**3GPP TSG- Meeting #**

**, , - Revision of S4-241050**

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| *CR-Form-v12.2* | | | | | | | | |
| **CHANGE REQUEST** | | | | | | | | |
|  | | | | | | | | |
|  |  | **CR** |  | **rev** |  | **Current version:** |  |  |
|  | | | | | | | | |
| *For* [***HE******LP***](http://www.3gpp.org/3G_Specs/CRs.htm#_blank)*on using this form: comprehensive instructions can be found at* [*http://www.3gpp.org/Change-Requests*](http://www.3gpp.org/Change-Requests)*.* | | | | | | | | |
|  | | | | | | | | |

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| ***Proposed change affects:*** | UICC apps |  | ME | **x** | Radio Access Network |  | Core Network | **x** |

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | | | | | | | | | | |
| ***Title:*** |  | | | | | | | | | |
|  |  | | | | | | | | | |
| ***Source to WG:*** | Dolby Sweden AB, Ericsson LM, Fraunhofer IIS, Huawei Technologies Co Ltd., Nokia Corporation, NTT, Orange, Panasonic Holdings Corporation, Philips International B.V., Qualcomm Incorporated, VoiceAge Corporation | | | | | | | | | |
| ***Source to TSG:*** |  | | | | | | | | | |
|  |  | | | | | | | | | |
| ***Work item code:*** |  | | | | |  | ***Date:*** | | |  |
|  |  | | | |  | |  | | |  |
| ***Category:*** |  |  | | | | | ***Release:*** | | |  |
|  | *Use one of the following categories:* ***F*** *(correction)* ***A*** *(mirror corresponding to a change in an earlier release)* ***B*** *(addition of feature),* ***C*** *(functional modification of feature)* ***D*** *(editorial modification)*  Detailed explanations of the above categories can be found in 3GPP [TR 21.900](http://www.3gpp.org/ftp/Specs/html-info/21900.htm). | | | | | | | | *Use one of the following releases: Rel-8 (Release 8) Rel-9 (Release 9) Rel-10 (Release 10) Rel-11 (Release 11) … Rel-17 (Release 17) Rel-18 (Release 18) Rel-19 (Release 19)  Rel-20 (Release 20)* | |
|  |  | | | | | | | | | |
| ***Reason for change:*** | | The new IVAS codec is introduced | | | | | | | | |
|  | |  | | | | | | | | |
| ***Summary of change:*** | | The IVAS codec is introduced by describing codec usage, session negotiation, RTP usage, packet loss handling, adaptation.  Examples are also given for SDP negotiation and bandwidth calculation. | | | | | | | | |
|  | |  | | | | | | | | |
| ***Consequences if not approved:*** | | WI is not completed | | | | | | | | |
|  | |  | | | | | | | | |
| ***Clauses affected:*** | | 2, 3.1, 3.2, 4.2, 5.1, 5.2.1.1, 5.2.1.5, 5.2.1.6, 5.2.1.x (new), 6.2.2.2, 6.2.2.3, 6.2.5.2, 7.4.2, 7.5.2.1.x (new), 8.2.1, 8.2.3.1, 10.2.0, 10.2.1.1, 10.2.1.2, 10.2.1.5, 10.2.1.7, 10.2.1.8, 10.2.1.9, 10.2.1.10, 10.2.1.11, 10.2.1.x1 (new), 10.2.1.x2 (new), [10.2.1.x3 (new)], [10.2.1.x4 (new)], [10.2.1.x5 (new)], [10.2.1.x8 (new)], 10.2.3, 11.1, 18.2.2.1, 18.2.2.3, 18.3.2.2, Annex A.xx (new), Annex A.xx.1 (new), Annex A.xx.1.1 (new), Annex A.xx.2 (new), Annex A.xx.2.1 (new), Annex M.4, Annex YY (new), Annex YY.1 (new), Annex YY.2 (new) | | | | | | | | |
|  | |  | | | | | | | | |
|  | | **Y** | **N** |  | | | |  | | |
| ***Other specs*** | |  | **x** | Other core specifications | | | | TS/TR ... CR ... | | |
| ***affected:*** | |  | **x** | Test specifications | | | | TS/TR ... CR ... | | |
| ***(show related CRs)*** | |  | **x** | O&M Specifications | | | | TS/TR ... CR ... | | |
|  | |  | | | | | | | | |
| ***Other comments:*** | |  | | | | | | | | |
|  | |  | | | | | | | | |
| ***This CR's revision history:*** | |  | | | | | | | | |

\*\*\* First change \*\*\*

# 2 References

…

[x01] 3GPP TS 26.250: "Codec for Immersive Voice and Audio Services – General Overview".

[x02] 3GPP TS 26.251: "Codec for Immersive Voice and Audio Services – C code (fixed-point)".

[x03] 3GPP TS 26.252: "Codec for Immersive Voice and Audio Services – Test Sequences".

[x04] 3GPP TS 26.253: "Codec for Immersive Voice and Audio Services – Detailed Algorithmic Description incl. RTP payload format and SDP parameter definitions".

[x05] 3GPP TS 26.254: "Codec for Immersive Voice and Audio Services – Rendering".

[x06] 3GPP TS 26.255: "Codec for Immersive Voice and Audio Services – Error Concealment of Lost Packets".

[x07] 3GPP TS 26.256: "Codec for Immersive Voice and Audio Services – Jitter Buffer Management".

[x08] 3GPP TS 26.258: "Codec for Immersive Voice and Audio Services – C code (floating-point)".

[x09] 3GPP TS 26.261: "Terminal audio quality performance requirements for immersive audio services".

…

\*\*\* Next change \*\*\*

## 3.1 Definitions

For the purposes of the present document, the terms and definitions given in TR 21.905 [1] and the following apply:

NOTE: A term defined in the present document takes precedence over the definition of the same term, if any, in TR 21.905 [1].

**example:** text used to clarify abstract rules by applying them literally.

**360-degree video:** A real-world visual scene captured by a set of cameras or a camera device with multiple lenses and sensors covering the sphere in all directions around the centre point of the camera set or camera device. The term 360-degree video may be used to include also limited 360-degree video.

**Limited 360-degree video:** A 360-degree video in which the visual scene does not cover the entire sphere around the center point of the camera set or camera device but only a part of it. A limited 360-degree video may be limited i) in the horizontal field to less than 360 degrees, or ii) in the vertical field to less than 180 degrees or iii) in both the vertical and horizontal fields.

**AMR, AMR-NB:** Both names refer to the AMR codec (TS 26.071 [11]) and are used interchangeably in this specification.

**Bitstream:** A bitstream that conforms to a video or audio encoding format.

**bitstream**: A sequence of bits that forms the representation of one or more coded video or audio sequences.

**CHEM:** The Coverage and Handoff Enhancements using Multimedia error robustness feature.

**Codec mode:** Used for the AMR and AMR-WB codecs to identify one specific bitrate. For example AMR includes 8 codec modes (excluding SID), each of different bitrate.

**Constrained terminal:** UE that is (i) operating in radio access capability category series "M" capable of supporting conversational services, and/or (ii) a wearable device which is constrained in size, weight or power consumption (e.g. connected watches), excluding smartphones and feature phones.

**DCMTSI client:** A data channel capable MTSI client supporting data channel media as defined in clause 6.2.10.

**DCMTSI client in terminal:** A DCMTSI client that is implemented in a terminal or UE. The term "DCMTSI client in terminal" is used in this document when entities such as MRFP, MRFC or media gateways are excluded.

**Dual-mono:** A variant of 2-channel stereo encoding where two instances of a mono codec are used to encode a 2-channel stereo signal.

**Evolved UTRAN:** Evolved UTRAN is an evolution of the 3G UMTS radio-access network towards a high-data-rate, low-latency and packet-optimized radio-access network.

**EVS codec:** The EVS codec includes two operational modes: EVS Primary operational mode (‘EVS Primary mode’) and EVS AMR-WB Inter-Operable (‘EVS AMR-WB IO mode’). When using EVS AMR-WB IO mode the speech frames are bitstream interoperable with the AMR-WB codec [18]. Frames generated by an EVS AMR-WB IO mode encoder can be decoded by an AMR-WB decoder, without the need for transcoding. Likewise, frames generated by an AMR-WB encoder can be decoded by an EVS AMR-WB IO mode decoder, without the need for transcoding.

**EVS Primary mode:** Includes 11 bit-rates for fixed-rate or multi-rate operation; 1 average bit-rate for variable bit-rate operation; and 1 bit-rate for SID (TS 26.441 [121]). The EVS Primary can encode narrowband, wideband, super-wideband and fullband signals. None of these bit-rates are interoperable with the AMR-WB codec.

**EVS AMR-WB IO mode:** Includes 9 codec modes and SID. All are bitstream interoperable with the AMR-WB codec (TS 26.171 ‎‎[17]).

**Field of View**: The extent of visible area expressed with vertical and horizontal angles, in degrees in the 3GPP 3DOF reference system as defined in TS 26.118 [180].

**Fisheye Video**: Video captured by a wide-angle camera lens that usually captures an approximately hemispherical field of view and projects it as a circular image.

**Frame Loss Rate (FLR):** The percentage of speech frames not delivered to the decoder. FLR includes speech frames that are not received in time to be used for decoding.

**Immersive audio:** Audio representation having more than mono (e.g. stereo, multi-channel, object-based audio, scene-based audio, metadata-assisted spatial audio), enabling enhanced experiences compared to mono encoding.

**ITT4RT client:** MTSI client supporting the Immersive Teleconferencing and Telepresence for Remote Terminals (ITT4RT) feature, as defined in Annex Y.

**ITT4RT-Tx client:** ITT4RT client only capable of sending immersive video.

**ITT4RT-Rx client:** ITT4RT client only capable of receiving immersive video

**ITT4RT MRF:** An ITT4RT client implemented by functionality included in the MRFC and the MRFP.

**ITT4RT client in terminal:** An ITT4RT client that is implemented in a terminal or UE. The term "ITT4RT client in terminal" is used in this document when entities such as ITT4RT MRF is excluded.

**Mode-set:** Used for the AMR and AMR-WB codecs to identify the codec modes that can be used in a session. A mode-set can include one or more codec modes.

**MSMTSI client:** A multi-stream capable MTSI client supporting multiple streams as defined in Annex S. An MTSI client may support multiple streams, even of the same media type, without being an MSMTSI client. Such an MTSI client may, for example, add a second video to an ongoing video telephony session as shown in Annex A.11. In that case, the MTSI client is an MSMTSI client only if it is fully compliant with Annex S.

**MSMTSI MRF:** An MSMTSI client implemented by functionality included in the MRFC and the MRFP.

**MSMTSI client in terminal:** An MSMTSI client that is implemented in a terminal or UE. The term "MSMTSI client in terminal" is used in this document when entities such as MRFP, MRFC or media gateways are excluded.

**MTSI client:** A function in a terminal or in a network entity (e.g. a MRFP) that supports MTSI.

**MTSI client in terminal:** An MTSI client that is implemented in a terminal or UE. The term "MTSI client in terminal" is used in this document when entities such as MRFP, MRFC or media gateways are excluded.

**MTSI media gateway (or MTSI MGW):** A media gateway that provides interworking between an MTSI client and a non MTSI client, e.g. a CS UE. The term MTSI media gateway is used in a broad sense, as it is outside the scope of the current specification to make the distinction whether certain functionality should be implemented in the MGW or in the MGCF.

**Multi-mono:** A variant of multi-channel encoding where several instances of a mono codec are used to encode a multi-channel signal.

**Omnidirectional media:** Media such as image or video and its associated audio that enable rendering according to the user's viewing orientation, if consumed with a head-mounted device, or according to user's desired viewport, otherwise, as if the user was in the spot where and when the media was captured.

