ETSI TS 126 117 V17.2.0 (2024-04)



5G; 5G Media Streaming (5GMS); Speech and audio profiles (3GPP TS 26.117 version 17.2.0 Release 17)



Reference RTS/TSGS-0426117vh20 Keywords 5G

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 - APE 7112B Association à but non lucratif enregistrée à la Sous-Préfecture de Grasse (06) N° w061004871

Important notice

The present document can be downloaded from: https://www.etsi.org/standards-search

The present document may be made available in electronic versions and/or in print. The content of any electronic and/or print versions of the present document shall not be modified without the prior written authorization of ETSI. In case of any existing or perceived difference in contents between such versions and/or in print, the prevailing version of an ETSI deliverable is the one made publicly available in PDF format at www.etsi.org/deliver.

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 https://portal.etsi.org/TB/ETSIDeliverableStatus.aspx

If you find errors in the present document, please send your comment to one of the following services: https://portal.etsi.org/People/CommiteeSupportStaff.aspx

If you find a security vulnerability in the present document, please report it through our Coordinated Vulnerability Disclosure Program:

https://www.etsi.org/standards/coordinated-vulnerability-disclosure

Notice of disclaimer & limitation of liability

The information provided in the present deliverable is directed solely to professionals who have the appropriate degree of experience to understand and interpret its content in accordance with generally accepted engineering or other professional standard and applicable regulations.

No recommendation as to products and services or vendors is made or should be implied.

No representation or warranty is made that this deliverable is technically accurate or sufficient or conforms to any law and/or governmental rule and/or regulation and further, no representation or warranty is made of merchantability or fitness for any particular purpose or against infringement of intellectual property rights.

In no event shall ETSI be held liable for loss of profits or any other incidental or consequential damages.

Any software contained in this deliverable is provided "AS IS" with no warranties, express or implied, including but not limited to, the warranties of merchantability, fitness for a particular purpose and non-infringement of intellectual property rights and ETSI shall not be held liable in any event for any damages whatsoever (including, without limitation, damages for loss of profits, business interruption, loss of information, or any other pecuniary loss) arising out of or related to the use of or inability to use the software.

Copyright Notification

No part may be reproduced or utilized in any form or by any means, electronic or mechanical, including photocopying and microfilm except as authorized by written permission of ETSI.

The content of the PDF version shall not be modified without the written authorization of ETSI.

The copyright and the foregoing restriction extend to reproduction in all media.

© ETSI 2024. All rights reserved.

Intellectual Property Rights

Essential patents

IPRs essential or potentially essential to normative deliverables may have been declared to ETSI. The declarations pertaining to these essential IPRs, if any, are 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 (https://ipr.etsi.org/).

Pursuant to the ETSI Directives including the ETSI IPR Policy, no investigation regarding the essentiality of IPRs, 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.

Trademarks

The present document may include trademarks and/or tradenames which are asserted and/or registered by their owners. ETSI claims no ownership of these except for any which are indicated as being the property of ETSI, and conveys no right to use or reproduce any trademark and/or tradename. Mention of those trademarks in the present document does not constitute an endorsement by ETSI of products, services or organizations associated with those trademarks.

DECTTM, **PLUGTESTS**TM, **UMTS**TM and the ETSI logo are trademarks of ETSI registered for the benefit of its Members. **3GPP**TM and **LTE**TM are trademarks of ETSI registered for the benefit of its Members and of the 3GPP Organizational Partners. **oneM2M**TM logo is a trademark of ETSI registered for the benefit of its Members and of the oneM2M Partners. **GSM**[®] and the GSM logo are trademarks registered and owned by the GSM Association.

Legal Notice

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. These shall be interpreted as being references to the corresponding ETSI deliverables.

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

Modal verbs terminology

In the present document "shall", "shall not", "should", "should not", "may", "need not", "will", "will not", "can" and "cannot" are to be interpreted as described in clause 3.2 of the <u>ETSI Drafting Rules</u> (Verbal forms for the expression of provisions).

"must" and "must not" are NOT allowed in ETSI deliverables except when used in direct citation.

Contents

Intelle	ectual Property Rights	2
Legal	Notice	2
•	ıl verbs terminology	
	vord	
1	Scope	
2	References	
3	Definitions of terms, symbols and abbreviations	
3.1	Terms	
3.1		
3.2 3.3	Symbols	
4	Overview	
5	Media Capabilities	10
5.1	Introduction	
5.2	Decoding Capabilities	
5.3	Encoding Capabilities	10
6	Operation Points	11
6.1	Introduction	
6.2	Speech Operation Points	
6.2.1	Introduction	
6.2.2	AMR	
6.2.2.1		
6.2.2.2		
6.2.2.3	1	
6.2.3	AMR-WB	
6.2.3.1		
6.2.3.2	1	
6.2.3.3	1	
6.2.4	EVS	
6.2.4.1		
6.2.4.2	The state of the s	
6.2.4.3	1	
6.3	Audio Operation Points	
6.3.1	Introduction.	
6.3.2		13
6.3.2.1		-
6.3.2.2		
6.3.2.3	1	
6.3.2.3	AMR-WB+	
6.3.3.1		
6.3.3.1		
6.3.3.3	1	
7	Mapping to 5GMS delivery	
7.1	Introduction	
7.1	AMR Media Profile	
7.2.1	Mapping to ISO BMFF	
7.2.1	Media Profile Definition	
7.2.2 7.2.2.1		
7.2.2.1 7.2.2.2		
7.2.2.2 7.2.2.3	e e e e e e e e e e e e e e e e e e e	
7.2.2.3 7.2.2.4		
7.2.2.4 7.2.2.5	7 1	
7.2.2.3 7.3	AMR-WB Media Profile	
1.5	2 11/11C 17 D 1/11/UH 1 1/UH/	

