ETSI TR 185 013 V3.1.1 (2009-05)

Technical Report

Telecommunications and Internet converged Services and Protocols for Advanced Networking (TISPAN); Codecs for customer network devices

Reference
DTR/TISPAN-05022-NGN-R3

Keywords
codec

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

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 (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 Report (TR) has been produced by ETSI Technical Committee Telecommunications and Internet converged Services and Protocols for Advanced Networking (TISPAN).

1 Scope

The objective of the present document is to define the set of codecs supported by CNDs for NGN services, on the basis of the service requirements contained in TS 181 005 [i.1] and TS 185 005 [i.4], and with reference to the CNDs architecture defined in TS 185 006 [i.11] (for communication services) and TS 185 009 [i.5] (for IPTV services).

Work will address AV/broadcast codecs for IPTV and Conversational codecs for telephony services. It will be fully consistent with the set of codecs identified in TS 181 005 [i.1]. When TS 181 005 [i.1] already specifies a list of codecs for voice services, no new codecs will be added, no codec will be removed nor downgraded in status (mandatory, recommended optional) within this list.

For each service (mainly person to person communication and IPTV) work will be focused on defining subset of default/mandatory which is the only way to guarantee interoperability and avoid transcoding a minimum set of codecs to be supported will be defined.

2 References

References are either specific (identified by date of publication and/or edition number or version number) or non-specific.

- For a specific reference, subsequent revisions do not apply.
- Non-specific reference may be made only to a complete document or a part thereof and only in the following cases:
  - if it is accepted that it will be possible to use all future changes of the referenced document for the purposes of the referring document;
  - for informative references.

Referenced documents which are not found to be publicly available in the expected location might be found at http://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.

2.1 Normative references

The following referenced documents are indispensable for the application of the present document. For dated references, only the edition cited applies. For non-specific references, the latest edition of the referenced document (including any amendments) applies.

Not applicable.

2.2 Informative references

The following referenced documents are not essential to the use of the present document but they assist the user with regard to a particular subject area**.** For non-specific references, the latest version of the referenced document (including any amendments) applies.

[i.1] ETSI TS 181 005: "Telecommunications and Internet Converged Services and Protocols for Advanced Networking (TISPAN); Service and Capability Requirements".

[i.2] ETSI TS 181 014: "Telecommunications and Internet converged Services and Protocols for Advanced Networking (TISPAN); Requirements for network transport capabilities to support IPTV services".

[i.3] ETSI TS 181 016: "Telecommunications and Internet converged Services and Protocols for Advanced Networking (TISPAN); Service Layer Requirements to integrate NGN Services and IPTV".

[i.4] ETSI TS 185 005: "Telecommunications and Internet converged Services and Protocols for Advanced Networking (TISPAN); Services requirements and capabilities for customer networks connected to TISPAN NGN".

[i.5] ETSI TS 185 009: "Telecommunications and Internet converged Services and Protocols for Advanced Networking (TISPAN) Architecture & reference points of a customer network device for IMS based IPTV services".

[i.6] ETSI TS 126 114: "Universal Mobile Telecommunications System (UMTS); LTE; IP Multimedia Subsystem (IMS); Multimedia telephony; Media handling and interaction (3GPP TS 26.114 version 8.2.1 Release 8)".

[i.7] ETSI TS 126 235: "Digital cellular telecommunications system (Phase 2+); Universal Mobile Telecommunications System (UMTS); LTE; Packet switched conversational multimedia applications; Default codecs (3GPP TS 26.235 Release 8)".

[i.8] ETSI TS 126 236: "Universal Mobile Telecommunications System (UMTS); LTE; Packet switched conversational multimedia applications; Transport protocols (3GPP TS 26.236 Release 8)".

[i.9] ETSI TS 102 005: "Digital Video Broadcasting (DVB); Specification for the use of Video and Audio Coding in DVB services delivered directly over IP protocols".

[i.10] ETSI TS 101 154: "Digital Video Broadcasting (DVB); Specification for the use of Video and Audio Coding in Broadcasting Applications based on the MPEG-2 Transport Stream".

[i.11] ETSI TS 185 006: "Telecommunications and Internet converged Services and Protocols for Advanced Networking (TISPAN); Customer Devices architecture and Reference Points".

[i.12] ITU-T Recommendation G.711: "Pulse Code Modulation (PCM) of voice frequencies".

[i.13] ITU-T Recommendation G.722: "7 kHz audio-coding within 64 kbit/s".

[i.14] ITU-T Recommendation G.191: "Software tools for speech and audio coding standardization".

[i.15] ITU-T Recommendation G.722 Appendix II: "Digital test sequences for the verification of the G.722 64 kbit/s SB-ADPCM 7 kHz codec".

[i.16] ETSI TS 102 527-1: "Digital Enhanced Cordless Telecommunications (DECT); New Generation DECT; Part 1: Wideband speech".

[i.17] ETSI TR 102 570: "Digital Enhanced Cordless Telecommunications (DECT); New Generation DECT; Overview and Requirements".

[i.18] IETF RFC 3551: "RTP Profile for Audio and Video Conferences with Minimal Control".

[i.19] ITU-T Recommendation G.722 Appendix III: "A high quality packet loss concealment algorithm for G.722".

[i.20] ITU-T Recommendation G.722 Appendix IV: "A low-complexity algorithm for packet loss concealment with G.722".

[i.21] ITU-T Recommendation G.729.1: "G.729-based embedded variable bit-rate coder: An 8-32 kbit/s scalable wideband coder bitstream interoperable with G.729".

[i.22] ITU-T Recommendation G.729.1 Amendment 3: "Extension of the G.729.1 low delay mode functionality to 14 kbit/s, and corrections to the main body and annex B".

[i.23] IETF RFC 4749: RTP Payload Format for the G.729.1 Audio Codec.

[i.24] ETSI TS 126 190: "Digital cellular telecommunications system (Phase 2+); Universal Mobile Telecommunications System (UMTS); LTE; Speech codec speech processing functions; Adaptive Multi-Rate - Wideband (AMR-WB) speech codec; Transcoding functions (3GPP TS 26.190 Release 8)".

[i.25] ITU-T Recommendation G.722.2: "Wideband coding of speech at around 16 kbit/s using Adaptive Multi-Rate Wideband (AMR-WB)".

[i.26] 3GPP2 C.S0014-B v1.0: "Enhanced Variable Rate Codec, Speech Service Option 3 and 68 for Wideband Spread Spectrum Digital Systems".

[i.27] VoIP Codecs and Protocols, ftp: //ftp.3gpp2.org/TSGC/Working/2007/2007-05-SanDiego/TSG-C-2007-05-SanDiego/WG1/SWG12/C12-20070514-012AR1--Proposed-SC-C.S0085-0-VoIP-Spec.doc.