**Operational mode:** Used for the EVS codec to distinguish between EVS Primary mode and EVS AMR-WB IO mode.

**Overlay:** A piece of visual media, rendered over omnidirectional video or image, or a viewport.

**Pose:** Position and rotation information associated to a viewport.

**Processing information data:** Provides additional information that may be used for the IVAS media receiver for rendering the audio signal, see TS 26.253 [x04].

**Projected picture:** Picture that has a representation format specified by an omnidirectional video projection format.

**Projection:** Inverse of the process by which the samples of a projected picture are mapped to a set of positions identified by a set of azimuth and elevation coordinates on a unit sphere.

**Simulcast:** Simultaneously sending different encoded representations (simulcast formats) of a single media source (e.g. originating from a single microphone or camera) in different simulcast streams.

**Simulcast format:** The encoded format used by a single simulcast stream, typically represented by an SDP format and all SDP attributes that apply to that particular SDP format, indicated in RTP by the RTP header payload type field.

**Simulcast stream:** The RTP stream carrying a single simulcast format in a simulcast.

**Viewport**: Region of omnidirectional image or video suitable for display and viewing by the user.

\*\*\* Next change \*\*\*

## 3.2 Abbreviations

For the purposes of the present document, the abbreviations given in TR 21.905 [1] and the following apply:

NOTE: An abbreviation defined in the present document takes precedence over the definition of the same abbreviation, if any, in TR 21.905 [1].

3DOF 3 Degrees of freedom

5GC 5G Core Network

AC Alternating Current

AL-SDU Application Layer - Service Data Unit

AMR Adaptive Multi-Rate

AMR-NB Adaptive Multi-Rate - NarrowBand

AMR-WB Adaptive Multi-Rate - WideBand

AMR-WB IO Adaptive Multi-Rate - WideBand Inter-operable Mode, included in the EVS codec

ANBR Access Network Bitrate Recommendation

ANBRQ Access Network Bitrate Recommendation Query

APP APPlication-defined RTCP packet

ARQ Automatic repeat ReQuest

AS Application Server

ATCF Access Transfer Control Function

ATGW Access Transfer GateWay

AVC Advanced Video Coding

BFCP Binary Floor Control Protocol

CCM Codec Control Messages

CDF Cumulative Distribution Function

cDRX Connected Mode DRX

CHEM Coverage and Handoff Enhancements using Multimedia error robustness feature

CMP Cube-Map Projection

CMR Codec Mode Request

cps characters per second

CS Circuit Switched

CSCF Call Session Control Function

CTM Cellular Text telephone Modem

CVO Coordination of Video Orientation

DBI Delay Budget Information

DRB Data Radio Bearer

DRX Discontinuous Reception

DTLS Datagram Transport Layer Security

DTMF Dual Tone Multi-Frequency

DTX Discontinuous Transmission

ECN Explicit Congestion Notification

ECN-CE ECN Congestion Experienced

ECT ECN Capable Transport

eNodeB E-UTRAN Node B

ERP EquiRectangular Projection

E-UTRAN Evolved UTRAN

EVS Enhanced Voice Services

FECC Far End Camera Control

FIR Full Intra Request

FLR Frame Loss Rate

FoIP Facsimile over IP

FOV Field Of View

GIP Generic IP access

GOB Group Of Blocks

H-ARQ Hybrid - ARQ

HEVC High Efficiency Video Coding

HMD Head Mounted Display HSPA High Speed Packet Access

ICM Initial Codec Mode

IDR Instantaneous Decoding Refresh

IFP Internet Facsimile Protocol

IFT Internet Facsimile Transfer

IMS IP Multimedia Subsystem

IP Internet Protocol

IPv4 Internet Protocol version 4

IRAP Intra Random Access Point

ITT4RT Immersive Teleconferencing and Telepresence for Remote Terminals

ITU-T International Telecommunications Union – Telecommunications

IVAS Immersive Voice and Audio Services

JBM Jitter Buffer Management

MGCF Media Gateway Control Function

MGW Media GateWay

MIME Multipurpose Internet Mail Extensions

MO Management Object

MPEG Moving Picture Experts Group

MRFC Media Resource Function Controller

MRFP Media Resource Function Processor

MSMTSI Multi-Stream Multimedia Telephony Service for IMS

MSRP Message Session Relay Protocol

MTSI Multimedia Telephony Service for IMS

MTU Maximum Transfer Unit

NACK Negative ACKnowledgment

NNI Network-to-Network Interface

NTP Network Time Protocol

OMAF Omnidirectional MediA Format

PCM Pulse Code Modulation

PDCP Packet Data Convergence Protocol

PDP Packet Data Protocol

PI Processing Information

PLI Picture Loss Indication

PLR Packet Loss Ratio

POI Point Of Interconnect

PSTN Public Switched Telephone Network

PTZF Pan, Tilt, Zoom and Focus

QCI QoS Class Identifier

QMC QoE Measurement Collection

QoE Quality of Experience

QoS Quality of Service

QP Quantization Parameter

RoHC Robust HeaderCompression

ROI Region of Interest

RR Receiver Report

RTCP RTP Control Protocol

RTP Real-time Transport Protocol

RWP Region-Wise Packing

SB-ADPCM Sub-Band Adaptive Differential PCM

SC-VBR Source Controlled VBR

SCTP Stream Control Transmission Protocol

SDAP Service Data Adaptation Protocol

SDP Session Description Protocol

SDPCapNeg SDP Capability Negotiation

SEI Supplemental Enhancement Information

SID SIlence Descriptor

SIP Session Initiation Protocol

SR Sender Report

SRVCC Single Radio Voice Call Continuity

TFO Tandem-Free Operation

TISPAN Telecoms and Internet converged Services and Protocols for Advanced Network

TMMBN Temporary Maximum Media Bit-rate Notification

TMMBR Temporary Maximum Media Bit-rate Request

TrFO Transcoder-Free Operation

UDP User Datagram Protocol

UDPTL Facsimile UDP Transport Layer (protocol)

UE User Equipment

VDP Viewport Dependent Processing

VoIP Voice over IP

VOP Video Object Plane

VR Virtual Reality

WebRTC Web Real-Time Communication

\*\*\* Next change \*\*\*

## 4.2 Client

The functional components of a terminal including an MTSI client in terminal using 3GPP access are shown in figure 4.2. An MTSI client in terminal using fixed access can have the same functional components except that it does not have any 3GPP Layer 2 protocol.



NOTE: The grey box marks the scope of the present document.

Figure 4.2: Functional components of a terminal including an MTSI client in terminal using 3GPP access

The scope of the present document is to specify media handling and interaction, which includes media control, media codecs, as well as transport of media and control data. General control-related elements of an MTSI client, such as SIP signalling (TS 24.229 [7]), fall outside this scope, albeit parts of the session setup handling and session control for conversational media are defined here:

- usage of SDP (RFC 4566 [8]) and SDP capability negotiation (RFC 5939 [69]) in SIP invitations for capability negotiation and media stream setup.

- set-up and control of the individual media streams between clients. It also includes interactivity, such as adding and dropping of media components.

Transport of media consists of the encapsulation of the coded media in a transport protocol as well as handling of coded media received from the network. This is shown in figure 4.2 as the "packet based network interface" and is displayed, for conversational media, in more detail in the user-plane protocol stack in figure 4.3. The basic MTSI client defined here specifies media codecs for speech, video, still images and text (see clause 5). A more advanced MTSI client may support several inputs and outputs of the same media type, see TS 26.250 [x01] for an example for immersive audio when the client uses several microphones and/or speakers. Data channels do not require use of any codec but allows for real-time interaction in parallel to the conversational media (see clause 6.2.10). The User interface in Figure 4.2 interacts with a web page and a related script received through a downlink data channel to handle the data channel I/O and data formatting. All conversational media components are transported over RTP with each respective payload format mapped onto the RTP (RFC 3550 [9]) streams. The data channels are using SCTP (RFC 4960 [173]) over DTLS (RFC 8261 [174]), used as specified for WebRTC data channels (RFC 8831 [175]).

A diagram of a conversational multimedia application

Description automatically generated

Figure 4.3: User plane protocol stack for a basic MTSI client

An MTSI client may also support non-conversational media, for example IMS messaging. The functional entities and the protocols used for IMS messaging are described in TS 24.247 [82].

The 3GPP Layer 2 protocol to be interfaced with MTSI client is PDCP [170] for EPC. For 5GC, another user-plane protocol, SDAP [171], is used on top of PDCP as shown in clause 4.4.1 of [164]. It is assumed that the SDAP would be configured without header for both directions in the typical MTSI cases, effectively interfacing with PDCP, as SDAP header would be needed only when more than one QoS flows are multiplexed in a DRB or reflective mapping is enabled.

\*\*\* Next change \*\*\*

## 5.1 Media components

The Multimedia Telephony Service for IMS supports simultaneous transfer of multiple media components with real-time characteristics. Media components denote the actual components that the end-user experiences.

The following media components are considered as core components. Multiple media components (including media components of the same media type) may be present in a session. At least one of the first three of these components is present in all conversational multimedia telephony sessions.

**- Speech/audio:** The sound that is picked up by one or more microphone and transferred from terminal A to terminal B and played out in one or more earphones/loudspeakers. Speech/audio includes detection, transport and generation of DTMF events. Immersive audio may be associated with Processing Information data (PI data) describing, for example, how the audio should be rendered.

**- Video:** The moving image that is, for example, captured by a camera of terminal A, transmitted to terminal B and, for example, rendered on the display of terminal B.

**- Text:** The characters typed on a keyboard or drawn on a screen on terminal A and rendered in real time on the display of terminal B. The flow is time-sampled so that no specific action is needed from the user to request transmission.

**- Data:** Any other data for real-time interaction, closely related to the multimedia telephony session that may be generated or consumed by either one of terminal A or terminal B, possibly via terminal external connections and/or physical connectors, optionally processed by application-specific logic at one or both terminals, and optionally presented on and controlled by the user interface at one or both terminals.

The first three of the above core media components are transported in real time from one MTSI client to the other using RTP (IETF RFC 3550 [9]). The "data" media component for real-time interaction is transported using SCTP (IETF RFC 4960 [173]) over DTLS (IETF RFC 8261 [174]), as described by WebRTC data channels [175]. All media components can be added or dropped during an ongoing session as required either by the end-user or by controlling nodes in the network, assuming that when adding components, the capabilities of the MTSI client support the additional component.

NOTE: The terms voice and speech are synonyms. The present document uses the term speech. The media type is called "audio" in SDP and therefore also the term "audio" is used as synonym.

MTSI specifications also support other media types than the core components described above, for example facsimile (fax) transmission.

Facsimile transmission is described in Annex L.

\*\*\* Next change \*\*\*

5.2.1.1 General codec requirements

MTSI clients in terminals offering speech communication shall support narrowband, wideband and super-wideband communication and should support immersive audio communication. The only exception to this requirement is for the MTSI client in constrained terminal offering speech communication, in which case the MTSI client in constrained terminal shall support narrowband and wideband, and should support super-wideband communication.

NOTE: MTSI clients in terminals refers to the definition in clause 3.1.

In addition, MTSI clients in terminals offering speech communication shall support:

- AMR speech codec (TS 26.071 [11], TS 26.090 [12], TS 26.073 [13] and TS 26.104 [14]) including all 8 modes and source controlled rate operation ‎TS 26.093 [15]. The MTSI client in terminal shall be capable of operating with any subset of these 8 codec modes. More detailed codec requirements for the AMR codec are defined in clause 5.2.1.2.