	ry				
Annes	ex B (informative): Change history	23			
A.1	3GPP Registered URIs	22			
Anne	ex A (informative): Registration Information	22			
7.7.2.5	5 Content Generation Requirements	21			
7.7.2.4	•				
7.7.2.3	11 6				
7.7.2.2	$\boldsymbol{\mathcal{U}}$				
7.7.2.1					
7.7.2	Media Profile Definition				
7.7.1	Mapping to ISO BMFF				
7.7	AMR-WB+ Media Profiles				
7.6.2.5	1 · · · · · · · · · · · · · · · · · · ·				
7.6.2.4	.,				
7.6.2.3	Tr 8				
7.6.2.2					
7.6.2.1					
7.6	eAAC+ stereo Media Profile				
7.5	void				
7.4.2.5	5 Content Generation Requirements	18			
7.4.2.4	<i>y</i> 1				
7.4.2.3	Mapping to DASH Adaptation Set	17			
7.4.2.2	2 CMAF Switching Set and Media Profile Definition	17			
7.4.2.1	1 CMAF Track Definition	17			
7.4.2	Media Profile Definition				
7.4.1	Mapping to ISO BMFF				
7.4	EVS Media Profile				
7.3.2.5	, 1				
7.3.2.4	TI 8				
7.3.2.3	· · · · · · · · · · · · · · · · · · ·				
7.3.2.2					
7.3.2.1					
7.3.1 7.3.2	Media Profile Definition				
7.3.1	Mapping to ISO BMFF	14			

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 document.

In the present document, modal verbs have the following meanings:

shall indicates a mandatory requirement to do something

shall not indicates an interdiction (prohibition) to do something

The constructions "shall" and "shall not" are confined to the context of normative provisions, and do not appear in Technical Reports.

The constructions "must" and "must not" are not used as substitutes for "shall" and "shall not". Their use is avoided insofar as possible, and they are not used in a normative context except in a direct citation from an external, referenced, non-3GPP document, or so as to maintain continuity of style when extending or modifying the provisions of such a referenced document.

should indicates a recommendation to do something

should not indicates a recommendation not to do something

may indicates permission to do something

need not indicates permission not to do something

The construction "may not" is ambiguous and is not used in normative elements. The unambiguous constructions "might not" or "shall not" are used instead, depending upon the meaning intended.

can indicates that something is possible

cannot indicates that something is impossible

The constructions "can" and "cannot" are not substitutes for "may" and "need not".

will indicates that something is certain or expected to happen as a result of action taken by an agency

the behaviour of which is outside the scope of the present document

will not indicates that something is certain or expected not to happen as a result of action taken by an

agency the behaviour of which is outside the scope of the present document

might indicates a likelihood that something will happen as a result of action taken by some agency the

behaviour of which is outside the scope of the present document

might not indicates a likelihood that something will not happen as a result of action taken by some agency

the behaviour of which is outside the scope of the present document

In addition:

is (or any other verb in the indicative mood) indicates a statement of fact

is not (or any other negative verb in the indicative mood) indicates a statement of fact

The constructions "is" and "is not" do not indicate requirements.

1 Scope

The present document specifies speech and audio media capabilities, operation points and media profiles for 5G Media Streaming in the context of 3GPP services and deployments. Speech and audio media capabilities, operation points and media profiles are also provided for usage in other streaming applications.

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*.
- 3GPP TR 21.905: "Vocabulary for 3GPP Specifications". [1] 3GPP TS 26.501: "5G Media Streaming (5GMS); General description and architecture". [2] [3] 3GPP TS 26.071: "Mandatory Speech Codec speech processing functions; AMR Speech CODEC; General description". 3GPP TS 26.090: "Mandatory Speech Codec speech processing functions; Adaptive Multi-Rate [4] (AMR) speech codec; Transcoding functions". 3GPP TS 26.073: "ANSI-C code for the Adaptive Multi Rate (AMR) speech codec". [5] 3GPP TS 26.104: "ANSI-C code for the floating-point Adaptive Multi Rate (AMR) speech codec". [6] [7] 3GPP TS 26.093: "Mandatory speech codec speech processing functions; Adaptive Multi-Rate (AMR) speech codec; Source controlled rate operation". [8] 3GPP TS 26.171: "Speech codec speech processing functions; Adaptive Multi-Rate - Wideband (AMR-WB) speech codec; General description". 3GPP TS 26.190: "Speech codec speech processing functions; Adaptive Multi-Rate - Wideband [9] (AMR-WB) speech codec; Transcoding functions". [10] 3GPP TS 26.173: "ANCI-C code for the Adaptive Multi Rate - Wideband (AMR-WB) speech codec". [11] 3GPP TS 26.204: "Speech codec speech processing functions; Adaptive Multi-Rate - Wideband (AMR-WB) speech codec; ANSI-C code". 3GPP TS 26.193: "Speech codec speech processing functions; Adaptive Multi-Rate - Wideband [12] (AMR-WB) speech codec; Source controlled rate operation". [13] 3GPP TS 26.441: "Codec for Enhanced Voice Services (EVS); General Overview". 3GPP TS 26.442: "Codec for Enhanced Voice Services (EVS); ANSI C code (fixed-point)". [14] 3GPP TS 26.443: "Codec for Enhanced Voice Services (EVS); ANSI C code (floating-point)". [15] 3GPP TS 26.445: "Codec for Enhanced Voice Services (EVS); Detailed Algorithmic Description". [16] [17] 3GPP TS 26.446: "Codec for Enhanced Voice Services (EVS); Adaptive Multi-Rate - Wideband (AMR-WB) backward compatible functions".