[i.28] 3GPP2, C11-20061204-005-Proposed V&V -Ballot-Text-EVRC-Release-C-Specification.zip.

[i.29] ISO/IEC 14496-3 (2007): "Information technology - Coding of audio-visual objects - Part 3: Audio, including Amd1 and Amd2".

[i.30] ETSI TS 126 290 (V7.0.0): "Digital cellular telecommunications system (Phase 2+); Universal Mobile Telecommunications System (UMTS); Audio codec processing functions; Extended Adaptive Multi-Rate - Wideband (AMR-WB+) codec; Transcoding functions (3GPP TS 26.290 version 7.0.0 Release 7)".

[i.31] ETSI TS 102 366: "Digital Audio Compression (AC-3, Enhanced AC-3) Standard".

[i.32] ETSI TS 126 304: "Digital cellular telecommunications system (Phase 2+); Universal Mobile Telecommunications System (UMTS); LTE; Extended Adaptive Multi-Rate - Wideband (AMR-WB+) codec; Floating-point ANSI-C code (3GPP TS 26.304 version 8.0.0 Release 8)".

[i.33] ETSI TS 126 273: "Digital cellular telecommunications system (Phase 2+); Universal Mobile Telecommunications System (UMTS); LTE; ANSI-C code for the fixed-point Extended Adaptive Multi-Rate - Wideband (AMR-WB+) speech codec (3GPP TS 26.273 version 8.0.0 Release 8)".

[i.34] IETF RFC 4352: "RTP Payload Format for the Extended Adaptive Multi-Rate Wideband (AMR-WB+) Audio Codec".

[i.35] ETSI TR 126 936: "Digital cellular telecommunications system (Phase 2+); Universal Mobile Telecommunications System (UMTS); LTE; Performance characterization of 3GPP audio codecs (3GPP TR 26.936 version 8.0.0 Release 8)".

[i.36] IETF RFC 4184: "RTP Payload Format for AC-3 Audio".

[i.37] IETF RFC 3690: "RTP Payload Format for Transport of MPEG-4 Elementary Streams".

[i.38] ETSI TS 126 401: "Digital cellular telecommunications system (Phase 2+); Universal Mobile Telecommunications System (UMTS); LTE; General audio codec audio processing functions; Enhanced aacPlus general audio codec; General description (3GPP TS 26.401 version 8.0.0 Release 8)".

[i.39] IETF RFC 4867: "RTP Payload Format and File Storage Format for the Adaptive Multi-Rate (AMR) and Adaptive Multi-Rate Wideband (AMR-WB) Audio Codecs".

[i.40] IETF RFC 4788: "RTP payload format for Enhanced Variable Rate Wideband Codec (EVRC-WB) and media subtype updates for EVRC-B codec".

[i.41] ITU-T Recommendation H.263 (2005): "Video coding for low bit rate communication".

[i.42] ITU-T Recommendation H.264 | ISO/IEC 14496-10: "Advanced video coding for generic audiovisual services".

[i.43] SMPTE 421M Television - VC-1 Compressed Video Bitstream Format and Decoding Process.

[i.44] ITU-T Recommendation G.729: "Coding of speech at 8 kbit/s using conjugate-structure algebraic-code-excited linear prediction (CS-ACELP)".

[i.45] ISO/IEC 13818-2: "Information technology -- Generic coding of moving pictures and associated audio information".

[i.46] ETSI TS 126 171: "Digital cellular telecommunications system (Phase 2+); Universal Mobile Telecommunications System (UMTS); LTE; Speech codec speech processing functions; Adaptive Multi-Rate - Wideband (AMR-WB) speech codec; General description (3GPP TS 26.171 Release 8)".

[i.47] ETSI TS 126 173: "Digital cellular telecommunications system (Phase 2+); Universal Mobile Telecommunications System (UMTS); LTE; ANSI-C code for the Adaptive Multi-Rate - Wideband (AMR-WB) speech codec (3GPP TS 26.173 Release 8)".

[i.48] ETSI TS 126 204: "Digital cellular telecommunications system (Phase 2+); Universal Mobile Telecommunications System (UMTS); LTE; Speech codec speech processing functions; Adaptive Multi-Rate - Wideband (AMR-WB) speech codec; ANSI-C code (3GPP TS 26.204 Release 8)".

[i.49] ETSI TS 126 194: "Digital cellular telecommunications system (Phase 2+); Universal Mobile Telecommunications System (UMTS); LTE; Speech codec speech processing functions; Adaptive Multi-Rate - Wideband (AMR-WB) speech codec; Voice Activity Detector (VAD) (3GPP TS 26.194 Release 8)".

[i.50] ETSI TS 126 192: "Digital cellular telecommunications system (Phase 2+); Universal Mobile Telecommunications System (UMTS); LTE; Speech codec speech processing functions; Adaptive Multi-Rate - Wideband (AMR-WB) speech codec; Comfort noise aspects (3GPP TS 26.192 Release 8)".

[i.51] ETSI TS 126 191: "Digital cellular telecommunications system (Phase 2+); Universal Mobile Telecommunications System (UMTS); LTE; Speech codec speech processing functions; Adaptive Multi-Rate - Wideband (AMR-WB) speech codec; Error concealment of erroneous or lost frames (3GPP TS 26.191 Release 8)".

[i.52] ISO/IEC 14496-10 (2008): "Information technology -- Coding of audio-visual objects -- Part 10: Advanced Video Coding".

[i.53] ITU-T Recommendation H.320:"Narrow-band visual telephone systems and terminal equipment".

[i.54] ITU-T Recommendation H.321:"Adaptation of H.320 visual telephone terminals to B-ISDN environments".

[i.55] ITU-T Recommendation H.322:"Visual telephone systems and terminal equipment for local area networks which provide a guaranteed quality of service".

[i.56] ITU-T Recommendation H.323:"Packet-based multimedia communications systems".

[i.57] ITU-T Recommendation G.726:"40, 32, 24, 16 kbit/s Adaptive Differential Pulse Code Modulation (ADPCM)".

[i.58] ITU-T Recommendation H.245: "Control protocol for multimedia communication".

3 Definitions, symbols and abbreviations

3.1 Definitions

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

Customer Network Device (CND): device enabling the final user to have a direct access to services through a specific user interface as defined in TS185006 [i.11]

NOTE 1: A dual-mode (fixed+mobile) CND is a fixed CND when connected to a fixed access and is a mobile CND when connected to a mobile access.