MTSI clients in terminals offering wideband speech communication at 16 kHz sampling frequency shall support:

- AMR-WB codec (TS 26.171 ‎‎[17], TS 26.190 ‎[18], TS 26.173 ‎[19] and TS 26.204 [20]) including all 9 modes and source controlled rate operation ‎TS 26.193 [21]. The MTSI client in terminal shall be capable of operating with any subset of these 9 codec modes. More detailed codec requirements for the AMR-WB codec are defined in clause 5.2.1.3. When the EVS codec is supported, the EVS AMR-WB IO mode may serve as an alternative implementation of AMR-WB as defined in clause 5.2.1.4.

MTSI clients in terminals offering super-wideband or fullband speech communication shall support:

- EVS codec ( TS 26.441 [121], TS 26.444 [124], TS 26.445 [125], TS 26.447 [127], TS 26.451 [131], TS 26.442 [122], TS 26.452 [165] and TS 26.443 [123]) as described below including functions for backwards compatibility with AMR-WB ( TS 26.446 [126]) and discontinuous transmission ( TS 26.449 [129] and TS 26.450 [130]). More detailed codec requirements for the EVS codec are defined in clause 5.2.1.4.

MTSI clients in terminals offering immersive audio communication:

- shall support IVAS codec (TS 26.250 [x01], TS 26.251 [x02], TS 26.252 [x03], TS 26.253 [x04], TS 26.254 [x05], TS 26.255 [x06], TS 26.256 [x07] and TS 26.258 [x08]) as described below, including functions for backwards compatibility with EVS and AMR-WB interoperable mode as described above. More detailed codec requirements for the IVAS codec are defined in clause 5.2.1.x;NOTE: Split rendering support of IVAS in MTSI is FFS.

- may support dual-mono based on super-wideband or fullband speech communication.

Encoding of DTMF is described in Annex G.

\*\*\* Next change \*\*\*

5.2.1.5 Offering multiple audio bandwidths and multiple channels

MTSI clients in terminals offering wideband speech communication shall also offer narrowband speech communications.

When offering super-wideband speech, both wideband speech and narrowband speech shall also be offered. When offering fullband speech, super-wideband speech, wideband speech and narrowband speech shall also be offered.

MTSI clients in terminals offering dual-mono, shall also offer mono.

MTSI clients in terminals offering immersive audio shall also offer mono audio with the same audio bandwidth(s) as offered for the immersive audio.

NOTE: Split rendering support of IVAS in MTSI is FFS.

\*\*\* Next change \*\*\*

5.2.1.6 Codec preference order

When offering both wideband speech and narrowband speech communication, payload types offering wideband shall be listed before payload types offering only narrowband speech in the ‘m=’ line of the SDP offer (RFC 4566 [8]).

When offering super-wideband speech, wideband and narrowband speech communication, payload types offering super-wideband shall be listed before payload types offering lower bandwidths than super-wideband speech in the ‘m=’ line of the SDP offer (RFC 4566 [8]).

For an MTSI client in terminal supporting EVS the following rules apply when creating the list of payload types on the m= line:

- When the EVS codec is offered for NB by an MTSI client in terminal supporting NB only, it shall be listed before other NB codecs.

- When the EVS codec is offered for up to WB, it shall be listed before other WB codecs.

When dual-mono is offered then this may be preferable over mono depending on the call scenario.

When offering immersive audio, the payload type shall be listed before payload types offering multi-mono, dual-mono and mono audio on the ‘m=’ line of the SDP offer.

\*\*\* Next change \*\*\*

5.2.1.x Detailed codec requirements, IVAS

When the IVAS codec is supported, the EVS mode of the IVAS codec is bitexact with the EVS codec, [125], both for EVS Primary mode and EVS AMR-WB IO mode.

When the IVAS codec is used, Processing Information (PI) data [x04] may need to be transmitted. The PI data may be transported in RTP packets together with the IVAS encoded audio frames.

NOTE: The PI data may also be transmitted as data, e.g., using the data channel, see Clause 6.2.10.

NOTE: Split rendering support of IVAS in MTSI is FFS.

\*\*\* Next change \*\*\*

6.2.2.2 Generating SDP offers

When speech is offered, an MTSI client in terminal sending a first SDP offer in the initial offer-answer negotiation shall include at least one RTP payload type for AMR-NB according to RFC4867 [28] and the MTSI client in terminal shall support and offer a configuration, where the MTSI client in terminal includes the parameter settings as defined in Table 6.1. When EVS-NB is also offered, the MTSI client in terminal shall support and offer a configuration, where the MTSI client in terminal includes the parameter settings for EVS (both EVS Primary and AMR-WB IO modes) as defined in Table 6.2a.

If wideband speech is also offered, then the SDP offer shall also include at least one RTP payload type for AMR-WB according to RFC4867 [28] and the MTSI client in terminal shall support and offer a configuration, where the MTSI client in terminal includes the parameter settings as defined in Table 6.1. When EVS-WB is also offered, the MTSI client in terminal shall support and offer a configuration, where the MTSI client in terminal includes the parameter settings for EVS (both Primary and AMR-WB IO modes) as defined in Table 6.2a. AMR-WB and EVS (including the EVS AMR-WB IO mode) are thus offered using different RTP payload types.

If super-wideband speech is also offered, the SDP offer shall include at least one RTP payload type for EVS and the MTSI client in terminal shall support a configuration where the MTSI client in terminal includes the parameter settings as defined in Table 6.2a.

If fullband speech is also offered, the SDP offer shall include at least one RTP payload type for EVS and the MTSI client in terminal shall support a configuration where the MTSI client in terminal includes the parameter settings as defined in Table 6.2a.

When EVS is offered, the RTP payload type for EVS shall also use parameters for EVS AMR-WB IO mode as defined in Table 6.2a, except for the ‘ecn-capable-rtp’ and ‘leap ect’ parameters. AMR-WB and EVS (including the EVS AMR-WB IO mode) are thus offered using different RTP payload types.

NOTE 1: RFC4867 can also be used for EVS AMR-WB IO when EVS is supported. This may happen after SRVCC when the EVS payload format is used between the ATGW and the MTSI client in terminal while RFC4867 is used between the CS-MGW and the ATGW.

NOTE 2: ECN-triggered adaptation is currently undefined for EVS. This does not prevent ECN-triggered adaptation from being negotiated and used for AMR or AMR-WB.

NOTE 3: When EVS is offered, the audio bandwidths may be different for different directions for the EVS Primary mode, even for ‘sendrecv’ media.

When IVAS is offered, the RTP payload type shall include parameters for immersive audio, EVS Primary mode and EVS AMR-WB IO mode as defined in Table 6.2b.

Clause 5.2.1.6 describes the preference order for how different configurations should be ordered in the list of payload type numbers that is given on the m= line.

**Table 6.1: SDP parameters for AMR-NB or AMR-WB, when the MTSI client in terminal offers the bandwidth-efficient payload format**

|  |  |
| --- | --- |
| **Parameter** | **Usage** |
| octet-align | Shall not be included |
| mode-set | Shall not be included |
| mode-change-period | Shall not be included |
| mode-change-capability | Shall be set to 2 |
| mode-change-neighbor | Shall not be included |
| maxptime | Shall be set to 240, see also Table 7.1 |
| crc | Shall not be included |
| robust-sorting | Shall not be included |
| interleaving | Shall not be included |
| ptime | Shall be set according to Table 7.1 |
| channels | Shall either be set to 1 or be omitted |
| max-red | Shall be included and shall be set to 220 or less |
| ecn-capable-rtp: leap ect=0 | Shall be included if offering to use ECN and if the session setup allows for bit-rate adaptation |

**Table 6.2: SDP parameters for AMR-NB or AMR-WB, when the MTSI client in terminal offers the octet-aligned payload format**

|  |  |
| --- | --- |
| **Parameter** | **Usage** |
| octet-align | Shall be set to 1 |
| mode-set | Shall not be included |
| mode-change-period | Shall not be included |
| mode-change-capability | Shall be set to 2 |
| mode-change-neighbor | Shall not be included |
| maxptime | Shall be set to 240, see also Table 7.1 |
| crc | Shall not be included |
| robust-sorting | Shall not be included |
| interleaving | Shall not be included |
| ptime | Shall be set according to Table 7.1 |
| channels | Shall either be set to 1 or be omitted |
| max-red | Shall be included and shall be set to 220 or less |
| ecn-capable-rtp: leap ect=0 | Shall be included if offering to use ECN and if the session setup allows for bit-rate adaptation |

**Table 6.2a: SDP parameters for EVS (both Primary and AMR-WB IO modes, when the MTSI client in terminal offers EVS**

|  |  |
| --- | --- |
| **Parameter** | **Usage** |
| ptime | Shall be set according to Table 7.1 |
| maxptime | Shall be set to 240, see also Table 7.1 |
| evs-mode-switch | MTSI client in terminal shall not include evs-mode-switch in the initial SDP offer. |
| hf-only | The SDP offer-answer considerations in TS 26.445 [125] apply. |
| dtx | MTSI client in terminal shall not include dtx in the initial SDP offer. |
| dtx-recv | MTSI client in terminal shall not include dtx-recv. |
| max-red | Shall be included and shall be set to 220 or less. |
| channels | The SDP offer-answer considerations in TS 26.445 [125] apply. |
| cmr | The SDP offer-answer considerations in TS 26.445 [125] apply. |
| br | An MTSI client in terminal supporting the EVS codec is required to support the entire bit-rate range but may offer a smaller bit-rate range or even a single bit-rate. |
| br-send | The SDP offer-answer considerations in TS 26.445 [125] apply. |
| br-recv | The SDP offer-answer considerations in TS 26.445 [125] apply. |
| bw | The SDP offer-answer considerations in TS 26.445 [125] apply. |
| bw-send | The SDP offer-answer considerations in TS 26.445 [125] apply. |
| bw-recv | The SDP offer-answer considerations in TS 26.445 [125] apply. |
| ch-send | The SDP offer-answer considerations in TS 26.445 [125] apply. |
| ch-recv | The SDP offer-answer considerations in TS 26.445 [125] apply. |
| ch-aw-recv | The SDP offer-answer considerations in TS 26.445 [125] apply. |
| mode-set | Shall not be included |
| mode-change-period | Shall not be included |
| mode-change-capability | The SDP offer-answer considerations in TS 26.445 [125] apply. |
| mode-change-neighbor | Shall not be included |

**Table 6.2b: SDP parameters for IVAS (including EVS Primary and EVS AMR-WB IO modes)**

|  |  |
| --- | --- |
| **Parameter** | **Usage** |
| ptime | Shall be set according to Table 7.1 |
| maxptime | Shall be set to 240, see also Table 7.1 |
| evs-mode-switch | Shall be set according to Table 6.2a |
| hf-only | If present, it shall be set to 1 |
| dtx | Shall be set according to Table 6.2a |
| dtx-send |  |
| dtx-recv | Shall be set according to Table 6.2a |
| max-red | Shall be included and shall be set to 220 or less. |
| channels | Shall not be present |
| cmr | The SDP offer-answer considerations in TS 26.253 [x04] apply. |
| br | Shall be set according to Table 6.2a |
| br-send | Shall be set according to Table 6.2a |
| br-recv | Shall be set according to Table 6.2a |
| bw | Shall be set according to Table 6.2a |
| bw-send | Shall be set according to Table 6.2a |
| bw-recv | Shall be set according to Table 6.2a |
| ch-send | Shall not be included |
| ch-recv | Shall not be included |
| ch-aw-recv | Shall be set according to Table 6.2a |
| mode-set | Shall be set according to Table 6.2a |
| mode-change-period | Shall be set according to Table 6.2a |
| mode-change-capability | Shall be set according to Table 6.2a |
| mode-change-neighbor | Shall be set according to Table 6.2a |
| ivas-mode-switch | The SDP offer-answer considerations in TS 26.253 [x04] apply. |
| ibr | An MTSI client in terminal supporting the IVAS codec is required to support the entire bit-rate range but may offer a smaller bit-rate range or even a single bit-rate. |
| ibr-send | The SDP offer-answer considerations in TS 26.253 [x04] apply. |
| ibr-recv | The SDP offer-answer considerations in TS 26.253 [x04] apply. |
| ibw | The SDP offer-answer considerations in TS 26.253 [x04] apply. |
| ibw-send | The SDP offer-answer considerations in TS 26.253 [x04] apply. |
| ibw-recv | The SDP offer-answer considerations in TS 26.253 [x04] apply. |
| cf | The SDP offer-answer considerations in TS 26.253 [x04] apply. |
| cf-send | The SDP offer-answer considerations in TS 26.253 [x04] apply. |
| cf-recv | The SDP offer-answer considerations in TS 26.253 [x04] apply. |
| pi-types | The SDP offer-answer considerations in TS 26.253 [x04] apply. |
| pi-types-send | The SDP offer-answer considerations in TS 26.253 [x04] apply. |
| pi-types-recv | The SDP offer-answer considerations in TS 26.253 [x04] apply. |
| pi-br | The SDP offer-answer considerations in TS 26.253 [x04] apply. |
| pi-br-send | The SDP offer-answer considerations in TS 26.253 [x04] apply. |
| pi-br-recv | The SDP offer-answer considerations in TS 26.253 [x04] apply. |

NOTE: Split rendering support of IVAS in MTSI is FFS.

