low bit rate communication".
- [10] ISO/IEC 14496-2:2001: "Information technology – Coding of audio-visual objects – Part 2: Visual".
- [11] 3GPP TS 26.071: "Mandatory Speech CODEC speech processing functions; AMR Speech CODEC; General description".
- [12] 3GPP TS 26.171: "AMR Wideband Speech Codec; General Description".
- [13] ISO/IEC 14496-3:2001: "Information technology – Coding of audio-visual objects – Part 3: Audio".
- [14] ISO/IEC 14496-14:2003: "Information technology – Coding of audio-visual objects – Part 14: MP4 file format".
- [15] IETF RFC 3267: "Real-Time Transport Protocol (RTP) Payload Format and File Storage Format for the Adaptive Multi-Rate (AMR) Adaptive Multi-Rate Wideband (AMR-WB) Audio Codecs", Sjöberg J. et al., June 2002.
- [16] 3GPP TS 26.101: "Mandatory Speech Codec speech processing functions; Adaptive Multi-Rate (AMR) speech codec frame structure".
- [17] 3GPP TS 26.201: "Speech Codec speech processing functions; AMR Wideband Speech Codec; Frame Structure".

- [18] ITU-T Recommendation H.263 – Annex X (2001): "Annex X: Profiles and levels definition".
- [19] IETF RFC 3711: "The Secure Real-time Transport Protocol", Baugher M. et al., Feb 2004.
- [20] ISO/IEC 14496-15: "Information technology – Coding of audio-visual objects – Part 15: Advanced Video Coding (AVC) file format".
- [21] [3GPP TS 26.290: "Extended AMR Wideband codec; Transcoding functions"](#).
- [22] [IETF Internet Draft: "Real-Time Transport Protocol \(RTP\) Payload Format for Extended AMR Wideband \(AMR-WB+\) Audio Codec", Sjoberg J., Westerlund M. and Lakaniemi A., <http://www.ietf.org/internet-drafts/draft-ietf-avt-rtp-amrwbplus-01.txt>, July 2004.](#)

< cut text >

3.2 Abbreviations

For the purposes of the present document, the abbreviations given in 3GPP TR 21.905 [6] and the following apply.

3GP	3GPP file format
AAC	Advanced Audio Coding
AMR-WB+	Extended Adaptive Multi-Rate Wideband Codec
BIFS	Binary Format for Scenes
ITU-T	International Telecommunications Union – Telecommunications
MIME	Multipurpose Internet Mail Extensions
MMS	Multimedia Messaging Service
MP4	MPEG-4 file format
PSS	Packet-switched Streaming Service
RTP	Real-time Transport Protocol
RTSP	Real-Time Streaming Protocol
SDP	Session Description Protocol
SRTSP	Secure Real-time Transport Protocol

< cut text >

5.2.1 Limitations to the ISO base media file format

The following limitations to the ISO base media file format [7] shall apply to a 3GP file:

- compact sample sizes ('stz2') shall not be used [for tracks containing H.263, MPEG-4 video, AMR, AMR-WB, AAC or Timed text](#);
- movie fragments shall not be used.

5.2.2 Registration of codecs

Code streams for H.263 video [9], MPEG-4 video [10], AMR narrow-band speech [11], AMR wide-band speech [12], [Extended AMR wide-band audio \[21\]](#), MPEG-4 AAC audio [13], and timed text [4] can be included in 3GP files as described in clause 6 of the present document.

< cut text >

6 Codec registration

6.1 General

The purpose of this clause is to define the necessary structure for integration of the H.263, MPEG-4 video, AMR, AMR-WB, [Extended AMR-WB \(AMR-WB+\)](#), and AAC media specific information in a 3GP file. Clause 6.2 gives some background information about the Sample Description box in the ISO base media file format [7] and clauses 6.3 and 6.4 about the MP4VisualSampleEntry box and the MP4AudioSampleEntry box in the MPEG-4 file format [14]. The definitions of the Sample Entry boxes for AMR, AMR-WB, [AMR-WB+](#) and H.263 are given in clauses 6.5 to 6.8. [10](#). The integration of timed text in a 3GP file is specified in [4].

AMR and AMR-WB data is stored in the stream according to the AMR and AMR-WB storage format for single channel header of Annex E [15], without the AMR magic numbers.