NOTE 2: In case the CND consists of more than one part, the codec function may be built in any part of the CND.

fixed CND: CND connected to the TISPAN NGN network either via a corded interface or a fixed-wireless interface (Wi-Fi, Bluetooth or DECT/DECT-NG)

fixed narrowband CND: fixed CND supporting narrowband speech

fixed wideband CND: fixed CND supporting wideband speech

mobile CND: CND connected to a mobile network such as a 3GPP or 3GPP2 network via the mobile interface

mobile narrowband CND: mobile CND supporting narrowband speech

mobile wideband CND: mobile CND supporting wideband speech

NOTE: Wideband CNDs are also required to support narrowband speech.

3.2 Abbreviations

For the purposes of the present document, the following abbreviations apply:

AAC	Advanced Audio Coding
ACELP	Algebraic Code-Excited Linear Prediction
ADPCM	Adaptive Differential Pulse-Code Modulation
AMR	Adaptive Multi-Rate
AMR-WB	Adaptive Multi-Rate - Wide Band
AVC	Advanced Video Coding
CELP	Code Excited Linear Prediction
CND	Customer Network Device
CPN	Customer Premises Network
DECT	Digital Enhanced Cordless Telecommunications
DECT-NG	Digital Enhanced Cordless Telecommunications - Next Generation
DTH	Direct To Home television
DTS	Digital Theatre System
DTT	Digital Terrestrial Television
DTX	Discontinuous Transmission system
DVB	Digital Video Broadcasting
EFR	Enhanced Full Rate
EVRC	Enhanced Variable Rate Coding
EVRC-B	Enhanced Variable Rate Coding-narrow Band
EVRC-WB	Enhanced Variable Rate Coding-Wide Band
FEC	Frame Erasure Concealment
HF	High Frequency
IMS	IP Multimedia Subsystem
IP	Internet Protocol
IPTV	Internet Protocol TeleVision
ISDN	Integrated Services Digital Network
LC	Low Complexity
LF	Low Frequency
LPC	Linear Predictive Coding
LSP	Line Spectrum Pair
MBMS	Multimedia Broadcast/Multicast Service
MDCT	Modified Discrete Cosine Transform
MIPS	Million Instructions Per Second
MMS	Multimedia Messaging Service
MPEG	Moving Picture Experts Group
MPEGLA	MPEG Lisenced Administrator
MPEGTS	MPEG Transport Stream
NGN	Next Generation Network
PCM	Pulse-Code Modulation
PoC	Push-to-talk over Cellular
PS	Parametric Stereo
PSS	Packet-switched Streaming Service
PSTN	Public Switched Telephone Network

QMF	Quadrature Mirror Filters
RTP	Real-Time Protocol
SBR	Spectral Band Replication
SNR	Signal-to-Noise Ratio
SVC	Scalable Video Coding
TDAC	Time-Domain Aliasing Cancellation
TDBWE	Time-Domain BandWidth Extension
VOD	Video On Demand
VoIP	Voice over IP
WMOPS	Weighted Millions of Operations Per Second

4 Codecs for telephony services

4.1 Services

Services to be supported by conversational speech codecs are IP multimedia services including a voice and/or audio conversational or interactive voice/audio session. Especially, the following services have been specified in TS 181 005 [i.1]:

- PSTN/ISDN emulation service.
- Video telephony service.

4.2 Codecs

4.2.1 General

Speech codecs for telephony services are supported by Customer Network Devices (CND) defined in draft TS 185 006 [i.11] as final devices allowing customers to have access to speech & audio services. This can be a non IP or IP (IMS capable or not) devices either fixed or mobile.

For speech conversational codecs for telephony services, both encoder and decoder have to be supported in CND whereas only the decoder has to be supported in receivers for broadcast/streaming services. This puts design constraints on speech codecs regarding encoding and decoding delay limitation (for conversation interactivity), encoder and decoder complexity limitation and optimization of quality performance for speech.

The voice encoded bandwidth can be narrow band (300 Hz to 3 400 Hz) or wide band (50 Hz to 7 000 Hz) range (or even further extended for some applications). Wideband coding allows enhancing decisively the voice quality: voice is better encoded over all its significant frequencies which produces a feeling of more transparent communication, a greatly improved sensation of presence and an increased intelligibility and listening comfort.

TS 181 005 [i.1] specifies that "*the NGN shall allow end-to-end negotiation of any codec between NGN entities (terminal, network elements)*". The interoperability without transcoding can be consequently achieved if one common codec can be negotiated from end to end between CNDs.

If no common codec is supported between CNDs, transcoding function (decoding and re encoding between the 2 coding formats) has to be implemented (in telco or customer network gateways). However, transcoding degrade quality, add delay, increase network costs and should be consequently reduced. As a consequence, encoding/decoding operations should occur only in CNDs and should be avoided as much as possible in all other devices of the Customer Premises Network (like Customer Network Gateway etc.).

In order to ensure minimum interoperability for narrow band voice services, ETSI 181 005 [i.1] specifies that "narrow band speech encoded format ITU-T Recommendation G.711 [i.12] must be supported". To further improve this interoperability ITU-T Recommendation G.729 [i.44], AMR and EVRC/EVRC-B are recommended. In addition and if wideband optional capability is supported, a restricted list of recommended wideband codecs consisting in ITU-T Recommendation G.722 [i.13], AMR-WB, ITU-T Recommendation G.729.1[i.21] and EVRC WB is specified for better wideband voice interoperability and quality.

NOTE: The standardized speech codecs for telephony services considered in this clause are subject to license costs according to ITU-T, ETSI/3GPP or 3GPP2 IPR Policies except ITU-T Recommendation G.711 [i.12] and ITU-T Recommendation G.722 [i.13] that are royalty free codecs.

4.2.2 Narrow band speech codecs

For narrow band speech codecs, ITU-T Recommendation G.711 [i.12] is already specified in TS 181 005 [i.1] as mandatory coding format. It is stated that: "*In order to enable interworking between the NGN and other networks (including the PSTN, mobile networks and other NGNs) the NGN must be capable of receiving and presenting G.711 coded speech when interconnected with another network*".

ITU-T Recommendation G.711 [i.12] as interoperability format is already widely used for interoperability between deployed networks and terminals (these are CNDs) including between fixed VoIP or PSTN and 3GPP mobile 2G/3G networks and terminals (CNDs). In order to ensure end to end interoperability and for consistency of ETSI standard specifications it is recommended that the Customer Premises Network ("CPN", as defined in draft TS 185 006 [i.11]) will be capable of receiving and presenting ITU-T Recommendation G.711 [i.12] coded speech.

Transcoding to ITU-T Recommendation G.711 [i.12] may be however needed inside the CPN if the CND does not support ITU-T Recommendation G.711 [i.12]. So the CND will support ITU-T Recommendation G.711 [i.12] except if another mandatory codec is already requested by the specific CND access technology (e.g. ITU-T Recommendation G.726 [i.57] for DECT terminals (CNDs) or AMR for mobile 3GPP terminals (CNDs)). In that case, support of ITU-T Recommendation G.711 [i.12] in addition to the mandatory codec is recommended.