When the channels parameter is omitted then this means that one channel is being offered.

The mode-set parameter is omitted, allowing maximum freedom for the visited network.

The mode-change-capability parameter is included and set to 2 for AMR-NB and AMR-WB, to support potential interworking with 2G radio access (GERAN). For EVS AMR-WB IO it is not required to include the mode-change-capability parameter.

An example of an SDP offer for AMR-NB is shown in Table A.1.1. An example of an SDP offer for both AMR-NB and AMR-WB is shown in Table A.1.2. An example of SDP offer for AMR-NB, AMR-WB, and EVS is shown in Table A.14.1.

An SDP example for offering and accepting a dual-mono session for EVS is shown in Annex A.14.1 and A.14.3.

SDP examples for IVAS are shown in Annex A.xx.

An MTSI client in terminal may divide the offer-answer negotiation into several phases and offer different configurations in different SDP offers. If this is done then the first SDP offer in the initial offer-answer negotiation shall include the most preferable configurations. For AMR-NB, this means that the first SDP offer in the initial offer-answer negotiation shall include at least one RTP payload type for AMR-NB with the parameters as defined in Table 6.1. If wideband speech is offered then the first SDP offer in the initial offer-answer negotiation shall include also at least one RTP payload type for AMR-WB with the parameters as defined in Table 6.1. This also means that offers for octet-aligned payload format do not need to be included in the first SDP offer. If super-wideband or fullband speech is offered, the first SDP offer in the initial offer-answer negotiation shall include at least one RTP payload type for EVS with the parameters as defined in [125]. One example of dividing the offer-answer negotiation into two phases, and the corresponding SDP offers, is shown in clause A.1.1.2.2.

NOTE 4: Dividing the offer-answer negotiation into several phases may lead to never offering the less preferred configurations, if the other end-point accepts to use at least one of the configurations offered in the initial SDP offer.

If the speech media is re-negotiated during the session then the knowledge from earlier offer-answer negotiations should be used in order to shorten the session re-negotiation time. I.e., failed offer-answer transactions shall not be repeated.

\*\*\* Next change \*\*\*

6.2.2.3 Generating SDP answer

An MTSI client in terminal must understand all the payload format options that are defined in RFC 4867 [28], and in [125]. It does not have to support operating according to all these options but must be capable to properly accepting or rejecting all options.

The SDP answer depends on many factors, for example:

- what is included in the SDP offer and in what preference order that is defined. The SDP offer will probably be different if it is generated by another MTSI client in terminal, by an MTSI MGW, a TISPAN client or some other VoIP client that does not follow this specification;

- if terminal and/or network resources are available; and:

- if there are other configurations, for example defined with OMA-DM, that mandate, recommend or prevent some configurations.

Table 6.3 describes requirements and recommendations for handling of the AMR payload format parameters and for how to generate the SDP answer.

NOTE 1: An MTSI client in terminal may support more features than what is required by this specification, e.g. crc, robust sorting and interleaving. Table 6.3 describes the handling of the AMR payload format parameters when the MTSI client implementation supports only those features that are required by this specification. Tables 6.3a-6.3c describe the handling of the EVS payload format parameters.

**Table 6.3: Handling of the AMR-NB and AMR-WB SDP parameters in the received SDP offer and in the SDP answer**

| **Parameter in the received SDP offer** | **Comments** | **Handling** |
| --- | --- | --- |
| Codec | Wide-band speech is preferable over narrow-band speech | If both AMR-WB and AMR-NB are offered and if AMR-WB is supported by the answering MTSI client in terminal then it shall select to use the AMR-WB codec and include this codec in the SDP answer, unless another preference order is indicated in the SDP offer. If the MTSI client in terminal only supports AMR-NB then this codec shall be selected to be used and shall be included in the SDP answer.  The SDP answer shall only include one RTP Payload Type for speech, see NOTE 1. |
| octet-align | Both the bandwidth-efficient and the octet-aligned payload formats are supported by the MTSI client in terminal.  MTSI MGWs for GERAN or UTRAN are likely to either not include the octet-align parameter or to offer octet-align=0.  The bandwidth-efficient payload format is preferable over the octet-aligned payload format. | The offer shall not be rejected purely based on the offered payload format variant.  If both bandwidth-efficient and octet-aligned are included in the received SDP offer then the MTSI client in terminal shall select the bandwidth-efficient payload format and include it in the configuration in the SDP answer. |
| mode-set | The MTSI client in terminal can interoperate properly with whatever mode-set the other end-point offers or if no mode-set is offered.  The possibilities to use the higher bit rate codec modes also depend on the offered bandwidth.  MTSI MGWs for GERAN or UTRAN inter-working are likely to include the mode-set in the offer if in case the intention is to use TFO or TrFO.  Mode sets that give more adaptation possibilities are preferable over mode-sets with fewer or no adaptation possibilities.  An MTSI client in terminal may be configured with a preferred mode set. Otherwise, the preferred mode-set for AMR-NB is {12.2, 7.4, 5.9, 4.75} and for AMR-WB it is {12.65, 8.85 and 6.60}. | The offer shall not be rejected purely based on the offered mode-set.  If only one mode-set is offered then the MTSI client in terminal shall select to use this and include the same mode-set in the SDP answer.  If several different payload types for the same codec with different mode-sets (possibly including one or more payload type without mode set) are included in the received SDP offer, then the MTSI client in terminal should select in the first hand the mode-set that provides the largest degrees of freedom for codec mode adaptation and in the second hand the mode-set that is closest to the preferred mode sets.  If only a payload type without mode-set has been offered, or if an MTSI client in terminal selects a payload type without mode-set from among the offered ones, and the MTSI client in terminal intends to use only some modes (e.g. one of the preferred mode sets defined at left), then the MTSI client in terminal should include these modes as the mode-set.  There are also dependencies between the mode-set and the SDP b=AS bandwidth parameter; see Clause 6.2.5.2. |
| mode-change-period | The MTSI client in terminal can interoperate properly with whatever mode-change-period the other end-point offers.  MTSI MGWs for GERAN or UTRAN inter-working are likely to include mode-change-period=2 in the offer if in case the intention is to use TFO or TrFO. | The offer shall not be rejected purely based on the offered mode-change-period.  If the received SDP offer defines mode-change-period=2 then this information shall be used to determine the mode changes for AMR-NB or AMR-WB encoded media that the MTSI client in terminal sends.  The MTSI client in terminal should not include the mode-change-period parameter in the SDP answer since it has no corresponding limitations. |
| mode-change-capability | The MTSI client in terminal can interoperate with whatever capabilities the other end-point declares. | The offer shall not be rejected purely based on the offered mode-change-capability.  The mode-change-capability information should be used to determine a proper value, or prevent using an improper value, for mode-change-period in the SDP answer, see above. If the offer includes mode-change-capability=1, then the MTSI client in terminal shall not offer mode-change-period=2 in the answer.  The MTSI client in terminal shall include mode-change-capability=2 in the SDP answer since it is required to support restricting mode changes to every other frame. |
| mode-change-neighbor | The MTSI client in terminal can interoperate with whatever limitations the other end-point offers. | The offer shall not be rejected purely based on the offered mode-change-neighbor.  The MTSI client in terminal shall use this information to determine how mode changes can be performed for AMR-NB or AMR-WB encoded media that the MTSI client in terminal sends.  The MTSI client in terminal shall not include the mode-change-neighbor parameter in the SDP answer since it has no corresponding limitations. |
| maxptime | The MTSI client in terminal can interoperate with whatever value that is offered.  The MTSI client in terminal may also use this information to determine a suitable value for max-red in the SDP answer. | The offer shall not be rejected purely based on the offered maxptime.  The MTSI client in terminal shall use this information to control the packetization when sending RTP packets to the other end-point, see also clause 7.4.2.  The maxptime parameter shall be included in the SDP answer and shall be an integer multiple of 20.  If the received SDP offer includes both the max-red and ptime parameter then the MTSI client in terminal may choose to use this information to define a suitable value for maxptime in the SDP answer, see NOTE 2. The MTSI client in terminal may also choose to set the maxptime value to 240, regardless of the ptime and/or max-red parameters in the SDP offer.  The maxptime value in the SDP answer shall not be smaller than ptime value in the SDP answer. The maxptime value should be selected to give at least some room for adaptation. |
| crc | The MTSI client in terminal is not required to support this option. | The MTSI client in terminal may have to reject offered RTP payload types including this option. |
| robust-sorting | The MTSI client in terminal is not required to support this option. | The MTSI client in terminal may have to reject offered RTP payload types including this option. |
| interleaving | The MTSI client in terminal is not required to support this option. | The MTSI client in terminal may have to reject offered RTP payload types including this option. |
| ptime | The MTSI client in terminal can interoperate with whatever value that is offered. | The offer shall not be rejected purely based on the offered ptime.  The MTSI client in terminal should use this information and should use the requested packetization when sending RTP packets to the other end-point. The MTSI client should use the ptime value to determine how many non-redundant speech frames that can be packed into the RTP packets. The requirements in clause 7.4.2 shall be followed even if ptime in the SDP offer is larger than 80.  The ptime parameter shall be included in the SDP answer and shall be an integer multiple of 20.  If the received SDP offer includes the ptime parameters then the MTSI client in terminal may choose to use this information to define a suitable value for ptime in the SDP answer, see NOTE 3. The MTSI client in terminal may also choose to set the ptime value in the SDP answer according to Table 7.1, regardless of the ptime parameter in the SDP offer.  The ptime value in the SDP answer shall not be larger than the maxptime value in the SDP answer. |
| channels | The number of channels may either be explicitly indicated in the SDP by including '/1', '/2', etc. on the a=rtpmap line, but the number of channels may also be omitted. When the number of channels is omitted then the default rule is that one channel is being offered.  The MTSI client in terminal is only required to support audio media using one channel. Offered RTP payload types with more than one channel may therefore have to be rejected. | When the MTSI client in terminal accepts an offer for single-channel audio then the SDP answer shall either explicitly indicate '/1' or omit the channels parameter.  When the MTSI client in terminal accepts an offer for multi-channel audio then the number of channels shall be included in the SDP answer. |
| max-red | The MTSI client in terminal may use this information to bound the delay for receiving redundant frames.  The MTSI client in terminal may also use this information to determine a suitable value for maxptime in the SDP answer. | The max-red parameter shall be included in the SDP answer and shall be an integer multiple of 20.  If the received SDP offer includes both the ptime and maxptime parameters then the MTSI client in terminal may choose to use this information to define a suitable value for max-red in the SDP answer, see NOTE 2. The MTSI client in terminal may also choose to set the max-red value to 220.  The max-red value in the SDP answer should be selected to give at least some room for adaptation. |
| ecn-capable-rtp: leap ect=0 | An MTSI client in terminal uses this SDP attribute to offer ECN for RTP-transported media | Shall be included in the SDP answer if accepting an offer to use ECN and if the session setup allows for bit-rate adaptation |
| NOTE 1: An MTSI client may include both a speech coded, e.g. AMR-NB or AMR-WB, and ‘telephone-events’ for DTMF in the SDP answer, see TS 24.229 Clause 6.1, [7].  NOTE 2: It is possible to use the following relationship between maxptime, ptime and max-red:  maxptime = ptime + max-red. There is however no mandatory requirement that these parameters must be aligned in this way.  NOTE 3: It may be wise to use the same ptime value in the SDP answer as was given in the SDP offer, especially if the ptime in the SDP offer is larger than 20, since a value larger than the frame length indicates that the other end-point is somehow packet rate limited. | | |