[The 3GPP file format is the native storage format AMR-WB+. The data stream, stored in samples of a 3GP file, shall be formatted according to clause 8.3 of \[21\] where the syntax of a sample for AMR-WB+ is defined by referring to the RTP payload format definition \[22\] and its basic mode.](#)

6.2 Sample Description box

In an ISO file, Sample Description Box gives detailed information about the coding type used, and any initialisation information needed for that coding. The Sample Description Box can be found in the ISO file format Box Structure Hierarchy shown in figure 6.1.

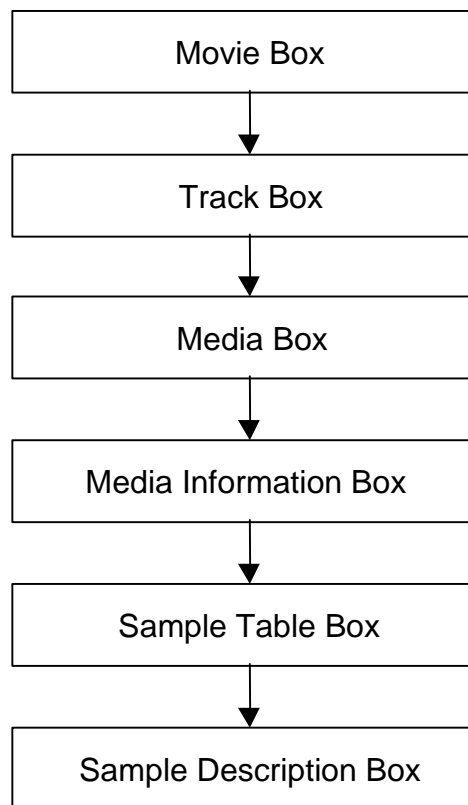


Figure 6.1: ISO File Format Box Structure Hierarchy

The Sample Description Box can have one or more Sample Entries. Valid Sample Entries already defined for ISO and MP4 include MP4AudioSampleEntry, MP4VisualSampleEntry and HintSampleEntry. The Sample Entries for AMR and AMR-WB shall be AMRSampleEntry, [for AMR-WB+ it shall be AMRWPSampleEntry](#), for H.263 it shall be H263SampleEntry, and for timed text it shall be TextSampleEntry.

The format of SampleEntry and its fields are explained as follows:

SampleEntry ::= MP4VisualSampleEntry |
 MP4AudioSampleEntry |
 AMRSampleEntry |
[AMRWPSampleEntry](#) |
 H263SampleEntry |
 TextSampleEntry |
 HintSampleEntry

Table 6.1: SampleEntry fields

Field	Type	Details	Value
MP4VisualSampleEntry		Entry type for visual samples defined in the MP4 specification.	
MP4AudioSampleEntry		Entry type for audio samples defined in the MP4 specification.	
AMRSampleEntry		Entry type for AMR and AMR-WB speech samples defined in clause 6.5 of the present document.	
AMRWPSampleEntry		Entry type for AMR-WB+ audio samples defined in clause 6.9 of the present document.	
H263SampleEntry		Entry type for H.263 visual samples defined in clause 6.6 of the present document.	
TextSampleEntry		Entry type for timed text samples defined in the timed text specification	
HintSampleEntry		Entry type for hint track samples defined in the ISO specification.	

From the above 6 Sample Entries, only the MP4VisualSampleEntry, MP4AudioSampleEntry, H263SampleEntry, ~~and~~ AMRSampleEntry [and AMRWPSampleEntry](#) are taken into consideration here. TextSampleEntry is defined in [4] and HintSampleEntry in [7].

< cut text >

6.9 AMRWPSampleEntry box

[The box type of the AMRWPSampleEntry Box shall be 'sawp'.](#)

[The AMRWPSampleEntry Box is defined as follows:](#)

```

AMRWPSampleEntry ::= BoxHeader
    Reserved_6
    Data-reference-index
    Reserved_8
    Reserved_2
    Reserved_2
    Reserved_4
    TimeScale
    Reserved_2
    AMRWPSpecificBox

```

Table 6.9: AMRWPSampleEntry fields

Field	Type	Details	Value
BoxHeader.Size	Unsigned int(32)		
BoxHeader.Type	Unsigned int(32)		'sawp'
Reserved_6	Unsigned int(8) [6]		0
Data-reference-index	Unsigned int(16)	Index to a data reference that to use to retrieve the sample data. Data references are stored in data reference boxes.	
Reserved_8	Const unsigned int(32) [2]		0
Reserved_2	Const unsigned int(16)		2
Reserved_2	Const unsigned int(16)		16
Reserved_4	Const unsigned int(32)		0
TimeScale	Unsigned int(16)	Copied from media header box of this media	
Reserved_2	Const unsigned int(16)		0
AMRWPSpecificBox		Information specific to the AMR-WB+ decoder.	