[18]	3GPP TS 26.450: "Codec for Enhanced Voice Services (EVS); Discontinuous Transmission (DTX)".
[19]	3GPP TS 26.401: "General audio codec audio processing functions; Enhanced aacPlus general audio codec; General description".
[20]	3GPP TS 26.402: "General audio codec audio processing functions; Enhanced aacPlus general audio codec; Additional decoder tools".
[21]	3GPP TS 26.403: "General audio codec audio processing functions; Enhanced aacPlus general audio codec; Encoder specification; Advanced Audio Coding (AAC) part".
[22]	3GPP TS 26.404: "General audio codec audio processing functions; Enhanced aacPlus general audio codec; Enhanced aacPlus encoder Spectral Band Replication (SBR) part".
[23]	3GPP TS 26.405: "General audio codec audio processing functions; Enhanced aacPlus general audio codec; Encoder specification parametric stereo part".
[24]	3GPP TS 26.410: "General audio codec audio processing functions; Enhanced aacPlus general audio codec; Floating-point ANSI-C code".
[25]	3GPP TS 26.411: "General audio codec audio processing functions; Enhanced aacPlus general audio codec; Fixed-point ANSI-C code".
[26]	3GPP TS 26.290: "Audio codec processing functions; Extended Adaptive Multi-Rate - Wideband (AMR-WB+) codec; Transcoding functions".
[27]	3GPP TS 26.304: "Extended Adaptive Multi-Rate - Wideband (AMR-WB+) codec; Floating-point ANSI-C code".
[28]	3GPP TS 26.273: "ANSI-C code for the fixed-point Extended Adaptive Multi-Rate - Wideband (AMR-WB+) speech codec".
[29]	3GPP TS 26.244: "Transparent end-to-end streaming service; 3GPP file format (3GP)".
[30]	ISO/IEC 23000-19: "Information Technology Multimedia Application Format (MPEG-A) – Part 19: Common Media Application Format (CMAF) for segmented media".
[31]	ISO/IEC 23009-1:2019/Amd.1:2020: "Information technology Dynamic adaptive streaming over HTTP (DASH) Part 1: Media presentation description and segment formats."
[32]	CTA-5003: "Web Application Video Ecosystem (WAVE): Device Playback Capabilities Specification" available here https://cdn.cta.tech/cta/media/media/resources/standards/pdfs/cta-5003-final.pdf.
[33]	3GPP TS 26.244: "Transparent end-to-end streaming service; 3GPP file format (3GP)".
[34]	3GPP TS 26.452: "Codec for Enhanced Voice Services (EVS); ANSI C code; Alternative fixed-point using updated basic operators".
[35]	3GPP TS 26.447: "Codec for Enhanced Voice Services (EVS); Error concealment of lost packets".
[36]	3GPP TS 26.511: "5G Media Streaming (5GMS); Profiles, Codecs and Formats".
[37]	ISO/IEC 14496-12: "Information technology - Coding of audio-visual objects - Part 12: ISO base media file format"
[38]	ISO/IEC 14496-14 "Information technology – Coding of audio-visual objects – Part 14: MP4 file format".

3 Definitions of terms, symbols and abbreviations

3.1 Terms

For the purposes of the present document, the terms given in 3GPP TR 21.905 [1] and the following apply. A term defined in the present document takes precedence over the definition of the same term, if any, in 3GPP TR 21.905 [1].

Bitstream: A media bitstream that conforms to an audio/speech encoding format and certain Operation Point.

Media Profile: A combination of a Bitstream encapsulated into a media container suitable for 5G Media Streaming Delivery.

Operation Point: A collection of discrete combinations of different content formats and the encoding format.

Receiver: A receiver that can decode and render any bitstream that is conforming to a certain Operation Point.

Sender: An entity that can process and encode formats associated to an Operation Point.

3.2 Symbols

Void.

3.3 Abbreviations

For the purposes of the present document, the abbreviations given in 3GPP TR 21.905 [1] and the following apply. An abbreviation defined in the present document takes precedence over the definition of the same abbreviation, if any, in 3GPP TR 21.905 [1].

AAC Advanced Audio Coding
ABR Adaptive BitRate
AMR Adaptive MultiRate

CMAF Common Media Application Format
DASH Dynamic Adaptive Streaming over HTTP

DTX Discontinuous Transmission
EVS Enhanced Voice Services
ISO BMFF ISO Base Media File Format
HTTP Hyper Text Transfer Protocol
SBR Spectral Band Replication
URN Universal Resource Name

WAVE Web Application Video Ecosystem

4 Overview

The speech and audio media capabilities defined in this specification are primarily introduced in order to be used as content format in the context of 5G Media Streaming, but not restricted to this use case. Parameters for audio encoder/decoder, content format and transport are defined.

The present document defines:

- Media decoding capabilities: the requirements for a receiver in terms of decoding
- Media encoding capabilities: the requirements for a sender in terms of encoding
- Operation Points: A collection of discrete combinations of different content formats and the encoding formats. Operation Points are supported by
 - Bitstream Requirements: A media bitstream that conforms to an audio or speech encoding format and certain Operation Point.

- Receiver Requirements: A function that can decode and playback any Bitstream that is conforming to a certain Operation Point in real-time.
- Sender Requirements: A function that can process and encode any Bitstream that is conforming to a certain Operation Point in real-time.
- The integration of each Operation Point in 5G Media Streaming as defined in TS 26.501 [2] and TS 26.511 [36].

5 Media Capabilities

5.1 Introduction

This clause documents speech and audio media capabilities in terms of decoding capabilities.

5.2 Decoding Capabilities

The following speech media decoding capabilities are defined:

- *AMR*: All decoding requirements for the AMR speech codec as specified in 3GPP TS 26.071 [3], 3GPP TS 26.090 [4], 3GPP TS 26.073 [5] and 3GPP TS 26.104 [6]) including all 8 modes and source-controlled rate operation 3GPP TS 26.093 [7].
- *AMR-WB*: All decoding requirements for the AMR-WB codec as specified in 3GPP TS 26.171 [8], 3GPP TS 26.190 [9], 3GPP TS 26.173 [10] and 3GPP TS 26.204 [11] including all 9 modes and source-controlled rate operation 3GPP TS 26.193 [12].
- *EVS*: All decoding requirements for the EVS codec as specified in 3GPP TS 26.441 [13], 3GPP TS 26.445 [16], 3GPP TS 26.442 [14] and 3GPP TS 26.443 [15] as described below including functions for backwards compatibility with AMR-WB (3GPP TS 26.446 [17]) and discontinuous transmission (3GPP TS 26.450 [18]).