If an SDP offer is received from another MTSI client in terminal using the AMR-NB or AMR-WB codec, then the SDP offer will include configurations as described in Table 6.1 and Table 6.2. If the MTSI client in terminal chooses to accept the offer for using the AMR-NB or AMR-WB codec, as configured in Table 6.1 or Table 6.2 then the MTSI client in terminal shall support a configuration where the MTSI client in terminal creates an SDP answer containing an RTP payload type for the AMR-NB and AMR-WB codec as shown in Table 6.4.

**Table 6.3a: Handling of SDP parameters common to EVS Primary and EVS AMR-WB IO in the received SDP offer and in the SDP answer**

| **Parameter** | **Comments** | **Handling** |
| --- | --- | --- |
| ptime |  |  |
| maxptime |  |  |
| evs-mode-switch | This parameter indicates the initial operational mode. This parameter is used by MTSI MGW either when starting in EVS AMR-WB IO mode instead of EVS Primary mode or when switching between EVS Primary mode and EVS AMR-WB IO mode, e.g., for SRVCC. | MTSI client in terminal shall not include evs-mode-switch in the initial SDP offer. When including evs-mode-switch in the SDP offer during a session, the offerer shall use the requested mode when sending EVS packets at the start of the session. However, if a media stream is already being received, the offerer needs to be prepared to receive packets in both EVS primary and EVS AMR-WB IO modes until receiving the answer. When including evs-mode-switch in the SDP answer during a session, the answerer shall use the requested mode when sending EVS packets at the start of the session. When receiving SDP answer including evs-mode-switch during a session, the offerer shall use the requested mode when sending EVS packets.  Note: the operational mode may be changed by adaptation during the session, irrespective of the value of the evs-mode-switch negotiated for the session. |
| hf-only |  | - |
| dtx |  | MTSI client in terminal shall not include dtx in the initial SDP offer. MTSI MGW may modify SDP offer to include dtx in order to disable DTX in the session. |
| dtx-recv |  | MTSI client in terminal shall not include dtx-recv. MTSI MGW may modify SDP offer or answer in order to disable DTX for the send direction of the receiver of dtx-recv. |
| cmr | In EVS AMR-WB IO mode, CMR to the bit-rates of EVS AMR-WB IO mode and NO\_REQ is always enabled. | If cmr=-1 and the session is in the EVS Primary mode, MTSI client in terminal shall not transmit CMR. If cmr=-1 and the session is in the EVS AMR-WB IO, MTSI client in terminal shall restrict CMR to values of EVS AMR-WB-IO bit-rates and NO\_REQ in the session.  MTSI client in terminal is required to accept CMR even when cmr=-1. MTSI client in terminal is required to accept RTP payload without CMR even when cmr=1. |
| max-red | See Table 6.3 | |
| channels | See Table 6.3 | |

**Table 6.3b: Handling of the EVS Primary SDP parameters in the received SDP offer and in the SDP answer**

| **Parameter** | **Comments** | **Handling** |
| --- | --- | --- |
| br |  | An MTSI client in terminal supporting the EVS codec is required to support the entire bit-rate range but may offer a smaller bit-rate range or even a single bit-rate. |
| br-send |  |  |
| br-recv |  |  |
| bw | The session should start with the maximum bandwidth supported by the initial bit-rate up to the maximum negotiated bandwidth. If a range of bandwidth is negotiated, the codec can operate in any bandwidth in the session but the maximum bandwidth in the range should be used after the start of or update of the session. If a single audio bandwidth higher than narrowband is negotiated, the codec operates in the negotiated bandwidth but can use lower bandwidth(s) in the session, depending on the input signal. | Both the offerer and the answerer shall send according to the bandwidth parameter in the answer. |
| bw-send |  |  |
| bw-recv |  |  |
| ch-send |  |  |
| ch-recv |  |  |
| ch-aw-recv |  | If a positive (2, 3, 5, or 7) value of ch-aw-recv is declared for a payload type and the payload type is accepted, the receiver of the parameter shall send partial redundancy (channel-aware mode) at the start of the session using the value as the offset. If ch-aw-recv=0 is declared or not present for a payload type and the payload type is accepted, the receiver of the parameter shall not send partial redundancy (channel-aware mode) at the start of the session. If ch-aw-recv=-1 is declared for a payload type and the payload type is accepted, the receiver of the parameter shall not send partial redundancy (channel-aware mode) in the session. If not present or a non-negative (0, 2, 3, 5, or 7) value of ch-aw-recv is declared for a payload type and the payload type is accepted, partial redundancy (channel-aware mode) can be activated or deactivated during the session based on the expected or estimated channel condition through adaptation signaling, such as CMR (see Annex A.2 of [125]) or RTCP based signalling (see clause 10.2). If not present or a non-negative (0, 2, 3, 5, or 7) value of ch-aw-recv is declared for a payload type and the payload type is accepted, the partial redundancy offset value can also be adjusted during the session based on the expected or estimated channel condition through adaptation signaling. |

**Table 6.3c: SDP parameters for the EVS AMR-WB IO parameters in the received SDP offer and in the SDP answer**

| **Parameter** | **Comments** | **Handling** |
| --- | --- | --- |
| mode-set | See Table 6.3 | |
| mode-change-period |
| mode-change-neighbor |
| mode-change-capability | The default value is re-defined in comparison to that in [28]. | As the default and the only allowed value of mode-change-capability is 2 in EVS AMR-WB IO, it is not required to include this parameter in the SDP offer or answer. |

NOTE 2: ECN-triggered adaptation is currently undefined for EVS. This does not prevent ECN-triggered adaptation from being negotiated and used for AMR or AMR-WB.

**Table 6.4: SDP parameters for AMR-NB or AMR-WB for SDP answer when the SDP offer is received from another MTSI client in terminal**

|  |  |
| --- | --- |
| **Parameter** | **Usage** |
| octet-align | Shall not be included |
| mode-set | See Table 6.3 |
| mode-change-period | Shall not be included |
| mode-change-capability | May be included. If it is included then it shall be set to 2 |
| mode-change-neighbor | Shall not be included |
| maxptime | Shall be set to 240, see also Table 7.1 |
| crc | Shall not be included |
| robust-sorting | Shall not be included |
| interleaving | Shall not be included |
| ptime | Shall be set according to Table 7.1 |
| channels | Shall either be set to 1 or be omitted |
| max-red | Shall be included and shall be set to 220 or less |
| ecn-capable-rtp: leap ect=0 | Shall be included in the SDP answer if accepting an offer to use ECN and if the session setup allows for bit-rate adaptation |

If an SDP offer is received from a MTSI MGW inter-working with CS GERAN/UTRAN, and when the MTSI MGW supports ECN (see also clause 12.3.3), then it is likely to be configured as shown in Table 6.5 if the MTSI MGW does not support redundancy.

**Table 6.5: Expected configuration of SDP parameters for AMR-NB or AMR-WB in an SDP offer from an MTSI MGW inter-working with CS GERAN/UTRAN**

|  |  |
| --- | --- |
| **Parameter** | **Usage** |
| octet-align | Either not included or set to 0 |
| mode-set | Included and indicates the codec modes that are allowed in the CS network |
| mode-change-period | Set to 2 |
| mode-change-capability | Set to 2 |
| mode-change-neighbor | Set to 1 if the CS network is GERAN |
| maxptime | Set to 80, see also Table 12.1 |
| crc | Not included |
| robust-sorting | Not included |
| interleaving | Not included |
| ptime | Set according to Table 12.1 |
| channels | Set to 1 or parameter is omitted |
| max-red | Set to 0 |
| ecn-capable-rtp: leap ect=0 | Shall be included in the SDP answer if accepting an offer to use ECN and if the session setup allows for bit-rate adaptation |

If the MTSI client in terminal accepts the offer included in Table 6.5 then the MTSI client in terminal shall support a configuration where the MTSI client in terminal creates an SDP answer containing an RTP payload type for the AMR-NB and AMR-WB codecs as shown in Table 6.6.

**Table 6.6: SDP parameters for AMR-NB or AMR-WB for SDP answer when the SDP offer is received from another MTSI MGW**

|  |  |
| --- | --- |
| **Parameter** | **Usage** |
| octet-align | Shall be set according to the offer |
| mode-set | See Table 6.3 |
| mode-change-period | Shall not be included |
| mode-change-capability | May be included. If it is included then it shall be set to 2 |
| mode-change-neighbor | Shall not be included |
| maxptime | Shall be set to 240, see also Table 7.1 |
| crc | Shall not be included |
| robust-sorting | Shall not be included |
| interleaving | Shall not be included |
| ptime | Shall be set according to Table 7.1 |
| channels | Shall be set according to the offer |
| max-red | Shall be included and shall be set to 220 or less |
| ecn-capable-rtp: leap ect=0 | Shall be included in the SDP answer if accepting an offer to use ECN and if the session setup allows for bit-rate adaptation |

An MTSI client accepting an offer for IVAS immersive audio shall in the SDP answer include parameters for immersive audio, EVS Primary mode and EVS AMR-WB IO mode.

NOTE: Split rendering support of IVAS in MTSI is FFS.

\*\*\* Next change \*\*\*

6.2.5.2 Speech

If an MTSI client includes an AMR or AMR-WB mode-set, EVS Primary mode br or br-recv parameter, or IVAS ibr or ibr-recv parameter in the SDP offer or answer, the MTSI client shall set the b=AS parameter to a value matching the maximum codec mode in the mode-set or the highest bit-rate in the br or br-recv parameter, or the highest IVAS bit-rate in the ibr or ibr-recv parameter, the packetization time (ptime), and the intended redundancy level. For example, b=AS for AMR-WB at IPv6 should be set to 38 if mode-set includes {6.60, 8.85, 12.65}, the packetization time is 20, and if no extra bandwidth is allocated for redundancy. Likewise, b=AS for EVS Primary mode at IPv4 should be set to 42 if br=7.2-24.4, the packetization is header-full payload format, ptime=20, and no extra bandwidth is allocated for redundancy.

If an MTSI client does not include an AMR or AMR-WB mode-set, or EVS Primary mode br or br-recv parameter, or IVAS ibr or ibr-recv parameter in the SDP offer or answer, the MTSI client shall set the b=AS parameter in the SDP to a value matching the highest AMR/AMR-WB mode, i.e., AMR 12.2 and AMR-WB 23.85, or the highest bit-rate of EVS Primary mode, or the highest bit-rate of IVAS depending on negotiated bandwidth(s), i.e., EVS 24.4 for NB and EVS 128 for WB, SWB and FB, respectively.