The ITU-T Recommendation G.711 [i.12] speech coding algorithm will conform to ITU-T Recommendation G.711 [i.12].

4.2.3 Wide band speech codecs

4.2.3.1 G.722

Application cases and standardization status and references

ITU-T Recommendation G.722 [i.13] has been standardized in 1988 by ITU-T with the purpose to enhance the audio quality of applications like video and audio conferencing over ISDN networks and has been used for some specific radio broadcast usage as well. It is recommended in ITU-T H.320[i.53], H.321[i.54], H.322[i.55] and H.323[i.56] recommendations.

ITU-T Recommendation G.722 [i.13] consists in a detailed codec description provided in [i.13]. Corresponding ANSI-C code is available in the G.722 module of the ITU-T Recommendation G.191 Software Tools Library [i.14] and Test vectors in G.722 Appendix II [i.15].

ITU-T Recommendation G.722 [i.13] has been specified as mandatory wideband codec for New Generation DECT (DECT-NG). The wideband speech service profile based on G.722 at 64 kbit/s is specified in TS 102 527-1 [i.16]. TR 102 570 [i.17] states that *"ITU-T Recommendation G.722 [i.13] codec is chosen as mandatory wideband codec for New Generation DECT in order to greatly increase the voice quality by extending the bandwidth from narrow band to wideband. G.722 provides a high wideband quality at bit rates of 64 kbit/s with low complexity and very low delay".*

ITU-T Recommendation G.722 [i.13] is gaining momentum for enhanced wideband voice services over IP networks thanks to some attractive features like very low delay, low complexity and license free status.

For usage of ITU-T Recommendation G.722 [i.13] over IP networks, format of RTP payload is specified in RFC 3551 [i.18].

ITU-T Recommendation G.722 Appendix III [i.19] and ITU-T Recommendation G.722 Appendix IV [i.20] propose two optional standardized packet loss mechanisms to strongly increase G.722 audio quality for usage over IP networks subject to packet losses.

Overview of main codec features

ITU-T Recommendation G.722 [i.13] has three modes of operation corresponding to the bit rates of 64 kbit/s, 56 kbit/s and 48 kbit/s.

ITU-T Recommendation G.722 [i.13] encoder produces an embedded 64 kbit/s bitstream structured in 3 layers corresponding to these 3 operating modes. The bits corresponding to the last 2 layers can be skipped by the decoder or any other component of the communication systems to dynamically reduce the bit rate to 56 kbit/s or 48 kbit/s which corresponds to 1 or 2 bits "stolen" from the low band.

ITU-T Recommendation G.722 [i.13] is based on sub band ADPCM technology that splits the wideband signal in 2 sub bands (0,4 kHz) and (4,8 kHz) by using Quadrature Mirror Filters (QMF). Both sub bands are then encoded and decoded separately using PCM differential adaptative coding.

Encoding/decoding operations are performed on a sample per sample basis which limits the algorithmic delay to 1 625 ms.

Complexity is limited and can be estimated to around 10 MIPS.

4.2.3.2 G.729.1

Standardization status and usage

ITU-T Recommendation G.729.1 [i.21] has been standardized by ITU-T in May 2006 to improve voice quality (narrow band voice quality and extension to high wideband voice quality) over widely deployed ITU-T Recommendation G.729 [i.44] based VoIP infrastructures: ITU-T Recommendation G.729 [i.44] codec is one of the most widely deployed VoIP codec especially in Enterprise environment due to high compression efficiency (8 kbit/s). ITU-T Recommendation G.729.1 [i.21] coding format includes ITU-T Recommendation G.729 [i.44] coding format to inter work with ITU-T Recommendation G.729 [i.44] installed basis at 8 kbit/s. Purpose is to increase the voice quality up to high quality wideband telephony services over fixed line access with limited impact on existing infrastructure for smooth transition from narrow band to wideband services.

ITU-T Recommendation G.729.1 [i.21] includes a detailed description of the codec, a set of test vectors (in ITU-T Recommendation G.729.1 [i.21] Amendment 1 "New annex A on ITU-T Recommendation G.729.1 [i.21] usage in ITU-T Recommendation H.245 [i.58], plus corrections to the main body and updated test vectors" published in [i.20]) and the fixed point simulation software in ANSI-C Code.

The Low Delay/Low complexity modes are specified in Amendment 3 of ITU-T Recommendation G.729.1 [i.22]

ITU-T Recommendation G.729.1 [i.21] has been specified as optional speech codec to ITU-T Recommendation G.722 [i.13] for DECT new generation. TR 102 570 [i.17] states that "*G.729.1 is recommended as an optional codec for wideband speech to provide even higher wideband quality and better robustness to packets/frames losses than G.722 at half the bit rate of G.722. This allows a better transport efficiency on the network side and over the DECT air interface (one full slot). In addition, it is seamless interoperable with largely deployed G.729 based VoIP networks and terminals.*" The optional wideband speech service profile based on ITU-T Recommendation G.729.1 [i.21] at 32 kbit/s is specified in TS 102 527-1 [i.16].

For usage of ITU-T Recommendation G.729.1 [i.21] over IP networks, format of RTP payload is specified in RFC 4749 [i.23].

The floating point C Code has been standardized in ITU-T Recommendation G.729.1 [i.21] Annex B "New Annex B on a reference floating-point implementation for ITU-T Recommendation G.729.1 [i.21] " and published in [i.21]

A discontinuous transmission system (DTX) with comfort noise generation is specified in G.729.1 Annex C to allow strong reduction of the coding rate during periods with no active speech. The comfort noise generation system generates a silence insertion description each time an update of the ambient background noise parameters is required to maintain the quality of the generated background noise.

Overview of codec technology and main features

The ITU-T Recommendation G.729.1 [i.21] coder is an 8 kbit/s to 32 kbit/s scalable wideband extension of ITU-T Recommendation G.729 [i.44]. It has been designed to support both narrow band (50 Hz/8 kHz to 4 000 Hz/8 kHz frequency sampling) and wideband (50 Hz/16 kHz to7 000 Hz/16 kHz frequency sampling and default sampling rate) to be used as one single codec suited for all VoIP applications to provide optimum state of the art quality for all types of signals (including music) at 32 kbit/s.

The encoder produces an embedded bitstream structured in 12 layers corresponding to 12 scalable operating modes at bit rates between 8 kbit/s to 32 kbit/s: in narrow band mode for the whole range from 8 kbit/s to 32 kbit/s or in wideband mode from 14 kbit/s to 32 kbit/s (by steps of 2 kbit/s) . The scalable architecture provides high flexibility to adapt bit rate, complexity and delay to network constraints and applications needs: the bitstream can be truncated at the decoder side or by any component of the communication systems to adjust "on the fly" the bit rate to the desired value during the session with no session interruption, no need for outband signalling nor audio artefacts .