The following audio media decoding capabilities are defined:

- *eAAC*+: All decoding requirements for the eAAC+ audio codec as specified in 3GPP TS 26.401 [19], 3GPP TS 26.402 [20], 3GPP TS 26.410 [24] and 3GPP TS 26.411 [25].
- *AMR-WB*+: All decoding requirements for the AMR-WB+ audio codec as specified in 3GPP TS 26.290 [26], 3GPP TS 26.304 [27] and 3GPP TS 26.273 [28].

5.3 Encoding Capabilities

The following speech media encoding capabilities are defined:

- AMR: The encoding requirements for the AMR speech codec as specified in 3GPP TS 26.401 [19], clause 7, as well as 3GPP TS 26.403 [21], 3GPP TS 26.404 [22] and 3GPP TS 26.405 [23].
- AMR-WB: The encoding requirements for the AMR-WB by one of the following methods:
 - according to 3GPP TS 26.173 [10]
 - according to 3GPP TS 26.204 [11];
 - the AMR-WB IO mode according to TS 26.442 [14] and TS 26.443 [15],
 - the AMR-WB IO mode according to TS 26.452 [34].
- EVS: The encoding requirements for the EVS codec by one of the following methods:
 - TS 26.442 [14] and TS 26.443 [15] encoding functions; or
 - TS 26.452 [34] encoding functions.

The following audio media encoding capabilities are defined:

- *eAAC*+: The encoding requirements for the AAC+ audio codec as specified 3GPP TS 26.401 [19], clause 7, as well as 3GPP TS 26.403 [21], 3GPP TS 26.404 [22] and 3GPP TS 26.405 [23].
- AMR-WB+: The encoding requirements for the AMR-WB+ audio codec by one of the following methods
 - according to 3GPP TS 26.273 [28]; or
 - according to 3GPP TS 26.304 [27].

6 Operation Points

6.1 Introduction

The speech and audio Operation Points defined in this clause are primarily introduced in order to be used as content format in the context of 5G Media Streaming, but not restricted to this use case.

An operation point is a combination of rendering formats and media decoding capabilities.

For each Operation Point, Bitstream and Receiver requirements are detailed in the remainder of clause 6.

Table 6.1 provides an overview of the Operation Points defined in the present document.

Table 6.1: Speech and Audio Operation Points

Operation Point name	Format Properties	Decoding and Encoding Capabilities	Reference
AMR speech	Sampling frequency: 8 kHz	AMR	6.2.2
AMR-WB speech	Sampling frequency: 16 kHz	AMR-WB	6.2.3
EVS mono	Sampling frequency: 8, 16, 32, 48 kHz	EVS	6.2.4
eAAC+ stereo	Sampling frequency: 32, 44.1, 48 kHz	eAAC+	6.3.2
AMR-WB+	Sampling frequency: 8, 16, 32, 48 kHz	AMR-WB+	6.3.3

6.2 Speech Operation Points

6.2.1 Introduction

This clause defines speech operation points. For each operation point, the requirements for the bitstream as well as for the receiver are defined.

6.2.2 AMR

6.2.2.1 Bitstream Encoding Requirements

The following requirements apply to the AMR Operation Point.

- The sampling frequency shall be 8 kHz.
- The bitstream shall be encoded according to either 3GPP TS 26.073 [5] or 3GPP TS 26.104 [6].

Note that the bitstream produced by the AMR encoder consists of 20ms encoded speech frames.

6.2.2.2 Receiver Requirements

Receivers conforming to the **AMR** Operation Point shall support the *AMR* speech media decoding capability according to clause 5.2 and shall support playback of the decoded signal.

6.2.2.3 Sender Requirements

Senders conforming to the **AMR** Operation Point shall support the *AMR* speech media encoding capability according to clause 5.3 in real-time for any speech source format with sampling frequency 8kHz.

6.2.3 AMR-WB

6.2.3.1 Bitstream Requirements

The following requirements apply to the AMR-WB Operation Point.

- The sampling frequency shall be 16 kHz.
- The bitstream shall be encoded by one of the following methods:
 - according to 3GPP TS 26.173 [10]
 - according to 3GPP TS 26.204 [11];
 - the AMR-WB IO mode according to TS 26.442 [14] and TS 26.443 [15],
 - the AMR-WB IO mode according to TS 26.452 [34].

Note that the bitstream produced by the AMR-WB encoder consists of 20 ms encoded speech frames.

6.2.3.2 Receiver Requirements

Receivers conforming to the **AMR-WB** Operation Point shall support the *AMR-WB* speech media decoding capability according to clause 5.2 and shall support playback of the decoded signal.

6.2.3.3 Sender Requirements

Senders conforming to the **AMR-WB** Operation Point shall support the *AMR-WB* speech media encoding capability according to clause 5.3 in real-time for any speech source format with sampling frequency 16kHz.

6.2.4 EVS

6.2.4.1 Bitstream Encoding Requirements

The following requirements apply to the ${\bf EVS}$ Operation Point:

- The sampling frequency shall be one of the following: 8, 16, 32, 48 kHz.
- The bitstream shall be encoded according to one of the following methods
 - TS 26.442 [14] and TS 26.443 [15] encoding functions; or
 - TS 26.452 [34] encoding functions.

Note that the bitstream produced by the EVS encoder consists of 20ms encoded speech frames.

6.2.4.2 Receiver Requirements

Receivers conforming to the **EVS** Operation Point shall support the *EVS* speech media decoding capability according to clause 5.2 and shall support playback of the decoded signal.

6.2.4.3 Sender Requirements

Senders conforming to the **EVS** Operation Point shall support the *EVS* speech media encoding capability according to clause 5.3 in real-time for any speech source format with sampling frequency 8, 16, 32, 48 kHz.

6.3 Audio Operation Points

6.3.1 Introduction

This clause defines audio operation points. For each operation point, the requirements for the bitstream as well as for the receiver are defined.

6.3.2 eAAC+ stereo

6.3.2.1 Bitstream Encoding Requirements

The following requirements apply to the eAAC+ stereo Operation Point.