NOTE 1: When no mode-set is defined, then this should be understood as that the offerer or answerer is capable of sending and receiving all codec modes of AMR or AMR-WB. An MTSI client in terminal will not include the mode-set parameter in SDP offer in the initial offer-answer negotiation. See Clause 6.2.2.2, Tables 6.1 and 6.2. It is however expected that the mode-set is defined when an SDP offer is received from an MTSI MGW inter-working with CS GERAN/UTRAN, see Clause 6.2.2.3, Table 6.5.

The bandwidth to use for b=AS for AMR and AMR-WB, and EVS Primary mode should be computed as shown in Annexes K and Q respectively. Tables 6.7 and 6.8 shows the bandwidth for the respective AMR and AMR-WB codec when the packetization time is 20 and no extra bandwidth is allocated for redundancy. The b=AS value is computed without taking statistical variations, e.g., the effects of DTX, into account. Such variations can be considered in the scheduling and call admission control. Detailed procedures to compute b=AS of AMR and AMR-WB, and EVS Primary mode can be found in Annexes K and Q.

NOTE 2: For any payload format, b=AS of EVS Primary mode at 5.9 kbps source controlled variable bit-rate (SC-VBR) coding is computed as the b=AS of its highest component bit-rate, 8 kbps.

NOTE 3: b=AS of EVS AMR-WB IO mode can be computed as in the octet-aligned payload format of AMR-WB as shown in Annex K.

b=AS of EVS shall be equal to the maximum of b=AS of the highest included EVS primary mode and b=AS of the highest included EVS AMR-WB IO mode, regardless of the presence and configuration of evs-mode-switch.

**Table 6.7: b=AS for each codec mode of AMR when ptime is 20**

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Payload format** | | **Codec mode** | | | | | | | |
| 4.75 | 5.15 | 5.9 | 6.7 | 7.4 | 7.95 | 10.2 | 12.2 |
| Bandwidth-efficient | IPv4 | 22 | 22 | 23 | 24 | 24 | 25 | 27 | 29 |
| IPv6 | 30 | 30 | 31 | 32 | 32 | 33 | 35 | 37 |
| Octet-aligned | IPv4 | 22 | 22 | 23 | 24 | 25 | 25 | 28 | 30 |
| IPv6 | 30 | 30 | 31 | 32 | 33 | 33 | 36 | 38 |

**Table 6.8: b=AS for each codec mode of AMR-WB when ptime is 20**

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Payload format** | | **Codec Mode** | | | | | | | | |
| 6.6 | 8.85 | 12.65 | 14.25 | 15.85 | 18.25 | 19.85 | 23.05 | 23.85 |
| Bandwidth-efficient | IPv4 | 24 | 26 | 30 | 31 | 33 | 35 | 37 | 40 | 41 |
| IPv6 | 32 | 34 | 38 | 39 | 41 | 43 | 45 | 48 | 49 |
| Octet-aligned | IPv4 | 24 | 26 | 30 | 32 | 33 | 36 | 37 | 40 | 41 |
| IPv6 | 32 | 34 | 38 | 40 | 41 | 44 | 45 | 48 | 49 |

**Table 6.9: b=AS for each bit-rate of EVS Primary mode when ptime is 20**

|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Payload format** | | **Bit-rate** | | | | | | | | | | |
| 7.2 | 8 | 9.6 | 13.2 | 16.4 | 24.4 | 32 | 48 | 64 | 96 | 128 |
| Header-full | IPv4 | 24 | 25 | 27 | 30 | 34 | 42 | 49 | 65 | 81 | 113 | 145 |
| IPv6 | 32 | 33 | 35 | 38 | 42 | 50 | 57 | 73 | 89 | 121 | 153 |

Tables 6.10-1 to 6.10-3 describe the setting of the bandwidth properties that should be used for the ‘a=bw-info’ attribute for a few possible combinations of codec, codec rate, packetization schemes and redundancy levels. The Minimum Supported Bandwidth does not prevent encoding the speech with an even lower bitrate, for example when EVS is used in the 5.9 kbps VBR mode or during DTX periods when SID frames are encoded with a very low bit rate and are generated with a reduced frame rate. Bit rates lower than the Minimum Supported Bandwidth may also be used when sending DTMF. Additional combinations and corresponding bandwidth properties are found in Annex K for AMR, AMR-WB and EVS AMR-WB IO mode and in Annex Q for EVS primary mode.

**Table 6.10-1: Recommended bandwidth properties for AMR to be used with the ‘a=bw-info’ attribute when codec modes up to 12.2 are negotiated**

|  |  |
| --- | --- |
| **Parameter** | **Assumed setting** |
| Negotiated codec modes | 4.75, 5.9, 7.4, 12.2 |
| Codec mode used without redundancy | 12.2 |
| Codec mode used with redundancy | 5.9 |
| Payload format | AMR/AMR-WB bandwidth-efficient |
| Minimum frame aggregation | 1 frame per packet |
| Maximum frame aggregation | 4 frames per packet |
| Maximum redundancy level | 100% |
| Redundancy offset | 0 |
| IP version | 6 |
| **Bandwidth property** | **Value** |
| Maximum Supported Bandwidth | 37 (see NOTE 1) |
| Maximum Desired Bandwidth | 37 |
| Minimum Desired Bandwidth | 31 |
| Minimum Supported Bandwidth | 13 (see NOTE 2) |
| NOTE 1: If redundancy is needed for higher codec modes, if additional redundancy is needed, or if redundancy offset is needed then the Maximum Supported Bandwidth needs to be set to a higher value.  NOTE 2: The Minimum Supported Bandwidth is calculated based on the lowest codec rate when maximum frame aggregation is used. | |

**Table 6.10-2: Recommended bandwidth properties for AMR-WB to be used with the ‘a=bw-info’ attribute when codec modes up to 12.65 are negotiated**

|  |  |
| --- | --- |
| **Parameter** | **Assumed setting** |
| Negotiated codec modes | 6.6, 8.85, 12.65 |
| Codec mode used without redundancy | 12.65 |
| Codec mode used with redundancy | 6.6 |
| Payload format | AMR/AMR-WB bandwidth-efficient |
| Minimum frame aggregation | 1 frame per packet |
| Maximum frame aggregation | 4 frames per packet |
| Maximum redundancy level | 100% |
| Redundancy offset | 0 |
| IP version | 6 |
| **Bandwidth property** | **Value** |
| Maximum Supported Bandwidth | 38 (see NOTE 1) |
| Maximum Desired Bandwidth | 38 |
| Minimum Desired Bandwidth | 32 |
| Minimum Supported Bandwidth | 13 (see NOTE 2) |
| NOTE 1: If redundancy is needed for higher codec modes, if additional redundancy is needed, or if redundancy offset is needed then the Maximum Supported Bandwidth needs to be set to a higher value.  NOTE 2: The Minimum Supported Bandwidth is calculated based on the lowest codec rate when maximum frame aggregation is used. | |

**Table 6.10-3: Recommended bandwidth properties for EVS to be used with the ‘a=bw-info’ attribute when codec modes up to 13.2 are negotiated**

|  |  |
| --- | --- |
| **Parameter** | **Assumed setting** |
| Negotiated codec modes | 5.9 – 13.2 |
| Codec mode used without redundancy | 13.2 |
| Codec mode used with redundancy | 7.2 |
| Payload format | EVS |
| Minimum frame aggregation | 1 frame per packet |
| Maximum frame aggregation | 4 frames per packet |
| Maximum redundancy level | 100% |
| Redundancy offset | 0 |
| IP version | 6 |
| **Bandwidth property** | **Value** |
| Maximum Supported Bandwidth | 40 (see NOTE 1 and NOTE 2) |
| Maximum Desired Bandwidth | 38 |
| Minimum Desired Bandwidth | 32 (see NOTE 2) |
| Minimum Supported Bandwidth | 14 (see NOTE 2) |
| NOTE 1: If redundancy is needed for higher codec modes, if additional redundancy is needed, or if redundancy offset is needed then the Maximum Supported Bandwidth needs to be set to a higher value.  NOTE 2: The bandwidth is calculated assuming EVS 7.2 kbps when maximum frame aggregation is used. | |

SDP examples using the ‘a=bw-info’ attribute for speech are shown in annex A.6.3.

When IVAS is offered, the bandwidth offered in b=AS shall consider the bandwidth needed for the highest negotiated IVAS bitrate using one frame per packet and for the PI data. An example procedure for calculating the required bandwidth is described in Annex [YY].

\*\*\* Next change \*\*\*

7.4.2 Speech

When the AMR codec is selected in the SDP offer-answer negotiation the AMR payload format [28] shall be used between RTP termination points.

When the AMR-WB is selected in the SDP offer-answer negotiation the AMR-WB payload format [28] shall be used between RTP termination points.

NOTE 1: It may happen that EVS AMR-WB IO encoded speech is transported using the AMR-WB payload format between an EVS-capable MTSI client and a legacy (not EVS capable) MTSI client. This may also happen after SRVCC (see Clause 12.3.4) when an EVS-capable MTSI client sends EVS AMR-WB IO encoded speech in EVS payload format to the ATGW and the ATGW then re-packetizes the EVS AMR-WB IO packet into AMR-WB payload format without performing transcoding of the media.

When the EVS codec is selected in the SDP offer-answer negotiation the EVS payload format [125] shall be used between RTP termination points.

NOTE 2: After SRVCC when a CS UE (not EVS capable) sends AMR-WB encoded speech to the ATGW, it may happen that the ATGW then re-packetizes this AMR-WB packet into the EVS payload format without performing transcoding of the media, see clause 12.3.4.

When the IVAS codec is selected in the SDP offer-answer negotiation the IVAS payload format [x04] shall be used between the RTP termination points.

NOTE: Split rendering support of IVAS in MTSI is FFS.

In case of ambiguity the present specification shall take precedence over RFC 4867 [28].

MTSI clients (except MTSI MGW) shall support both the bandwidth-efficient and the octet-aligned payload format of the AMR/AMR-WB payload format [28]. The bandwidth‑efficient payload format shall be preferred over the octet-aligned payload format.

When sending AMR or AMR-WB encoded media, the RTP Marker Bit shall be set according to Section 4.1 of the AMR/AMR-WB payload format [28]. When sending EVS encoded media, the RTP Marker Bit shall be set as described in the EVS payload format [125]. When sending IVAS encoded media, the RTP Marker Bit shall be set as described in the IVAS payload format [x04].

The MTSI clients (except MTSI MGW) should use the SDP parameters defined in table 7.1 for the session. For all access technologies, and for normal operating conditions, the MTSI client should encapsulate the number of non-redundant (a.k.a. primary) speech frames in the RTP packets that corresponds to the ptime value received in SDP from the other MTSI client, or if no ptime value has been received then according to "Recommended encapsulation" defined in table 7.1. The MTSI client may encapsulate more non-redundant speech frames in the RTP packet but shall not encapsulate more than 4 non-redundant speech frames in the RTP packets. The MTSI client may encapsulate any number of redundant speech frames in an RTP packet but the length of an RTP packet, measured in ms, shall never exceed the maxptime value.

NOTE 3: The terminology "non-redundant speech frames" refers to speech frames that have not been transmitted in any preceding packet.