The ITU-T Recommendation G.729.1 [i.21] "core layer" at 8 kbit/s is fully interoperable with ITU-T Recommendation G.729 [i.44], ITU-T Recommendation G.729 [i.44] Annex A, and ITU-T Recommendation G.729 [i.44] Annex B.

ITU-T Recommendation G.729.1 [i.21] coder operates on 20 ms frames. The input signal is first split into two sub bands using a QMF filterbank and then decimated. The high-pass filtered lower band signal is coded by an 8 kbit/s to12 kbit/s narrowband embedded CELP encoder. The difference between the input and local synthesis signal of the CELP encoder at 12 kbit/s is processed by the perceptual weighting filter. The weighted difference signal is then transformed into frequency domain by MDCT. The spectral folded higher band signal is pre-processed by a low-pass filter with 3 000 Hz cut-off frequency. The resulting signal is coded by the TDBWE encoder and the signal is also transformed into frequency domain by MDCT. The MDCT coefficients of lower band and higher band signal are finally coded by the TDAC encoder. In addition, some parameters are transmitted by the Frame Erasure Concealment (FEC) encoder in order to introduce parameter-level redundancy in the bitstream. This redundancy allows improving quality in the presence of erased frames.

ITU-T Recommendation G.729 [i.44] is implementable on any current platform with a "scalable" maximum complexity of around 35 WMOPS at 32 kbit/s .It has an algorithmic delay of 48,94 ms.

It can be implemented and used in low delay/low complexity configuration modes for narrow band at bit rates below 12 kbit/s or in wideband mode at bit rate limited to 14 kbit/s. For wideband low delay/low complexity mode, the algorithmic delay is reduced to 28,94 ms and the complexity is reduced from 35 WMOPS to below 23,5 WMOPS. For narrow band low delay/low complexity mode, the algorithmic delay is reduced to 25 ms and the complexity is below 20 WMOPS.

4.2.3.3 AMR-WB/G.722.2

Standardization status and usage

AMR-WB has been standardized in 3GPP Release 5 (2002) with detailed description provided in TS 126 190 [i.24] with purpose to provide high quality wideband voice over mobile 3GPP 2G/3G systems. Same specification has been reproduced by ITU-T in ITU-T Recommendation G.722.2 [i.25] and its Annexes.

AMR-WB consists of the multi-rate speech coder, a source controlled rate scheme including a voice activity detector and a comfort noise generation system, and an error concealment mechanism to combat the effects of transmission errors and lost packets.

In 3GPP, AMR-WB is the mandatory codec for several services when wideband speech sampled at 16 kHz is used. These services include circuit switched and packet-switched telephony, 3G-324H multimedia telephony, Multimedia Messaging Service (MMS), Packet-switched Streaming Service (PSS), Multimedia Broadcast/Multicast Service (MBMS), IP multimedia Subsystem (IMS) Messaging and Presence, and Push-to-talk over Cellular (PoC).

For usage over IP networks, format of RTP payload is specified in RFC 4867 [i.39].

In 3GPP, the AMR WB codec has been specified in several specifications:

- TS 126 171 [i.46] gives a general overview of the AMR-WB standards.
- The algorithmic detailed description is given in TS 126 190 [i.24].
- the fixed point and floating point source code are given in TS 126 173 [i.47] and TS 126 204 [i.48], respectively.
- Voice Activity detection is given in TS 126 194 [i.49] and comfort noise aspects are detailed in TS 126 192 [i.50].
- Frame erasure concealment is specified in TS 126 191 [i.51].

Overview of codec technology and main features

AMR-WB is a multi-rate codec that encodes wideband audio signals sampled at 16 kHz (with a signal bandwidth of 50 Hz to 7 000 Hz).

The AMR-WB codec consists of nine modes with bit rates of 23 kbit/s, 85 kbit/s, 23,05 kbit/s, 19,85 kbit/s, 18,25 kbit/s, 15,85 kbit/s, 14,25 kbit/s, 12,65 kbit/s, 8,85 kbit/s and 6,6 kbit/s. The speech coder is capable of switching its bit-rate every 20 ms speech frame upon command.

AMR-WB is capable to provide high quality wideband voice for usage over mobile radio channels. Especially, good quality wideband voice can be supported at 12,65 kbit/s corresponding to almost the same bit rate as currently widely used in GSM/3GPP mobile networks for narrow band voice (12,2 kbit/s) with AMR/EFR codec. Fall back modes down to very low bit rate allow coping with radio access channel constraints.

AMR-WB also includes a 1,75 kbit/s background noise mode that is designed for the Discontinuous Transmission (DTX) operation in GSM and can be used as a low bit rate source-dependent back ground noise mode in other systems.

It has a complexity estimated to 39 WMOPS.

The algorithmic delay is 25 ms.

The codec is based on the Code Excited Linear Prediction (CELP) coding model using Algebraic Codebook (ACELP technology). At each frame, the speech signal is analysed to extract the parameters of the CELP model (LP filter coefficients, adaptive and fixed codebooks' indices and gains). In addition to these parameters, high-band gain indices are computed in 23,85 kbit/s mode. These parameters are encoded and transmitted. At the decoder, these parameters are decoded and speech is synthesized by filtering the reconstructed excitation signal through the LP synthesis filter.

4.2.3.4 EVRC-WB

Standardization status and usage

EVRC-WB is a wideband voice codec standardized by 3GPP2 in February 2007 with recommendation published in 3GPP2 C.S0014-C [i.26]

EVRC-WB has been standardized for wideband voice services in cdma2000. "VoIP Codecs and Protocols for cdma2000" [i.27] states that: "Wideband-capable VoIP terminals (CNDs) shall support EVRC-WB in addition to EVRC-B and EVRC codecs".

For usage over IP networks, format of RTP payload is specified in RFC 4788 [i.40].

EVRC-WB is expected to be deployed widely in 3GPP2 networks starting from mid-2007.

Overview of codec technology and main features

EVRC-WB compresses each 20 ms frame into one of the three sizes: full rate - 171 bits, rate 1/2 bit to 80 bits, and rate 1/8 bit to 16 bits. The input/output sampling rate for EVRC-WB is 16 khz but the codec can also support 8 Khz sampling without re-sampling to 16 KHz. EVRC-WB also supports native VoIP features such as time-warping and DTX. The average active speech bit rate for EVRC-WB is 7,5 kbit/s.

The algorithmic delay is 35 ms.