If one compares the MP4AudioSampleEntry Box - AMRWPSampleEntry Box the main difference is in the replacement of the ESDBox, which is specific to MPEG-4 systems, with a box suitable for AMR-WB+. The **AMRWPSpecificBox** field structure is described in clause 6.10.

NOTE 1: In order to maintain backward compatibility with Release 4 and 5, the AMRWPSampleEntry should not be used for AMR-WB+ streams that only contain AMR-WB modes. Such streams should be stored as AMR-WB, i.e. by using the AMRSampleEntry with box type 'sawb', defined in clause 6.5, and the storage format for single channel header of Annex E [15], without the AMR magic numbers. This way file readers of previous releases will always be able to read AMR-WB streams stored in 3GP files.

NOTE 2: In order to enhance interoperability in Release 6, file readers capable of parsing tracks with AMR-WB+ should also be capable of parsing AMR-WB tracks (see note 1).

6.10 AMRWPSpecificBox field for AMRWPSampleEntry box

The AMRWPSpecificBox fields for AMR-WB+ shall be as defined in table 6.10. The AMRWPSpecificBox for the AMRWPSampleEntry Box shall always be included if the 3GP file contains AMR-WB+ media.

Table 6.10: The AMRWPSpecificBox fields for AMRWPSampleEntry

Field	Type	Details	Value
BoxHeader.Size	Unsigned int(32)		
BoxHeader.Type	Unsigned int(32)		'dawp'
DecSpecificInfo	AMRWPDDecSpecStruc	Structure which holds the AMR-WB+ Specific information	

BoxHeader Size and Type: indicate the size and type of the AMR-WB+ decoder-specific box. The type must be 'dawp'.

DecSpecificInfo: the structure where the AMR-WB+ stream specific information resides.

The AMRWPDDecSpecStruc is defined as follows:

```
struct AMRWPDDecSpecStruc{
    _____ Unsigned int (32)   vendor
    _____ Unsigned int (8)    decoder version
}
```

The definitions of AMRWPDDecSpecStruc members are as follows:

vendor: four character code of the manufacturer of the codec, e.g. 'VXYZ'. The vendor field gives information about the vendor whose codec is used to create the encoded data. It is an informative field, which may be used by the decoding end. If a manufacturer already has a four-character code, it is recommended that it uses the same code in this field. Else, it is recommended that the manufacturer creates a four character code which best addresses the manufacturer's name. It can be safely ignored.

decoder version: version of the vendor's decoder which can decode the encoded stream in the best (i.e. optimal) way. This field is closely tied to the vendor field. It may give advantage to the vendor which has optimal encoder-decoder version pairs. The value is set to 0 if decoder version has no importance for the vendor. It can be safely ignored.

< cut text >

10.2 Sample entries for encrypted media tracks

The sample entries stored in the sample description box of a media track in a 3GP file identify the format of the encoded media, i.e. codec and other coding parameters. All valid sample entries for unencrypted media in a 3GP file are described in Clause 6. The principle behind storing encrypted media in a track is to “disguise” the original sample entry with a generic sample entry for encrypted media. Table 10.1 gives an overview of the formats (identifying sample entries) that can be used in 3GP files for signalling encrypted video, audio and text.

Table 10.1: Formats for encrypted media tracks

Format	Original format	Media content
'encv'	's263', 'mp4v', ...	encrypted video: H.263, MPEG-4 visual, ...
'enca'	'samr', 'sawb', ' sawp ', 'mp4a', ...	encrypted audio: AMR, AMR-WB, AMR-WB+ , AAC, ...
'enct'	'tx3g', ...	encrypted text: timed text, ...

- [20] ISO/IEC 14496-15: "Information technology – Coding of audio-visual objects – Part 15: Advanced Video Coding (AVC) file format".
- [21] [3GPP TS 26.401: "General audio codec audio processing functions; Enhanced aacPlus general audio codec; General description"](#).
- [22] [3GPP TS 26.410: "General audio codec audio processing functions; Enhanced aacPlus general audio codec; Floating-point ANSI-C code"](#).
- [23] [3GPP TS 26.xxx: "General audio codec audio processing functions; Enhanced aacPlus general audio codec; Fixed-point ANSI-C code"](#).
- [24] [ISO/IEC 14496-3:2001/Amd.1:2003, Bandwidth Extension](#)

3.2 Abbreviations

For the purposes of the present document, the abbreviations given in 3GPP TR 21.905 [6] and the following apply.