**Table 7.1: Encapsulation parameters (to be used as defined above)**

|  |  |  |  |
| --- | --- | --- | --- |
| **Radio access bearer technology** | **Recommended encapsulation (if no ptime and no RTCP\_APP\_REQ\_AGG has been received)** | **ptime** | **maxptime** |
| Default | 1 non-redundant speech frame per RTP packet  Max 12 speech frames in total but not more than a received maxptime value requires | 20 | 240 |
| HSPA  E-UTRAN  NR | 1 non-redundant speech frame per RTP packet  Max 12 speech frames in total but not more than a received maxptime value requires | 20 | 240 |
| EGPRS | 2 non-redundant speech frames per RTP packet, but not more than a received maxptime value requires  Max 12 speech frames in total but not more than a received maxptime value requires | 40 | 240 |
| GIP | 1 to 4 non-redundant speech frames per RTP packet but not more than a received maxptime value requires.  Max 12 speech frames in total but not more than a received maxptime | 20, 40, 60 or 80 | 240 |

NOTE 4: It is possible to send only redundant speech frames in one RTP packet.

When the radio access bearer technology is not known to the MTSI client, the default encapsulation parameters defined in Table 7.1 shall be used.

When the AMR/AMR-WB payload formats are used, the bandwidth-efficient payload format should be used unless the session setup concludes that the octet-aligned payload format is the only payload format that all parties support. The SDP offer shall include an RTP payload type where octet-align=0 is defined or where octet-align is not specified and should include another RTP payload type with octet-align=1. MTSI client offering wide-band speech shall offer these parameters and parameter settings also for the RTP payload types used for wide-band speech.

For examples of SDP offers and answers, see annex A.

The RTP payload format for DTMF events ís described in Annex G.

\*\*\* Next change \*\*\*

##### 7.5.2.1.x Initial codec mode for IVAS

When the EVS codec embedded in the IVAS is used from the start of the session, then clause 7.5.2.1.8 applies.

When the IVAS Immersive mode is used from the start of the session, the following principles apply for the selection of the Initial Codec Mode bit-rate (ICMbr):

- If GBR is known and if GBR is less than MBR, the ICMbr should be aligned with the GBR or should be lower than the GBR.

When IVAS Immersive mode is used from the start of the session, the Initial IVAS Coded Format (ICMicf) should be the most preferred immersive format negotiated for the session and the Initial Codec Mode audio bandwidth (ICMab) should be the highest audio bandwidth negotiated for the Initial Codec Mode bit-rate (ICMbr).

\*\*\* Next change \*\*\*

8.2.1 Terminology

In the following paragraph(s), Jitter Buffer Management (JBM) denotes the actual buffer as well as any control, adaptation and media processing algorithm (excluding speech decoder) used in the management of the jitter induced in the transport channel. An illustration of an exemplary structure of an MTSI speech receiver with adaptive jitter buffer is shown in figure 8.1 to clarify the terminology and the relation between different functional components.

****

**Figure 8.1: Example structure of an MTSI speech receiver**

The blocks "network analyzer" and "adaptation control logic" together with the information on buffer status form the actual buffer control functionality, whereas "speech decoder" and "adaptation unit" provide the media processing functionality. Note that the external playback device control driving the media processing is not shown in figure 8.1.

The grey dashed lines indicate the measurement points for the jitter buffer delay, i.e. the difference between the decoder consumption time and the arrival time of the speech frame to the JBM.

The functional processing blocks are as follows:

**- Buffer:** The jitter buffer unpacks the incoming RTP payloads and stores the received speech frames. The buffer status may be used as input to the adaptation decision logic. Furthermore, the buffer is also linked to the speech decoder to provide frames for decoding when they are requested for decoding.

**- Network analyser:** The network analysis functionality is used to monitor the incoming packet stream and to collect reception statistics (e.g. jitter, packet loss) that are needed for jitter buffer adaptation. Note that this block can also include e.g. the functionality needed to maintain statistics required by the RTCP if it is being used.

**- Adaptation control logic:** The control logic adjusting playback delay and operating the adaptation functionality makes decisions on the buffering delay adjustments and required media adaptation actions based on the buffer status (e.g. average buffering delay, buffer occupancy, etc.) and input from the network analyser. Furthermore, external control input, including RTCP from the sender, can be used e.g. to enable inter-media synchronisation, to adapt the jitter buffer, or other external scaling requests. The control logic may utilize different adaptation strategies such as fixed jitter buffer (without adaptation and time scaling), simple adaptation during comfort noise periods or buffer adaptation also during active speech. The general operation is controlled with desired proportion of frames arriving late, adaptation strategy and adaptation rate.

**- Speech decoder:** The standard AMR, AMR-WB, EVS or IVAS speech decoder. Note that the speech decoder is also assumed to include error concealment / bad frame handling functionality. Speech decoder may be used with or without the adaptation unit.

**- Adaptation unit:** The adaptation unit shortens or extends the output signal length according to requests given by the adaptation control logic to enable buffer delay adjustment in a transparent manner. The adaptation is performed using the frame based or sample based time scaling on the decoder output signal during comfort noise periods only or during active speech and comfort noise. The buffer control logic should have a mechanism to limit the maximum scaling ratio. Providing a scaling window in which the targeted time scale modifications are performed improves the situation in certain scenarios - e.g. when reacting to the clock drift or to a request of inter-media (re)synchronization - by allowing flexibility in allocating the scaling request on several frames and performing the scaling on a content-aware manner. The adaptation unit may be implemented either in a separate entity from the speech decoder or embedded within the decoder.

\*\*\* Next change \*\*\*

8.2.3.1 General

An MTSI client in terminal supporting speech shall use a JBM fulfilling the minimum performance requirements defined in this clause. The JBM specified in [128] fulfils these minimum performance requirements and should be used for EVS. The JBM specified in [x04] should be used for IVAS.

The jitter buffering time is the time spent by a speech frame in the JBM. It is measured as the difference between the decoding start time and the arrival time of the speech frame to the JBM. The frames that are discarded by the JBM are not counted in the measure.

The minimum performance requirements consist of objective criteria for delay and jitter-induced concealment operations. In order for a JBM implementation to pass the minimum performance requirements all objective criteria shall be met.

A JBM implementation used in MTSI shall comply with the following design guidelines:

1. The overall design of the JBM shall be to minimize the buffering time at all times while still conforming to the minimum performance requirements of jitter induced concealment operations and the design guidelines for sample-based timescaling (as set in bullet point 3);

2. If the limit of jitter induced concealment operations cannot be met, it is always preferred to increase the buffering time in order to avoid growing jitter induced concealment operations going beyond the stated limit above. This guideline applies even if that means that end-to-end delay requirement given in TS 22.105 [34] can no longer be met;

3. If sample-based time scaling is used (after speech decoder), then artefacts caused by time scaling operation shall be kept to a minimum. Time scaling means the modification of the signal by stretching and/or compressing it over the time axis. The following guidelines on time scaling apply:

- Use of a high-quality time scaling algorithm is recommended;

- The amount of scaling should be as low as possible;

- Scaling should be applied as infrequently as possible;

- Oscillating behaviour is not allowed.

NOTE: If the end-to-end delay for the ongoing session is known to the MTSI client in terminal and measured to be less than 150 ms (as defined in TS 22.105 [34]), the JBM may relax its buffering time minimization criteria in favour of reduced JBM adaptation artefacts if such a relaxation will improve the media quality. Note that a relaxation is not allowed when testing for compliance with the minimum performance requirements specified in clauses 8.2.3.2.2 and 8.2.3.2.3.

\*\*\* Next change \*\*\*

10.2.0 General

To reduce the risk for confusion in the media-sender, it is beneficial if the signaling from media-receiver back to media-sender for the media adaptation is the same regardless of which triggers are used in the adaptation-decision in the media-receiver. The ANBR described in clause 10.7 should, if supported by both the access network and the MTSI client in terminal, be used as one such trigger.

NOTE 1: The media-receiver is aware that other nodes in the media path may also influence the media adaptation. A media-receiver sending a specific CMR value X can expect that (after some time) no media is received with a mode higher than X, but modes lower than X may be received any time.

The adaptation for AMR, AMR-WB, EVS and IVAS includes adapting the media bit-rate, the frame aggregation, the redundancy level and the redundancy offset. The domain of adaptation for EVS and IVAS furthermore includes adapting audio bandwidth, partial redundancy, switching between EVS primary mode and EVS AMR-WB IO mode.

NOTE: For IVAS, adapting between different coded formats may require resetting the encoder and/or decoder, which may be audible.

When the AMR codec or the AMR-WB codec is used, two signaling mechanisms are defined:

- CMR in the AMR/AMR-WB RTP payload, [28].   
CMR in RTP can be used by the media-receiver to restrict the codec mode in the remote media-sender to an upper limit (maximum mode).

- RTCP-APP, see clause 10.2.1.   
If the media-sender supports RTCP-APP, then the media-receiver can use it in the following way:  
CMR in RTCP-APP can be used by the media-receiver to restrict the codec mode in the remote media-sender to an upper limit (maximum mode), in addition to CMR in RTP.   
RTCP-APP can further be used by the media-receiver for the adaptation of frame aggregation, redundancy level and redundancy offset in the RTP packets to be sent by the remote media-sender.

When the EVS codec is used, the following signaling mechanism is defined:

- CMR in the EVS RTP payload, [125].

- RTCP-APP, see clause 10.2.1.

When the IVAS codec is used, the following signaling mechanism is defined:

- CMR in the IVAS RTP payload, [x04].

- RTCP-APP, see clause 10.2.1.

In response to received DL ANBR, a speech media receiver should trigger sending CMR requesting bitrate adaptation in the corresponding media sender RTP stream. If RTCP-APP is supported, then a speech media receiver should trigger sending CMR or RTCP-APP requesting bitrate adaptation in the corresponding media sender RTP stream based on the received DL ANBR.

When adapting frame aggregation and/or redundancy, the MTSI client must verify that the maximum packetization, defined by the maxptime SDP parameter, is not exceeded. The MTSI client must also verify that the IP packet sizes does not exceed the Maximum Transfer Unit (MTU).

The boundaries of the adaptation may be controlled by a set of parameters. These parameters may be configured into the MTSI client based on operator policy, for example using OMA-DM.

Table 10.1 defines a mandatory set of parameters that are used by the ECN triggered adaptation for AMR and AMR-WB. The default values for the parameters are also specified. Alternate values for these parameters may be configured into the MTSI client based on operator policy, for example using OMA-DM.

**Table 10.1: Configuration parameters when ECN is used as a trigger**

|  |  |
| --- | --- |
| **Parameter** | **Description** |
| ECN\_min\_rate | Lower boundary for the media bit-rate adaptation in response to ECN-CE marking. The media bit-rate shall not be reduced below this value as a reaction to the received ECN-CE.  The ECN\_min\_rate should be selected to maintain an acceptable service quality while reducing the resource utilization.  Default value: For AMR and AMR-WB, the default value shall be the rate of the recommended Initial Codec Mode, see Clause 7.5.2.1.6. |
| ECN\_congestion\_wait | The waiting time after an ECN-CE marking for which an up-switch shall not be attempted.  A negative value indicates an infinite waiting time, i.e. to prevent up-switch for the whole remaining session.  Default value: 5 seconds |

The configuration of adaptation parameters, and the actions taken during the adaptation, are specific to the particular triggers. For example, the adaptation may be configured to reduce the media bit-rate to AMR5.9 when ECN-CE is detected, while it may reduce the media bit-rate to AMR4.75 for bad radio conditions when high PLR is detected.

Multiple ECN-CE markings within one RTP-level round-trip time is considered as the same congestion event. Each time an MTSI client detects a congestion event it shall send an adaptation request to reduce the media bit-rate unless already operating at the ECN\_min\_rate or below. An MTSI client detecting a congestion event shall not send an adaptation request to increase the media bit-rate for a time period ECN\_congestion\_wait after the end of the congestion event.

Multiple adaptation trigger algorithms can be used in parallel, for example ECN-triggered adaptation, adaptation based on ANBR, and PLR-triggered adaptation. When multiple adaptation algorithms are used for the rate adaptation, the rate that the MTSI client is allowed to use should be no higher than any of the rates determined by each adaptation algorithm.