The EVRC-WB is based on a split-band coding approach in which the wideband input speech (16 KHz sampled) signal is separated into a low frequency (LF) band (0 KHz to 4 KHz) signal and the high frequency (HF) band (3,5 KHz to 7 KHz) signal using an analysis filterbank. The LF signal is encoded using an appropriate coding mode from the EVRC-B (narrow band) [i.40] coding modes, modified to free up bits for the HF band signal coding. The HF signal is encoded using a LPC based coding scheme where the excitation is derived from the coded LF band excitation using non-linear processing. The parameters transmitted corresponding to the HF signal include a set of LSP coefficients, and a set of gain parameters that are obtained by comparing the input HF signal and the HF excitation (derived by non-linear processing) filtered by the LPC synthesis filter.

4.2.3.5 Wideband speech codecs recommendation status

TS 181 005 [i.1] specifies the ITU-T Recommendation G.722 [i.13], ITU-T Recommendation G.729.1 [i.21] AMR-WB/ITU-T Recommendation G.722.2 [i.25] and EVRC WB codecs with a recommended status: it recommends that all 4 codecs should be supported in network and that terminals (CNDs) should support at least one of these codecs. This level of specification does not however guarantee that one same common wideband encoding format is shared between all terminals (CNDs) connected to TISPAN IMS/NGN systems and cannot consequently guarantee wideband interoperability between wideband terminals (CNDs). Transcoding may consequently be needed to ensure such wideband interoperability at a price of additional costs in network (transcoding gateways), additional latency (addition of the encoding/decoding algorithmic delay of 2 codecs) and slight quality degradation (typically between 0,2 MOS and 0,4 MOS).

This interoperability is however currently already ensured by ETSI/3GPP standards between the following possible CNDs:

- Between 3GPP mobile terminals (CNDs) thanks to AMR-WB/ ITU-T Recommendation G.722.2 [i.25] mandatory status in 3GPP wideband terminals (CNDs).

- Between ETSI/DECT NG terminals (CNDs) thanks to ITU-T Recommendation G.722 [i.13] mandatory status in DECT NG wideband terminals (CNDs).

Interoperability within 3GPP2 systems is also ensured between 3GPP2 mobile terminals thanks to EVRC-WB mandatory status in 3GPP2 wideband terminals.

In order to extend this guaranteed interoperability and minimize communication set up failures or fall back to narrow band quality between wideband terminals (CNDs), it is proposed that the support of these codecs be extended to any Wideband CNDs that may be part or connected to TISPAN/NGN R3 systems as specified below:

- ITU-T Recommendation G.722 [i.13] as by default codec to be supported by any Fixed CND for TISPAN R3 NGN wideband telephony services.

- AMR-WB/ ITU-T Recommendation G.722.2 [i.25] to interoperate with 3GPP wideband user equipment and/or user equipment with mobility according to 3GPP access for TISPAN R3 NGN wideband telephony services.

In addition, it is recommended that ITU-T Recommendation G.729.1 [i.21], EVRC WB or both of them be supported by CNDs for TISPAN R3 NGN wideband telephony services:

- ITU-T Recommendation G.729.1 [i.21] where required to support DECT NG user equipment, VoIP and/or legacy user equipment and/or interworking to some VoIP and legacy networks for TISPAN R3 NGN wideband telephony services.

- EVRC-WB where required to interoperate with 3GPP2 wideband user equipment and/or user equipment with mobility according to 3GPP2 access for TISPAN R3 NGN wideband telephony services.

For interworking purpose a "by default" speech packetization size of 20 ms is proposed to be specified if no other packetization size is agreed by bilateral arrangement

Terminals (CNDs) may provide any other codecs in addition to the above list.

4.2.4 Video

4.2.4.1 General

Video telephony services historically imply video capable mobile phones for which all the terminals (CNDs) share similar capabilities. Future video telephony services will target terminals (CNDs) with heterogeneous capabilities such as laptops, PCs and/or set-top-boxes.

Video codecs to be used for video telephony services and relying over 3GPP guidelines are specified in TS 126 114 [i.6].

These guidelines specify the video codecs which are mandatory and recommended, including profiling restrictions and the way the video decoder will behave.

4.2.4.2 H.263

ITU-T Recommendation H.263 [i.41] is the legacy video codec for multimedia telephony that will be supported in Profile 0 Level 45 and below.

Profile 3 of ITU-T Recommendation H.263 [i.41] is only recommended.

4.2.4.3 MPEG-4 Visual (Part2)

MPEG-4 Visual (Part 2) [i.42] should be supported under some constraints specified in clause 5.2.2 of TS 126 114 [i.6].

4.2.4.4 H.264/AVC

ITU-T Recommendation H.264/AVC [i.42] is the state of the art video codec and should be supported under the constraints specified in clause 5.2.2 of TS 126 114 [i.6].

4.2.4.5 SVC amendment 3 of MPEG4-AVC/H.264

SVC is the scalable video coding extension of MPEG4-AVC/H.264 (ISO/IEC 14496-10, ITU Recommendation H.264 [i.42]) allowing for efficient video adaptation without the need of transcoding. SVC provides three types of scalability which can be combined:

- spatial scalability (resolutions);
- temporal scalability (frame rate);
- SNR quality (bit rate).

Scalable coding consists in adding enhancement layers to a core layer interoperable with existing legacy formats. Bits related to enhancement layers can be decoded for enhanced quality or skipped just at decoder side.

For future video telephony services SVC is a candidate technology for achieving best in class services by implementing adaptation mechanisms. SVC avoids the risk of bottom levelled services (less capable terminal (CND) resolution/available bit rate/decoding capabilities). It also avoids the need for transcoding which may impact latency constraints.

Therefore SVC can be used on top of ITU Recommendation H.264 [i.42] as specified in TS 126 114 [i.6] in order to guaranty a minimum level of interoperability.

Further profiling work is still to be done.

5 Codecs for IPTV services

5.1 Services

This clause refer to services define in TS 181 016 [i.3]; the services are described in annex A.

5.2 Codecs

5.2.1 Video

5.2.1.1 General

Video codecs to be used in IPTV services and relying over ETSI guidelines are specified in:

- TS 101 154 [i.10] in the case of IPTV services using MPEG-2 TS at the transport level even if the service is effectively delivered over IP networks. Actually, this applies to the vast majority of current IPTV services for both live TV as well as VOD streaming.
- TS 102 005 [i.9].

These guidelines also specify how video is transported over MPEG2-TS for TS 101 154 [i.10] or directly over IP (RTP) for TS 102 005 [i.9].

TS 101 154 [i.10] provides a very detailed profiling of audio and video codecs while TS 102 005 [i.9] mainly defines general devices capabilities mainly in terms of resolution or picture format and codecs profiles and levels dependencies.

