ETSI TS 102 563 V2.1.1 (2017-01)

TECHNICAL SPECIFICATION

Digital Audio Broadcasting (DAB);
DAB+ audio coding (MPEG HE-AACv2)

European Broadcasting Union
Union Européenne de Radio-Télévision
EBU·UER

Reference
RTS/JTC-DAB-82

Keywords
audio, broadcasting, coding, DAB, digital

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

The present document can be downloaded from:
http://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 only prevailing document is the print of the Portable Document Format (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
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

Contents

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 (https://ipr.etsi.org/).

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 Joint Technical Committee (JTC) Broadcast of the European Broadcasting Union (EBU), Comité Européen de Normalisation ELECtrotechnique (CENELEC) and the European Telecommunications Standards Institute (ETSI).

NOTE 1: The EBU/ETSI JTC Broadcast was established in 1990 to co-ordinate the drafting of standards in the specific field of broadcasting and related fields. Since 1995 the JTC Broadcast became a tripartite body by including in the Memorandum of Understanding also CENELEC, which is responsible for the standardization of radio and television receivers. The EBU is a professional association of broadcasting organizations whose work includes the co-ordination of its members' activities in the technical, legal, programme-making and programme-exchange domains. The EBU has active members in about 60 countries in the European broadcasting area; its headquarters is in Geneva.

European Broadcasting Union
CH-1218 GRAND SACONNEX (Geneva)
Switzerland
Tel: +41 22 717 21 11
Fax: +41 22 717 24 81

The Eureka Project 147 was established in 1987, with funding from the European Commission, to develop a system for the broadcasting of audio and data to fixed, portable or mobile receivers. Their work resulted in the publication of European Standard, ETSI EN 300 401 [1], for DAB (see note 2) which now has worldwide acceptance.

NOTE 2: DAB is a registered trademark owned by one of the Eureka Project 147 partners.

The DAB family of standards is supported by World DAB, an organization with members drawn from broadcasting organizations and telecommunication providers together with companies from the professional and consumer electronics industry.

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 ETSI Drafting Rules (Verbal forms for the expression of provisions).

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

1 Scope

The present document defines the method to code and transmit audio services using the HE-AACv2 [2] audio coder for Eureka-147 Digital Audio Broadcasting (DAB) (ETSI EN 300 401 [1]) and details the necessary mandatory requirements for decoders. The permitted audio modes and the data protection and encapsulation are detailed. This audio coding scheme permits the full use of the PAD channel for carrying dynamic labels and user applications.

2 References

2.1 Normative references

References are either specific (identified by date of publication and/or edition number or version number) or non-specific. For specific references, only the cited version applies. For non-specific references, the latest version of the referenced document (including any amendments) applies.

Referenced documents which are not found to be publicly available in the expected location might be found at https://docbox.etsi.org/Reference.

NOTE: While any hyperlinks included in this clause were valid at the time of publication, ETSI cannot guarantee their long term validity.

The following referenced documents are necessary for the application of the present document.

[1] ETSI EN 300 401 (V2.1.1): "Radio Broadcasting Systems; Digital Audio Broadcasting (DAB) to mobile, portable and fixed receivers".

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

2.2 Informative references

References are either specific (identified by date of publication and/or edition number or version number) or non-specific. For specific references, only the cited version applies. For non-specific references, the latest version of the referenced document (including any amendments) applies.

NOTE: While any hyperlinks included in this clause were valid at the time of publication, ETSI cannot guarantee their long term validity.

The following referenced documents are not necessary for the application of the present document but they assist the user with regard to a particular subject area.

[i.1] ISO/IEC 23003-1: "Information technology - MPEG audio technologies - Part 1: MPEG Surround".

3 Definitions, abbreviations and arithmetic operators

3.1 Definitions

For the purposes of the present document, the terms and definitions given in ETSI EN 300 401 [1] and the following apply:

access unit: access unit contains the audio samples for 20 ms, 30 ms, 40 ms or 60 ms of audio depending on the sampling rate of the AAC core, respectively 48 kHz, 32 kHz, 24 kHz or 16 kHz

audio super frame: audio super frame contains a number of AUs which together contain the encoded audio for 120 ms

subchannel_index: subchannel_index is derived from the size of the sub-channel carrying the audio service and defines the number of Reed-Solomon code words in each audio super frame

3.2 Abbreviations

For the purposes of the present document, the abbreviations given in ETSI EN 300 401 [1] and the following apply:

AAC	Advanced Audio Coding
AU	Access Unit
CRC	Cyclic Redundancy Check
CU	Capacity Unit
DAB	Digital Audio Broadcasting
DAC	Digital Analogue Converter
DMB	Digital Multimedia Broadcasting
DRC	Dynamic Range Control
DVB	Digital Video Broadcasting
FIC	Fast Information Channel
FIG	Fast Information Group
GF	Galois Field
HE-AAC	High Efficiency AAC
MCI	Multiplex Configuration Information
MPEG	Moving Pictures Experts Group
MSC	Main Service Channel
PAD	Programme Associated Data
PS	Parametric Stereo
RS	Reed-Solomon
SAC	Spatial Audio Coding
SBR	Spectral Band Replication

3.3 Arithmetic operators

+	addition
-	subtraction
$\times$	multiplication
$\div$	division
m DIV p	denotes the quotient part of the division of m by p (m and p are positive integers)
m MOD p	denotes the remainder of the division of m by p (m and p are positive integers)
$\sum_{i=p}^{q} f(i)$	denotes the sum: $f(p) + f(p + 1) + f(p + 2) \ldots + f(q)$
$\prod_{i=p}^{q} f(i)$	denotes the product: $f(p) \times f(p + 1) \times f(p + 2) \ldots \times f(q)$

4 Introduction

The DAB system standard [1] allows audio (programme) services to be carried using either DAB audio or DAB+ audio. The present document defines the way that audio (programme) services are carried when using DAB+ audio (MPEG 4 HE-AACv2 audio coding).

MPEG 4 HE-AACv2 specifies two transforms, but only the 960 transform is permitted for DAB+ audio, with core sampling rates of 48 kHz, 32 kHz, 24 kHz and 16 kHz. Each AU (audio frame) contains samples for 20 ms, 30 ms, 40 ms or 60 ms respectively. AUs are built into audio super frames of 120 ms duration which are then carried in five DAB logical frames. In order to provide additional error control, Reed Solomon coding and virtual interleaving is applied. The overall scheme is shown in figure 1.

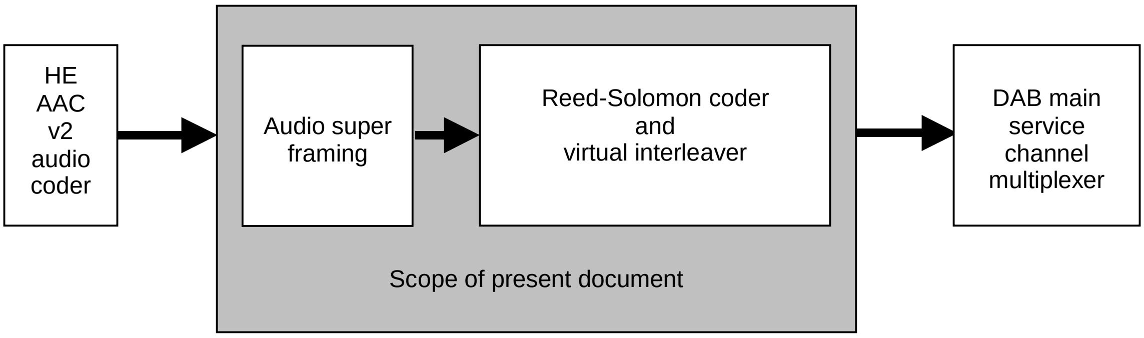

Figure 1: Conceptual diagram of the outer coder and interleaver

5 Audio

5.1 HE-AACv2 audio coding

For generic audio coding, a subset of the MPEG-4 High Efficiency Advanced Audio Coding v2 (HE-AACv2) profile chosen to best suit the DAB system environment is used. The HE-AACv2 Profile, Level 2 according to [2] shall apply with the following additional restrictions for the DAB system:

- Sampling rates: permitted output sampling rates of the HE-AACv2 decoder are 32 kHz and 48 kHz, i.e. when SBR is enabled the AAC core shall be operated at 16 kHz or 24 kHz, respectively. If SBR is disabled then the AAC core shall be operated at 32 kHz or 48 kHz respectively.
- Transform length: the number of samples per channel per AU is 960. This is required to harmonize HE-AAC AU lengths to allow the combination of an integer number of AUs to build an audio super frame of 120 ms duration.
- Audio bit rates are restricted to fit within a maximum sub-channel size of 192 kbps (approximately 175 kbps for audio, assuming no PAD).
- Audio super framing: AUs are composed into audio super frames, which always correspond to 120 ms in time. The AUs in the audio super frames are encoded together such that each audio super frame is of constant length, i.e. that bit exchange between AUs is only possible within an audio super frame. The number of AUs per super frame are: two (16 kHz AAC core sampling rate with SBR enabled), three (24 kHz AAC core sampling rate with SBR enabled), four (32 kHz AAC core sampling rate) or six (48 kHz AAC core sampling rate).