- The sampling frequency shall be either 32 kHz, 44.1 kHz or 48 kHz.
- The bitstream shall be encoded according to 3GPP TS 26.401 [19], clause 7, as well as 3GPP TS 26.403 [21], 3GPP TS 26.404 [22] and 3GPP TS 26.405 [23].

NOTE: The specified eAAC+ encoder consists of AAC-LC with additional tools that can be enabled (SBR, PS and more), see [19].

6.3.2.2 Receiver Requirements

Receivers conforming to the **eAAC+ stereo** Operation Point shall support the *eAAC+* media decoding capability according to clause 5.2 and shall support playback of the decoded signal.

NOTE: The eAAC+ decoder supports decoding of streams encoded with AAC-LC or aacPlus, see [19].

6.3.2.3 Sender Requirements

Senders conforming to the **eAAC+ stereo** Operation Point shall support the *eAAC+ stereo* audio media encoding capability according to clause 5.3 in real-time for any stereo audio source format with sampling frequency 32kHz, 44.1kHz, 48kHz.

6.3.3 AMR-WB+

6.3.3.1 Bitstream Encoding Requirements

The following requirements apply to the AMR-WB+ Operation Point.

- The sampling frequency shall be either 8, 16, 32 or 48 kHz.
- The bitstream shall be encoded by one of the following methods
 - according to 3GPP TS 26.273 [28]; or
 - according to 3GPP TS 26.304 [27].

6.3.3.2 Receiver Requirements

Receivers conforming to the **AMR-WB+** Operation Point shall support the *AMR-WB+* media decoding capability according to clause 5.2 and shall support playback of the decoded signal.

6.3.3.3 Sender Requirements

Senders conforming to the **AMR-WB+** Operation Point shall support the *AMR-WB+* audio media encoding capability according to clause 5.3 in real-time for any stereo audio source format with sampling frequency 8, 16, 32 or 48 kHz.

7 Mapping to 5GMS delivery

7.1 Introduction

This clause defines the mapping of the Operation Points as defined in clause 6 to 5G Media Streaming delivery. In particular the following aspects are addressed:

- Encapsulation of a bitstream into an ISO BMFF track.
- Definition of media content and receivers conforming to media profile including:
 - Encapsulation of a bitstream into a CMAF track
 - Providing the content in a CMAF Switching Set
 - Mapping to DASH-based distribution
 - Playback Requirements for a receiver conforming to this media profile

7.2 AMR Media Profile

7.2.1 Mapping to ISO BMFF

If media is provided following the operation point **AMR** and is encapsulated in the ISO BMFF [37], then the file format track shall contain the sample entry AMRSampleEntry with box_type 'samr' as defined in TS 26.244 [29] clause 6.5 and conform to the requirements of the sample entry 'samr' as defined in TS 26.244 [29].

7.2.2 Media Profile Definition

7.2.2.1 CMAF Track Definition

If media is provided following the operation point **AMR** and is encapsulated in a CMAF track, then the CMAF track shall conform to clause 7.2.1, and conform to the general CMAF Track constraints in ISO/IEC 23000-19 [30] as well as the general audio track constraints defined in ISO/IEC 23000-19 [30], clause 10.

7.2.2.2 CMAF Switching Set and Media Profile Definition

If media is provided following the operation point **AMR** and is provided in a CMAF Switching Set, then every CMAF track in the CMAF Switching Set shall conform to clause 7.2.2.1, and shall conform to the general CMAF Switching Set constraints in ISO/IEC 23000-19 [30], clause 7. CMAF Switching Sets that follow these requirements conform to the CMAF AMR media profile 'camr' defined in this clause.

7.2.2.3 Mapping to DASH Adaptation Set

If media is provided following the operation point **AMR**, a CMAF Switching Set conforming to clause 7.2.2.2 may be provided in a DASH Media Presentation Description in an Adaptation Set, in that case, the Adaptation Set shall conform to the Adaptation Set constraints of the DASH profile for CMAF as defined in ISO/IEC 23009-1 [31]. The following parameters shall be present on Adaptation Set level and set:

- @codecs is set to 'samr'
- @mimeType is set to be compatible with "audio/mp4 profiles='camr'"
- @audioSamplingRate is set to '8000'

If the Adaptation Set conforms to the constraints for the **AMR** Operation Point as defined in this clause, then the @profiles parameter in the Adaptation Set may signal conformance to this Media Profile by using "urn:3GPP:audio:mp:amr".

7.2.2.4 Playback Requirements

For a receiver supporting the **AMR** media profile the following applies:

- It shall support the receiver requirements as documented in clause 6.2.2.2 for any CMAF Track conforming to the CMAF **AMR** media profile 'camr' as defined in clause 7.2.2.1.
- It shall support the following playback requirements as documented in clause 8 of CTA-WAVE 5003 [32] for any content conforming to a CMAF Switching Set according to CMAF AMR media profile 'camr' as defined in clause 7.2.2.2, namely:
 - 8.2 Sequential Track Playback
 - 8.3 Random Access to Fragment
 - 8.4 Random Access to Time
 - 8.5 Switching Set Playback
 - 8.6 Regular Playback of Chunked Content
 - 8.7 Regular Playback of Chunked Content, non-aligned append
- It should support the following playback requirements as documented in clause 8 of CTA-WAVE 5003 [29] for any content conforming to a CMAF Switching Set according to CMAF AMR media profile 'camr' as defined in clause 7.2.2.2, namely:
 - 8.9 Out-Of-Order Loading
 - 8.10 Overlapping Fragments
 - 8.12 Playback of Encrypted Content

7.2.2.5 Content Generation Requirements

For a transmitter supporting the AMR media profile the following applies:

- It shall support all media encoding capabilities for AMR as defined in clause 5.3.
- It shall support the sender requirements for AMR as defined in clause 6.2.2.3.
- It shall support the generation of a CMAF Track as defined in clause 7.2.2.1 that conforms to the CMAF Media Profile 'camr' as defined in clause 7.2.2.2.
- If used for Adaptive Bit Rate (ABR) distribution, it shall support the generation of a CMAF Switching Set as defined in clause 7.2.2.4.