The TS 101 154 [i.10] and TS 102 005 [i.9] are both based on a codec toolbox approach. There is no mandatory codec with TS 101 154 [i.10]. The introduction of a new codec into the DVB guidelines is subject to the following rules:

- there is a DVB market demand from at least 5 constituencies and an end-to-end value chain exists;
- there is a published specification;
- the IPR is fair, reasonable and non-discriminatory;
- the performance has been independently verified.

The following table summarizes the toolboxes:

		DVB market demand	Published Specification	IPR FRaND	Performance independently verified
Video	MPEG-2 (TS 101 154 [i.10] only)	Legacy			
	ITU Recommendation H.264 [i.42]/AVC	✓	✓	(✓) (see note 1)	✓
	ITU Recommendation H.264 [i.42]/SVC	(✓) (see note 1)	(✓) (see note 2)	(✓) (see note 3)	(✓) (see note 4)
	VC-1	✓	✓	✓	(✓) (see note 5)
NOTE 1: SVC market demand - a show of hands at CM-AVC meeting 17 showed 10 companies in support. Documented supported now includes 12 companies. NOTE 2 The video codec specification is published, but encapsulation with MPEGTS and RTP payload are not yet published. NOTE 3: Assumption is that SVC will come under the same licensing terms as ITU Recommendation H.264 [i.42]/AVC - to be verified with MPEGLA. NOTE 4: Test results are available from MPEG - performance to be confirmed by TM-AVC. NOTE 5: Independent testing of VC-1 has been conducted by multiple organizations but results are not publicly available. VC-1 is currently in widespread commercial use and TM-AVC is not aware of any feedback from commercial deployments that is inconsistent with the technical claim that the coding efficiency of VC-1 is significantly better than that of MPEG-2.					

NOTE: TS 102 005 [i.9] is designed in such a way that a specific annex exist given a particular application. Each annex (i.e. each application) can optionally mandate the use of any codec from the generic toolbox. At the moment, the only available annex (annex-B) refers to IP datacast applications such as DVB-H applications. There is no annex dedicated to IPTV services. Typically, while MPEG-2 exist in the toolbox, its support is not required for DVB-H services while it should be for IPTV services for legacy reasons and hybrid DTH/DTT/IPTV devices.

For consistency and backward interoperability with existing ETSI standard for broadcasting services the following codecs are considered in this Study Report

- MPEG-2 only when backward compatibility is required with legacy services. This mainly applies to hybrid broadcast (DTH, DTT) and IPTV services.
- MPEG-4 AVC/ITU Recommendation H.264 [i.42] as the default video codec for deploying IPTV services.

It should be noticed that VC-1 [i.43], while documented in TS 101 154 [i.10], has only been used by few IPTV operators and most of them have migrated to MPEG-4 AVC/ITU Recommendation H.264 [i.42].

5.2.1.2 MPEG-2

MPEG-2 is specified in ISO/IEC 13818-2 [i.45] (ITU Recommendation H.262).

MPEG-2 is the legacy standard definition video codec for DTH, cable and, for some countries, DTT deployments.

As per TS 101 154 [i.10], it is recommended to be used at main profile main level for standard definition TV and at main profile high level for high definition (720p/1 080i formats).

5.2.1.3 MPEG-4 AVC/H.264

MPEG-4 AVC/ITU Recommendation H.264 [i.42] is defined in ISO/IEC 14496-10 [i.52] (ITU Recommendation H.264 [i.42]). This video codec allows achieving the same picture quality than MPEG-2 at half the bit rate of MPEG-2. MPEG-4 AVC/ITU Recommendation H.264 [i.42] is the state of the art video codec for IPTV and high definition DTH deployments.

As per TS 101 154 [i.10], it is recommended to be used at main profile, level 3, optionally at high profile, level 3 for standard definition TV. For high definition, it is recommended to be used at high profile, level 4 for high definition (720p/1 080i). Currently, TS 101 154 [i.10] does not provide any recommendation for 1 080p50/60 high definition (under consideration).

5.2.1.4 SVC amendment 3 of MPEG4-AVC/H.264

SVC is the scalable video coding extension of MPEG4-AVC/ITU Recommendation H.264 [i.42] (ISO/IEC 14496-10 [i.52], ITU Recommendation H.264 [i.42]). SVC provides three types of scalability which can be combined:

- spatial scalability (resolutions);
- temporal scalability (frame rate);
- SNR quality (bit rate).

SVC is currently investigated by DVB for introduction in TS 101 154 [i.10] and TS 102 005 [i.9] guidelines.

5.2.2 Audio

5.2.2.1 General

Audio codecs to be used in DVB services delivered directly over IP protocols are specified in TS 102 005 [i.9].

It is specified in clause 6 of TS 102 005 [i.9]. that "*each IP Integrated Receiver-Decoder (IRD) shall be capable of decoding either audio bitstreams conforming to HE AAC v2 as specified in ISO/IEC 14496-3 or else audio bitstreams conforming to Extended AMR WB (AMR WB+) as specified in TS 126 290 or else audio bitstreams conforming to AC-3 or Enhanced AC-3 as specified in TS 102 366 or any combination of the four.*"

For consistency and backward interoperability with existing ETSI standard for broadcasting services the following codecs are considered in this Study Report:

- HE AAC v2.
- AMR-WB+.
- AC-3/Enhanced AC-3 (also known as Dolby Digital and Dolby Digital Plus).

Besides, for the case of IPTV services using MPEG-2 TS at the transport level even if the service is effectively delivered over IP networks, guidelines for audio encoding are described in clause 6 of TS 101 154 [i.10]. These guidelines are provided for the following codecs:

- MPEG 1 Layer 2.
- MPEG 2 Layer 2.
- AC-3/Enhanced AC-3.
- DTS audio.
- MPEG-4 AAC audio, or MPEG-4 HE-AAC audio, or MPEG-4 HE AACv2.

However the clause does not imply that one or several of these codecs will be supported.

NOTE: All these standardized audio codecs considered in this clause are subject to license costs according to ISO-MPEG or ETSI/3GPP IPR policies.

5.2.2.2 HE AAC v2

NOTE: HE AACv2 is specified by 3GPP under the name eAAC+ (Enhanced aacPlus).

Standardization status and usage

The MPEG 4 High Efficiency AAC v2 audio codec is specified in ISO/IEC 14496-3 [i.29] and TS 126 401 [i.38].

The transport has been specified in ISO/IEC 14496-3 [i.29] in RFC 3690 [i.37].

HE AAC v2 is recommended by 3GPP to be supported in media decoder for the 3GPP transparent end-to-end packet-switched streaming service PSS if audio is supported with the name Enhanced aacPlus.

HE AAC v2 is one of the codecs from the list of codecs in TS 102 005 [i.9] for DVB services delivered directly over IP protocols and specifying that one or more codecs out of this list will be supported.