Each audio super frame is carried in five consecutive logical DAB frames (see clause 7) which enables simple synchronization and management of reconfigurations. The size of the audio super frame is defined by the size of the MSC sub-channel (see ETSI EN 300 401 [1], clause 6.2.1) which carries the audio super frame. Sub-channels are multiples of 8 kbps in size. The size of the audio super frame in bytes is given by the expressions below:

$$\text{subchannel\_index} = \text{MSC sub-channel size (kbps)} \div 8$$

$$\text{audio\_super\_frame\_size (bytes)} = \text{subchannel\_index} \times 110$$

The first byte of the audio super frame is byte 0 and the last byte is byte (audio_super_frame_size - 1).

NOTE: The subchannel_index parameter may take the values 1 to 24 due to the restriction limiting the maximum sub-channel size to 192 kbps.

5.2 Audio super framing syntax

Table 1: Syntax of he_aac_super_frame()

Syntax	No. of bits	Note

he_aac_super_frame(subchannel_index)
{
    he_aac_super_frame_header()
    for (n = 0; n < num_aus; n++) {
        au[n]
        au_crc[n]
    }
}
	
	8 × au_size[n]
	16
		
		determines num_aus

NOTE: au corresponds to one single access unit. Each au is protected by one CRC. The size of he_aac_super_frame() is equal to audio_super_frame_size.		

he_aac_super_frame_header()

The header contains the audio parameters for the audio super frame and the respective start positions of each AU within the audio super frame, along with an error protection word. The au_start values for the second and subsequent AUs are stored consecutively in the header. Depending on the number of AUs, 4 padding bits are added to achieve byte-alignment.

num_aus

The number of AUs in the audio super frame is determined by the settings of the audio parameters. num_aus may take the values 2, 3, 4 or 6 (see table 2).

au[n]

The AU contains the audio samples for 20 ms, 30 ms, 40 ms or 60 ms of audio depending on the core sampling rate, respectively 48 kHz, 32 kHz, 24 kHz or 16 kHz.

au_size[n]

This is the size in bytes of the AU.

au_crc[n]

Each AU is protected by a 16-bit CRC.

The CRC shall be generated according to the procedure defined in ETSI EN 300 401 [1], annex E. The generation shall be based on the polynomial:

$$G(x) = x^{16} + x^{12} + x^5 + 1$$

The CRC word shall be complemented (1s complement) prior to transmission. At the beginning of each CRC word calculation, all register stages shall be initialized to "1".

Table 2: Syntax of he_aac_super_frame_header()

Syntax	No. of bits	Note
he_aac_super_frame_header()		
{		
header_firecode	**16**	
// start of audio parameters		
rfa	**1**	
dac_rate	**1**	
sbr_flag	**1**	
aac_channel_mode	**1**	
ps_flag	**1**	
mpeg_surround_config	**3**	
// end of audio parameters		
if ((dac_rate == 0) && (sbr_flag == 1)) num_aus = 2;		AAC core sampling rate 16 kHz
if ((dac_rate == 1) && (sbr_flag == 1)) num_aus = 3;		AAC core sampling rate 24 kHz
if ((dac_rate == 0) && (sbr_flag == 0)) num_aus = 4;		AAC core sampling rate 32 kHz
if ((dac_rate == 1) && (sbr_flag == 0)) num_aus = 6;		AAC core sampling rate 48 kHz
for (n = 1; n < num_aus; n++) {		
au_start[n];	**12**	AU start position
}		
if !((dac_rate == 1) && (sbr_flag == 1))		
alignment	**4**	byte-alignment
}		
NOTE: The au_start for the first AU in the audio super frame (au_start[0]) is not transmitted. The first AU always starts immediately after the he_aac_super_frame_header().		

header_firecode

The header_firecode is a 16-bit field containing a Fire code capable of detecting and correcting most single error burst of up to 6 bits. The error pattern 101111 (where "1" indicates "bit error", "0" indicates "no bit error") can be detected but not corrected. The Fire code shall be generated using the polynomial:

$$G(x) = (x^{11} + 1)(x^5 + x^3 + x^2 + x + 1) = x^{16} + x^{14} + x^{13} + x^{12} + x^{11} + x^5 + x^3 + x^2 + x + 1$$

The Fire code word shall be calculated over the nine bytes from byte 2 to byte 10 of the audio super frame.

NOTE 1: Except in the case where num_aus = 6, the Fire code calculation will include some bytes from the first AU.

At the beginning of each Fire code word calculation, all register stages shall be initialized to "0".

NOTE 2: Other definitions call the above code an error-correcting CRC code.

audio parameters

The audio parameters comprise the rfa, dac_rate, sbr_flag, aac_channel_mode, ps_flag and mpeg_surround_config fields.

NOTE 3: When the audio parameters are changed, some interruption to the audio output should be expected. Broadcasters should therefore plan audio parameter changes carefully.

rfa

The rfa is a 1-bit field reserved for future addition. This bit shall be set to zero for the currently specified application.