7.3 AMR-WB Media Profile

7.3.1 Mapping to ISO BMFF

If media is provided following the operation point **AMR-WB** and is encapsulated in the ISO BMFF, then the file format track shall contain the AMRSampleEntryBox and shall conform to the requirements of the sample entry 'sawb' as defined in TS 26.244 [29] clause 6.5.

7.3.2 Media Profile Definition

7.3.2.1 CMAF Track Definition

If media is provided following the operation point **AMR-WB** and is encapsulated in a CMAF track, then the CMAF track shall conform to clause 7.3.1, and conform to the general CMAF Track constraints in ISO/IEC 23000-19, clause 7 as well as the general audio track constraints defined in ISO/IEC 23000-19 [30], clause 10.

7.3.2.2 CMAF Switching Set and Media Profile Definition

If media is provided following the operation point **AMR-WB** and is provided in a CMAF Switching Set, then every CMAF track in the CMAF Switching Set shall conform to clause 7.3.2.1, and shall conform to the general CMAF Switching Set constraints in ISO/IEC 23000-19 [30], clause 7. CMAF Switching Sets that follow these requirements conform to the CMAF AMR WB media profile 'camw' defined in this clause.

7.3.2.3 Mapping to DASH Adaptation Set

If media is provided following the operation point **AMR-WB**, a switching set conforming to clause 7.3.2.2 may be provided in a DASH Media Presentation Description in an Adaptation Set, in that case, the Adaptation Set shall conform to the Adaptation set constraints of the DASH profile for CMAF as defined in ISO/IEC 23009-1 [31] clause 8.12.4.3. The following parameters shall be present on Adaptation Set level:

- @codecs is set to 'sawb'
- @mimeType is set to be compatible with "audio/mp4 profiles='camw'"
- @audioSamplingRate is set to '16000'

If the Adaptation Set conforms to the constraints for the **AMR-WB** Operation Point as defined in this clause, then the @profiles parameter in the Adaptation Set may signal conformance to this Media Profile by using "urn:3GPP:audio:mp:amr-wb".

7.3.2.4 Playback Requirements

For a receiver supporting the AMR-WB media profile the following applies:

- It shall support the receiver requirements as documented in clause 6.2.3.2 for any CMAF Track conforming to the CMAF AMR-WB media profile 'camw' as defined in clause 7.3.2.2.
- It shall support the following playback requirements as documented in clause 8 of CTA-WAVE 5003 [32] for any content conforming to a CMAF Switching Set according to CMAF AMR-WB media profile 'camw' as defined in clause 7.3.2.2, namely:
 - 8.2 Sequential Track Playback
 - 8.3 Random Access to Fragment
 - 8.4 Random Access to Time
 - 8.5 Switching Set Playback
 - 8.6 Regular Playback of Chunked Content
 - 8.7 Regular Playback of Chunked Content, non-aligned append
- It should support the following playback requirements as documented in clause 8 of CTA-WAVE 5003 [32] for any content conforming to a CMAF Switching Set according to CMAF AMR-WB media profile 'camw' as defined in clause 7.3.2.2, namely:
 - 8.9 Out-Of-Order Loading
 - 8.10 Overlapping Fragments
 - 8.12 Playback of Encrypted Content

7.3.2.5 Content Generation Requirements

For a transmitter supporting the AMR-WB media profile the following applies:

- It shall support all media encoding capabilities for AMR-WB as defined in clause 5.3.
- It shall support the sender requirements for AMR-WB as defined in clause 6.2.3.3.

- It shall support the generation of a CMAF Track as defined in clause 7.3.2.1 that conforms to the CMAF Media Profile 'camw' as defined in clause 7.3.2.2.
- If used for Adaptive Bit Rate (ABR) distribution, it shall support the generation of a CMAF Switching Set as defined in clause 7.3.2.4.

7.4 EVS Media Profile

7.4.1 Mapping to ISO BMFF

If media is provided following the operation point **EVS** and is encapsulated in the ISO BMFF, then the file format track shall contain the EVSSampleEntryBox with box_type and conform to the requirements of the sample entry 'sevs' as defined in TS 26.244 [29].

7.4.2 Media Profile Definition

7.4.2.1 CMAF Track Definition

If media is provided following the operation point **EVS** and is encapsulated in a CMAF track, then the CMAF track shall conform to clause 7.3.1, and conform to the general CMAF Track constraints in ISO/IEC 23000-19 [30], clause 7 as well as the general audio track constraints defined in ISO/IEC 23000-19 [30], clause 10.

7.4.2.2 CMAF Switching Set and Media Profile Definition

If media is provided following the operation point **EVS** and is provided in a CMAF Switching Set, then every CMAF track in the CMAF Switching Set shall conform to clause 7.4.2.1, and shall conform to the general CMAF Switching Set constraints in ISO/IEC 23000-19 [30], clause 7. A CMAF Switching Set that follows these requirements conforms to the CMAF EVS media profile 'cevs' defined in this clause.

7.4.2.3 Mapping to DASH Adaptation Set

If media is provided following the operation point **EVS** and is provided in a DASH Media Presentation in an Adaptation Set, a switching set conforming to clause 7.4.2.2 may be provided in a DASH Media Presentation Description in an Adaptation Set. In that case, the Adaptation Set shall conform to the Adaptation set constraints of the DASH profile for CMAF as defined in ISO/IEC 23009-1 [31]. The following parameters shall be present on Adaptation Set level:

- @codecs is set to 'sevs'
- @mimeType is set to be compatible with "audio/mp4 profiles='cevs'"
- @audioSamplingRate is set to one of the following: '8000', '16000', '24000', '32000'

If the Adaptation Set conforms to the constraints for the **EVS** Operation Point as defined in this clause, then the @profiles parameter in the Adaptation Set may signal conformance to this Media Profile by using "urn:3GPP:audio:mp:evs.