HE AAC v2 is widely used for TV, DVB-H services and music distribution on mobile environment. It is widely implemented in mobile 3G terminals (CNDs).

Typical operational bit rate range is 24 kbit/s to 64 kbit/s for stereo and around 160 kbit/s for 5.1 representations.

Overview of codec technology and main features

MPEG 4 High Efficiency AAC (HE AAC) is the combination of the MPEG 4 Audio Object Types AAC Low Complexity (LC) and Spectral Band Replication (SBR). It is not a replacement for AAC (Advanced Audio Coding), but rather a superset which extends the range of high quality MPEG 4 Audio to much lower bitrates. HE AAC decoders will decode both, plain AAC and the enhanced AAC plus SBR. The result is a backward compatible extension of the AAC standard.

MPEG 4 HE AAC v2 is the combination of the HE AAC and Audio Object Type Parametric Stereo (PS), which enables stereo coding at very low bitrates. The principle behind the PS tool is to transmit a mono signal coded in HE AAC format together with a description of the stereo image. The PS tool is used at bit rates in the low bit rate range. An enhanced aacPlus decoder is also able to decode AAC-LC content.

The encoded audio bandwidth depends on the sampling frequency and bit rate. The sampling frequency can vary typically from 8 kHz to 96 kHz with bit rates range as follows :

Sample Rate (kHz)	8	11,025	12	16	22,05	24	32	44,1	48
Max Bit Rate/Channel (kbit/s)	48	66,15	72	96	132,3	144	192	264,6	288

Decoder complexity is around 34 MIPS for a typical HE AACv2 stereo content.

5.2.2.3 AMR-WB+

Standardization status and usage

The AMR WB+ codec has been specified in TS 126 290 [i.30] and includes error concealment and also contains a user's guide.

AMR WB+ is recommended by 3GPP to be supported in media decoder for the 3GPP transparent end-to-end packet-switched streaming service PSS if audio is supported

AMR WB+ is one of the codecs from the list of codecs in TS 102 005 [i.9] for DVB services delivered directly over IP protocols and specifying that one or more codecs out of this list will be supported.

The source code for both encoder and decoder has been fully specified in TS 126 304 [i.32] and TS 126 273 [i.33].

The transport has been specified in RFC 4352 [i.34].

This codec is supported in at least one major mobile platform since autumn 2006 and is gaining momentum in mobile 3G terminals (CNDs).

Typical operational bit rate range is 16 kbit/s to 24 kbit/s for stereo.

Overview of codec technology and main features

The extended AMR WB audio codec can encode mono and stereo content, up to 48 kbit/s for stereo. It supports also downmixing to mono at a decoder. Numerous sampling frequencies are supported by the encoder, from 8 kHz up to 48 kHz. In mono, bit rates from 10,4 kbit/s to 24 kbit/s are supported with additional 2 kbit/s to 8 kbit/s bit rate for stereo. Extended AMR-WB decoder is also able to decode AMR-WB content.

Decoder complexity is around 25 MIPS.

5.2.2.4 AC3 and enhanced AC3

Standardization status and usage

The AC-3 and enhanced AC-3 audio codecs are specified in TS 102 366 [i.31].

To transport AC-3 audio, over RTP, the RTP payload specified in RFC 4184 [i.36] is used.

AC3 is used for broadcast TV DVB services. It used for instance for Home Cinema 5.1 products.

Typical operating range for AC3 is 224 kbit/s for stereo and 384 kbit/s for 5.1.

Overview of codec technology and main features

The AC-3 digital compression algorithm can encode from 1 channel to 5,1 channels of source audio from a PCM representation into a serial bit stream, at data rates from 32 kbit/s to 640 kbit/s. The 0,1 channel refers to a fractional bandwidth channel intended to convey only low frequency effect signals.

Bit rates supported are: 32 to 80 (steps of 8), 96, 112, 128, 150 to 256 (steps of 32), 300 to 640 (steps of 64)

Enhanced AC-3 is an evolution of the AC-3 coding system. The addition of a number of low data rate coding tools enables use of Enhanced AC-3 at a lower bit rate than AC-3 for high quality, and use at much lower bit rates than AC-3 for medium quality.

Decoder complexity is around 22 MIPS.

5.2.2.5 Audio codecs recommendation status

The two following ETSI and ETSI/3GPP specifications are the ones already specifying audio codage usage for broadcast/streaming services:

- For DVB services, TS 102 005 [i.9] mandates the DVB IP Receiver support at least one of the codecs listed in clause 5.2.4.1.

- For the 3GPP transparent end-to-end packet-switched streaming service (PSS), if audio is supported, then support of one or both AMR-WB+/HE AAC v2 audio decoders is recommended.

The ETSI/DVB and ETSI/3GPP specifications listed above do not ensure however that one same common encoding format can be decoded by any receiver. In order to reduce the number of content formats to be handled in media servers, and terminals (CNDs) and improve interoperability between contents, HEAACv2 is proposed to be selected as the mandatory default format to be supported by CND devices for TISPAN R2 NGN IPTV services:

- HE AAC v2 is capable to cover a wide range of bit rates for usage over many access technologies/broadcast channels :

  - It can operate efficiently at sufficiently reduced bit rates (down to 24 kbit/s for stereo) to be used over mobile radio channel for mobile and fixed/mobile convergent CNDs. This represents a significant compression efficiency improvement with respect to AAC (typical range for usage is 128 kbit/s for stereo and 320 kbit/s for 5.1) and AC3 (typical range for usage is 224 kbit/s for stereo and 384 kbit/s for 5.1)

  - It is capable to encode at high bit rates up to "transparency" quality with best state of the art quality from bit rates greater than 24 kbit/s to 36 kbit/s depending on the types of signals.

  - 5.1 format is supported.

In addition, support of AMR-WB+ format is recommended for usage in mobile environments allowing to maximize the quality especially on speech and mixed content when bit rates have to be decreased below 24 kbit/s to 36 kbit/s and possibly even further down to 14 kbit/s to 16 kbit/s. Detailed information about quality performance of 3GPP recommended audio codecs HE AACv2 and AMR-WB+ can be found in TR 126.936 [i.35]. These results show that Extended AMR-WB provides better quality than HeAACV2 at low bit rates and for speech and mixed contents (speech + music) whereas HE AACv2 provides better quality performance than Extended AMR-WB at high bit rates and for music contents. The crossover point is between 24 kbit/s and 36 kbit/s for mono depending on the type of content.

CND may implement other codecs in addition to the above list for instance for interoperability with legacy systems like audio codecs specified for MPEG1 Layer II(MP2) or MPEG1 Layer III (MP3) or other codecs listed in [i.10] for transport over MPEG 2 TS.

History

Document history		
V3.1.1	May 2009	Publication