3GP	3GPP file format
AAC	Advanced Audio Coding
BIFS	Binary Format for Scenes
Enhanced aacPlus	MPEG-4 High Efficiency AAC plus MPEG-4 Parametric Stereo
ITU-T	International Telecommunications Union – Telecommunications
MIME	Multipurpose Internet Mail Extensions
MMS	Multimedia Messaging Service
MP4	MPEG-4 file format
PSS	Packet-switched Streaming Service
RTP	Real-time Transport Protocol
RTSP	Real-Time Streaming Protocol
SDP	Session Description Protocol
SRTP	Secure Real-time Transport Protocol

5.2.2 Registration of codecs

Code streams for H.263 video [9], MPEG-4 video [10], AMR narrow-band speech [11], AMR wide-band speech [12], [Enhanced aacPlus audio](#) [21, 22, 23], MPEG-4 AAC audio [13], and timed text [4] can be included in 3GP files as described in clause 6 of the present document.

6.1 General

The purpose of this clause is to define the necessary structure for integration of the H.263, MPEG-4 video, AMR, AMR-WB, [Enhanced aacPlus](#) and AAC media specific information in a 3GP file. Clause 6.2 gives some background information about the Sample Description box in the ISO base media file format [7] and clauses 6.3 and 6.4 about the MP4VisualSampleEntry box and the MP4AudioSampleEntry box in the MPEG-4 file format [14]. The definitions of the Sample Entry boxes for AMR, AMR-WB and H.263 are given in clauses 6.5 to 6.8. The integration of timed text in a 3GP file is specified in [4].

AMR and AMR-WB data is stored in the stream according to the AMR and AMR-WB storage format for single channel header of Annex E [15], without the AMR magic numbers.

6.4 MP4AudioSampleEntry box

MP4AudioSampleEntryBox is defined as follows:

```

MP4AudioSampleEntry ::= BoxHeader
    Reserved_6
    Data-reference-index
    Reserved_8
    Reserved_2
    Reserved_2
    Reserved_4
    TimeScale
    Reserved_2
ESDBox
  
```

Table 6.3: MP4AudioSampleEntry fields

Field	Type	Details	Value
BoxHeader .Size	Unsigned int(32)		
BoxHeader .Type	Unsigned int(32)		'mp4a'
Reserved_6	Unsigned int(8) [6]		0
Data-reference-index	Unsigned int(16)	Index to a data reference that to use to retrieve the sample data. Data references are stored in data reference boxes.	
Reserved_8	Const unsigned int(32) [2]		0
Reserved_2	Const unsigned int(16)		2
Reserved_2	Const unsigned int(16)		16
Reserved_4	Const unsigned int(32)		0
TimeScale	Unsigned int(16)	Copied from track	
Reserved_2	Const unsigned int(16)		0
ESDBox		Box containing an elementary stream descriptor for this stream.	

The stream type specific information is in the ESDBox structure, as defined in [14]. [Enhanced aacPlus stored in .3GP files shall not use implicit signalling \(as defined in \[24\]\).](#)

10.2 Sample entries for encrypted media tracks

The sample entries stored in the sample description box of a media track in a 3GP file identify the format of the encoded media, i.e. codec and other coding parameters. All valid sample entries for unencrypted media in a 3GP file are described in Clause 6. The principle behind storing encrypted media in a track is to “disguise” the original sample entry with a generic sample entry for encrypted media. Table 10.1 gives an overview of the formats (identifying sample entries) that can be used in 3GP files for signalling encrypted video, audio and text.

Table 10.1: Formats for encrypted media tracks

Format	Original format	Media content
'encv'	's263', 'mp4v', ...	encrypted video: H.263, MPEG-4 visual, ...
'enca'	'samr', 'sawb', 'mp4a', ...	encrypted audio: AMR, AMR-WB, Enhanced aacPlus , AAC, ...
'enct'	'tx3g', ...	encrypted text: timed text, ...

The generic sample entries for encrypted media replicate the original sample entries and include a Protection scheme information box with details on the original format, as well as all requirements for decrypting the encoded media. The EncryptedVideoSampleEntry and the EncryptedAudioSampleEntry are defined in Tables 10.2 and 10.3, where the ProtectionSchemeInfoBox (defined in clause 10.2) is simply added to the list of boxes contained in a sample entry.