7.4.2.4 Playback Requirements

For a receiver supporting the EVS media profile the following applies:

- It shall support the receiver requirements as documented in clause 6.2.4.2 for any CMAF Track conforming to the CMAF EVS media profile 'cevs' as defined in clause 7.4.2.2.
- It shall support the following playback requirements as documented in clause 8 of CTA-WAVE 5003 [32] for any content conforming to a CMAF Switching Set according to CMAF EVS media profile 'cevs' as defined in clause 7.4.2.2, namely:
 - 8.2 Sequential Track Playback

- 8.3 Random Access to Fragment
- 8.4 Random Access to Time
- 8.5 Switching Set Playback
- 8.6 Regular Playback of Chunked Content
- 8.7 Regular Playback of Chunked Content, non-aligned append
- It should support the following playback requirements as documented in clause 8 of CTA-WAVE 5003 [32] for any content conforming to a CMAF Switching Set according to CMAF EVS media profile 'cevs' as defined in clause 7.2.2.2, namely:
 - 8.9 Out-Of-Order Loading
 - 8.10 Overlapping Fragments
 - 8.12 Playback of Encrypted Content

7.4.2.5 Content Generation Requirements

For a transmitter supporting the EVS media profile the following applies:

- It shall support all media encoding capabilities for EVS as defined in clause 5.3.
- It shall support the sender requirements for EVS as defined in clause 6.2.4.3.
- It shall support the generation of a CMAF Track as defined in clause 7.4.2.1 that conforms to the CMAF Media Profile 'cevs' as defined in clause 7.4.2.2.
- If used for Adaptive Bit Rate (ABR) distribution, it shall support the generation of a CMAF Switching Set as defined in clause 7.4.2.2.

7.5 void

7.6 eAAC+ stereo Media Profile

7.6.2.1 CMAF Track Definition

If media is provided following the operation point **eAAC+ stereo** and is encapsulated in a CMAF track, then the CMAF track shall conform to and contain the MP4AudioSampleEntry 'mp4a' as defined in ISO/IEC 14496-14 [38], the general CMAF Track constraints in ISO/IEC 23000-19 [30], clause 7, the general audio track constraints defined in ISO/IEC 23000-19 [30], clause 10 as well as AAC core constraints in clause 10 of ISO/IEC 23000-19 [30].

7.6.2.2 CMAF Switching Set and Media Profile Definition

If media is provided following the operation point **eAAC+ stereo** and is provided in a CMAF Switching Set, then every CMAF track in the CMAF Switching Set shall conform to clause 7.6.2.1, and the tracks shall conform to the general CMAF Switching Set constraints in ISO/IEC 23000-19 [30], clause 7, the general CMAF audio track Switching Set constraints defined in ISO/IEC 23000-19 [30], clause 10 as well as the AAC core Switching Set constraints in clause 10 of ISO/IEC 23000-19 [30]. A CMAF Switching Set following these requirements is defined as the CMAF eAAC+ stereo media profile 'ceac'.

7.6.2.3 Mapping to DASH Adaptation Set

If media is provided following the operation point **eAAC+ stereo**, a switching set conforming to clause 7.6.2.2 may be provided in a DASH Media Presentation Description in an Adaptation Set. In that case, the Adaptation Set shall conform to the Adaptation set constraints of the DASH profile for CMAF as defined in ISO/IEC 23009-1 [31] clause 8.12.4.3. The following parameters shall be present on Adaptation Set level and set:

- @codecs is set to 'mp4a'
- @mimeType is set to be compatible with "audio/mp4 profiles='ceac'"
- @audioSamplingRate is set to '32000', '44100', or '48000'

If the Adaptation Set conforms to the constraints for the **eAAC+ stereo** Operation Point as defined in this clause, then the @profiles parameter in the Adaptation Set may signal conformance to this Media Profile by using "urn:3GPP:audio:mp:eAAC+".

7.6.2.4 Playback Requirements

For a receiver supporting the **eAAC+ stereo** media profile the following applies:

- It shall support the receiver requirements as documented in clause 6.3.1.2 for any CMAF Track conforming to the CMAF eAAC+ stereo media profile 'ceac' as defined in clause 7.6.2.2.
- It shall support the following playback requirements as documented in clause 8 of CTA-WAVE 5003 [32] for any content conforming to a CMAF Switching Set according to CMAF eAAC+ media profile 'ceac' as defined in clause 7.6.2.2, namely:
 - 8.2 Sequential Track Playback
 - 8.3 Random Access to Fragment
 - 8.4 Random Access to Time
 - 8.5 Switching Set Playback
 - 8.6 Regular Playback of Chunked Content
 - 8.7 Regular Playback of Chunked Content, non-aligned append
- It should support the following playback requirements as documented in clause 8 of CTA-WAVE 5003 [32] for any content conforming to a CMAF Switching Set according to CMAF AMR-WB media profile 'ceac' as defined in clause 7.6.2.2, namely:
 - 8.9 Out-Of-Order Loading
 - 8.10 Overlapping Fragments
 - 8.12 Playback of Encrypted Content

7.6.2.5 Content Generation Requirements

For a transmitter supporting the eAAC+ stereo media profile the following applies:

- It shall support all media encoding capabilities for eAAC+ stereo as defined in clause 5.3.
- It shall support the sender requirements for eAAC+ stereo as defined in clause 6.6.2.3.
- It shall support the generation of a CMAF Track as defined in clause 7.6.2.1 that conforms to the CMAF Media Profile 'ceac' as defined in clause 7.6.2.2.
- If used for Adaptive Bit Rate (ABR) distribution, it shall support the generation of a CMAF Switching Set as defined in clause 7.6.2.2.

7.7 AMR-WB+ Media Profiles

7.7.1 Mapping to ISO BMFF

If media is provided following the operation point **AMR-WB**+ and is encapsulated in the ISO BMFF, then the file format track shall contain the AMRWPSampleEntry with box_type and conform to the requirements of the sample entry 'sawp' as defined in TS 26.244 [29] clause 6.9.

7.7.2 Media Profile Definition

7.7.2.1 CMAF Track Definition

If media is provided following the operation point **AMR-WB**+ and is encapsulated in a CMAF track, then the CMAF track shall conform to clause 7.7.1 and the general CMAF Track constraints in ISO/IEC 23000-19 [30], clause 7 as well as the general audio track constraints defined in ISO/IEC 23000-19, clause 10.

7.7.2.2 CMAF Switching Set and Media Profile Definition

If media is provided following the operation point **AMR-WB**+ and is provided in a CMAF Switching Set, then every CMAF track in the CMAF Switching Set shall conform to the requirements of clause 7.7.2.1 and the general CMAF Switching Set constraints in ISO/IEC 23000-19 [30], clause 7. A CMAF Switching Set following these requirements is defined as the CMAF AMR-WB+ media profile 'camp'.

7.7.2.3 Mapping to DASH Adaptation Set

If media is provided following the operation point **AMR-WB**+, a switching set conforming to clause 7.7.2.2 may be provided in a DASH Media Presentation in an Adaptation Set. In that case, the Adaptation Set shall conform to the Adaptation set constraints of the DASH profile for CMAF as defined in ISO/IEC 23009-1 [31] clause 8.12.4.3. The following parameters shall be present on Adaptation Set level:

- @codecs is set to 'sawp'
- @mimeType is set to be compatible with "audio/mp4 profiles='camp'"
- @audioSamplingRate is set to any of the following values: '8000', '16000', '32000', or '38400'

If the Adaptation Set conforms to the constraints for the **AMR-WB+** Operation Point as defined in this clause, then the @profiles parameter in the Adaptation Set may signal conformance to this Media Profile by using "urn:3GPP:audio:mp:amr-wb+".

7.7.2.4 Playback Requirements

For a receiver supporting the AMR-WB+ media profile the following applies:

- It shall support the receiver requirements as documented in clause 6.3.2.2 for any CMAF Track conforming to the CMAF AMR-WB+ stereo media profile 'camp' as defined in clause 7.7.2.2.
- It shall support the following playback requirements as documented in clause 8 of CTA-WAVE 5003 [32] for any content conforming to a CMAF Switching Set according to CMAF AMR-WB media profile 'camp' as defined in clause 7.7.2.2, namely:
 - 8.2 Sequential Track Playback
 - 8.3 Random Access to Fragment
 - 8.4 Random Access to Time
 - 8.5 Switching Set Playback
 - 8.6 Regular Playback of Chunked Content

- 8.7 Regular Playback of Chunked Content, non-aligned append
- It should support the following playback requirements as documented in clause 8 of CTA-WAVE 5003 [32] for any content conforming to a CMAF Switching Set according to CMAF AMR-WB media profile 'camp' as defined in clause 7.7.2.2, namely:
 - 8.9 Out-Of-Order Loading
 - 8.10 Overlapping Fragments
 - 8.12 Playback of Encrypted Content

7.7.2.5 Content Generation Requirements

For a transmitter supporting the AMR-WB+ media profile the following applies:

- It shall support all media encoding capabilities for AMR-WB+ stereo as defined in clause 5.3.
- It shall support the sender requirements for AMR-WB+ as defined in clause 6.3.3.3.
- It shall support the generation of a CMAF Track as defined in clause 7.7.2.1 that conforms to the CMAF Media Profile 'cawp' as defined in clause 7.7.2.2.
- If used for Adaptive Bit Rate (ABR) distribution, it shall support the generation of a CMAF Switching Set as defined in clause 7.7.2.2.

Annex A (informative): Registration Information

A.1 3GPP Registered URIs

The clause documents the registered URIs in the present document following the process in http://www.3gpp.org/specifications-groups/34-uniform-resource-name-urn-list

Table A-1 lists all registered URN values as well as:

- a brief description of its functionality;
- a reference to the specification or other publicly available document (if any) containing the definition;
- the name and email address of the person making the application; and
- any supplementary information considered necessary to support the application.

Table A-1: 3GPP Registered URNs

URN	Description	Reference	Contact	Remarks
urn:3GPP:audio:mp:amr	AMR Media Profile	TS 26.117,	Thomas Stockhammer	none
		clause 7.2.2.3	tsto@qti.qualcomm.com	
urn:3GPP:audio:mp:amr-wb	AMR-WB Media Profile	TS 26.117,	Thomas Stockhammer	none
		clause 7.3.2.3	tsto@qti.qualcomm.com	
urn:3GPP:audio:mp:evs	EVS Media Profile	TS 26.117,	Thomas Stockhammer	none
		clause 7.4.2.3	tsto@qti.qualcomm.com	
urn:3GPP:audio:mp:eAAC+	eAAC+ stereo Media	TS 26.117,	Thomas Stockhammer	none
	Profile	clause 7.6.2.3	tsto@qti.qualcomm.com	
urn:3GPP:audio:mp:amr-	AMR-WB+ Media Profile	TS 26.117,	Thomas Stockhammer	none
+dw		clause 7.7.2.3	tsto@qti.qualcomm.com	

Annex B (informative): Change history

Change history							
Date	Meeting	TDoc	CR	Rev	Cat	Subject/Comment	New version
2019-12	SA#86	SP-190990				Presented to TSG SA#86 (for information)	1.0.0
2020-01	SA4#107	S4-200245				Version agreed by TSG SA WG 4 to be sent for SA approval	1.1.0
2020-03	SA#87-e	SP-200049				Presented to TSG SA#87-e (for approval)	2.0.0
2020-03	SA#87-e	SP-200049				Approved by TSG SA#87-e	16.0.0
2022-04	-	-	-	-	-	Update to Rel-17 version (MCC)	17.0.0
2023-06	SA#100	SP-230546	000 4	-	Α	Corrections to references	17.1.0
2024-03	SA#103	SP-240461	000 7	4	Α	[5GMS3] Correction on Audio encapsulation in ISO BMFF, 3GP file format, CMAF and DASH	17.2.0

History

Document history						
V17.0.0	0.0 May 2022 Publication					
V17.1.0	July 2023	Publication				
V17.2.0	April 2024	Publication				