NOTE 2: For example, if the ECN-triggered adaptation indicates that AMR5.9 should be used and if the PLR-triggered adaptation indicates that AMR4.75 should be used then the rate that the MTSI client uses should be no higher than min(AMR5.9, AMR4.75) = AMR4.75.An example adaptation scheme is described in Annex C.

When additional transport bandwidth information is provided using the ‘a=bw-info’ attribute defined in clause 19, the Minimum Desired Bandwidth should be aligned with the ECN\_min\_rate configuration parameter.

\*\*\* Next change \*\*\*

10.2.1.1 General

When signalling adaptation requests for speech in MTSI, an RTCP-APP packet should be used. This application-specific packet format supports three different adaptation requests when the AMR or AMR-WB codec is used; bit-rate requests, frame aggregation requests and redundancy requests. The requests for frame aggregation and redundancy are also used when the EVS codec or the IVAS codec is used. The codec mode request used for AMR-WB is also used when the EVS AMR-WB IO mode is used and when EVS AMR-WB IO mode is used in the IVAS codec. The application specific format supports additionally five requests that are used for the EVS codec or when IVAS is used in EVS mode. An additional adaptation requests are defined for the IVAS codec. The RTCP-APP packet is put in a compound RTCP packets according to the rules outlined in RFC 3550 [9] and RFC 4585 [40]. In order to keep the size of the RTCP packets as small as possible it is strongly recommended that the RTCP packets are transmitted as minimal compound RTCP packets, meaning that they contain only the items:

- SR or RR;

- SDES CNAME item;

- APP (when applicable).

The recommended RTCP mode is RTCP-AVPF early mode since it will enable transmission of RTCP reports when needed and still comply with RTCP bandwidth rules. The RTCP-APP packets should not be transmitted in each RTCP packet, but rather as a result in the transport characteristics which require end-point adaptation.

The signalling allows for a request that the other endpoint modifies the packet stream to better fit the characteristics of the current transport link. The request in the received RTCP-APP is valid until a new request is received. Note that the media sender can, if having good reasons, choose to not comply with the request received from the media receiver. One such reason could be knowledge of that the local conditions do not allow the requested format.

\*\*\* Next change \*\*\*

#### 10.2.1.2 General Format of RTCP-APP packet with codec control requests

The RTCP-APP packet defined to be used for adaptation signalling for speech in MTSI is constructed as shown in figure 10.1.



Figure 10.1: RTCP-APP formatting

The RTCP-APP specific fields are defined as follows:

- Subtype - the subtype value shall be set to "0".

- Name - the name shall be set to "3GM7", meaning 3GPP MTSI Release 7.

The application-dependent data field contains the requests listed below. The length of the application-dependent data shall be a multiple of 32 bits. The unused bytes shall be set to zero.



Figure 10.2: Basic syntax of the application-dependent data fields when only ID is used

The ID field identifies the request type. ID Code points [0000] ... [1111] are specified in the present document, whereas the other ID code points are reserved for future use.

[

When ID = ‘1111’, the ID field is extended with a 4-bit IDe field. In this case, a 1-byte length field is also added before the codec control request field.

0 1 2

0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 ...

+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+

|ID=1111| IDe | Length | Codec Control Request...

+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+

Figure 10.2.0: Basic syntax of the application-dependent data fields when both ID and IDe are used

When only the ID field is used, the length of the messages is 1 or 2 bytes depending on request type and is specified in the following subclauses.

When the ID field is extended with the IDe field, a 1-byte length field is also included describing the number of bytes used for the codec control request field, i.e., not including the ID, IDe and length fields. An implementation not supporting the IDe and the associated codec control messages may disregard the codec control message while continuing with reading of subsequent codec control messages, if included.

]

The signalling for several different adaptation requests is defined below. For each request, the codecs that can use the request are also specified.

The requests that can be used in a session are negotiated with SDP, see clause 10.2.3.

\*\*\* Next change \*\*\*

10.2.1.5 Codec Mode Request

**RTCP\_APP\_CMR**: Codec Mode Request

****

**Figure 10.5: Codec mode request**

Codecs: This request can only be used for the AMR codec, the AMR-WB codecs, for the EVS codec when operating in AMR-WB IO mode and for the IVAS codec when operating in EVS mode using the AMR-WB IO mode.

The definition of the CMR bits in the RTCP\_APP\_CMR message is identical to the definition of the CMR bits defined in [28]. The CMR indicates the maximum codec mode (highest bit-rate) that the receiver wants to receive. The sender may very well use a lower codec mode (lower bit-rate) when sending.

An MTSI client in terminal that requests mode adaptation should transmit the CMR in an RTCP\_APP\_CMR, unless specified otherwise in Clause 7.3.2.

When the MTSI MGW has an interworking session with a circuit-switched (CS) system using transcoding and requests mode adaptation, the MTSI MGW should transmit CMR in an RTCP\_APP\_CMR, unless specified otherwise in Clause 7.3.2, and should set the CMR in the AMR payload to 15 (no mode request present [28]).

When the MTSI MGW has an interworking session with a circuit-switched (CS) system using TFO/TrFO, then the MTSI media gateway should translate the CMR bits (in GERAN case) or the Iu/Nb rate control messages (in UTRAN case) from the CS client into the CMR bits in the AMR payload. If the MTSI media gateway prefers to receive a lower codec mode rate from the MTSI client in terminal than what the CMR from the CS side indicates, then the MTSI media gateway may replace the CMR from the CS side with the CMR that the MTSI media gateway prefers. The value 15 (no mode request present [28]) shall be used in the CMR bits in the AMR payload towards the PS side if on the CS side no mode request has been received and if the MTSI media gateway has no preference on the used codec mode. The RTCP\_APP\_CMR should not be used in the direction from the MTSI media gateway towards the MTSI client when TFO/TrFO is used.

If an MTSI client receives CMR bits both in the AMR payload and in an RTCP\_APP\_CMR message, the mode with the lowest bit rate of the two indicated modes should be used. A codec mode request received in a RTCP\_APP\_CMR is valid until the next received RTCP\_APP\_CMR.

\*\*\* Next change \*\*\*

10.2.1.7 EVS Primary Rate Request

**RTCP\_APP\_REQ\_EPRR**: EVS Primary Rate Request

****

**Figure 10.6a: EVS primary rate request**

Codecs: This request can be used for the EVS codec and the IVAS codec when operating in EVS Primary mode.

The DATA field a 4-bit field and is encoded as described in the table below. The rate request indicates the maximum codec mode (highest bit-rate) that the receiver wants to receive. The sender may use a lower codec mode (lower bit-rate) when sending. The rate request shall comply with the media type parameters that are negotiated in the session.

**Table 10.1a Encoding of the DATA field in the EVS Primary Rate Request.**

|  |  |
| --- | --- |
| **Index** | **EVS Primary rate request** |
| 0000 | 5.9 |
| 0001 | 7.2 |
| 0010 | 8 |
| 0011 | 9.6 |
| 0100 | 13.2 |
| 0101 | 16.4 |
| 0110 | 24.4 |
| 0111 | 32 |
| 1000 | 48 |
| 1001 | 64 |
| 1010 | 96 |
| 1011 | 128 |
| 1100 | Not used |
| 1101 | Not used |
| 1110 | Not used |
| 1111 | Not used |

10.2.1.8 EVS Bandwidth Request

**RTCP\_APP\_REQ\_EBWR**: EVS Bandwidth Request

****

**Figure 10.6b: EVS bandwidth request**

Codecs: This request can be used for the EVS codec when operating in EVS Primary mode and for the IVAS codec when operating in IVAS immersive mode or EVS Primary mode.

The DATA field is a 4-bit field b0…b3, corresponding to bit 4 to bit 7 in the octet:

- b0 set to ‘1’ = request for narrowband.

- b1 set to ‘1’ = request for wideband.

- b2 set to ‘1’ = request for super-wideband.

- b3 set to ‘1’ = request for fullband.

Each bit in the DATA field indicates a bandwidth that the receiver wants to receive. One or several of these four bits can be set to ‘1’. For example, a request for ‘1110’ indicates that the receiver wants to receive narrowband, wideband or super-wideband speech but not fullband speech. The bandwidth request shall comply with the media type parameters that are negotiated in the session.

A media receiver operating in IVAS Immersive mode should not send a request for narrowband unless also sending a request for switching to EVS Primary mode.

10.2.1.9 EVS Channel Aware Request

**RTCP\_APP\_REQ\_EPRED**: EVS Channel Aware Request

****

**Figure 10.6d: EVS partial redundancy request**

Codecs: This request can be used for the EVS codec when operating in EVS Primary mode and for the IVAS codec when operating in IVAS immersive mode or EVS Primary mode.

The DATA field is a 4-bit field and is encoded as described in the table below.

**Table 10.1.b Encoding of the DATA field in the EVS Channel Aware Request.**

|  |  |
| --- | --- |
| **Index** | **Partial Redundancy request** |
| 0000 | 13.2 CA-L-O2 |
| 0001 | 13.2 CA-L-O3 |
| 0010 | 13.2 CA-L-O5 |
| 0011 | 13.2 CA-L-O7 |
| 0100 | 13.2 CA-H-O2 |
| 0101 | 13.2 CA-H-O3 |
| 0110 | 13.2 CA-H-O5 |
| 0111 | 13.2 CA-H-O7 |
| 1000 | Not used |
| 1001 | Not used |
| 1010 | Not used |
| 1011 | Not used |
| 1100 | Not used |
| 1101 | Not used |
| 1110 | Not used |
| 1111 | Not used |

Since channel-aware mode is only defined for the EVS Primary 13.2 kbps mode then sending an EVS Channel Aware Request also implies changing to the EVS Primary mode and to the 13.2 kbps bit-rate and possibly also changing the audio bandwidth to either WB or SWB. When the IVAS codec isused, this also implies changing to mono audio. In addition, the following applies:

- A media receiver operating in EVS Primary mode for narrowband or fullband audio, IVAS Immersive mode or in IVAS EVS Primary mode for narrowband or fullband audio, and sending a request for EVS channel aware mode shall also include an EVS bandwidth request (subclause 10.2.1.8) for wideband or super-wideband.

The channel aware request shall comply with the media type parameters that are negotiated for the session.

10.2.1.10 EVS Primary mode to EVS AMR-WB IO mode Switching Request

**RTCP\_APP\_REQ\_EP2I**: EVS Primary mode to EVS AMR-WB IO mode Switching Request

****

**Figure 10.6e: EVS primary mode to EVS AMR-WB IO mode switching request**

Codecs: This request can be used for the EVS codec when operating in EVS Primary mode and for the IVAS codec when operating in IVAS immersive mode or EVS Primary mode.

The DATA field is an 11-bit field where the first 9 bits (b4-b12) are used to indicate the AMR-WB codec modes that are allowed and the 2 last bits (b13 and b14) are flags to set mode-change-period and mode-change-neighbor as follows:

- first 9 bits for mode-set:

- b4 = ‘0’: AMR-WB 6.60 not allowed  
b4 = ‘1’: AMR-WB 6.60 allowed,

- b5 = ‘0’: AMR-WB 8.85 not allowed  
b5 = ‘1’: AMR-WB 8.85 allowed,

- b6 = ‘0’: AMR-WB 12.65 not allowed  
b6 = ‘1’: AMR-WB 12.65 allowed,

- b7 = ‘0’: AMR-WB 14.25 not allowed  
b7 = ‘1’: AMR-WB 14.25 allowed,

- b8 = ‘0’: AMR-WB 15.85 not allowed  
b8 = ‘1’: AMR-WB 15.85 allowed,

- b9 = ‘0’: AMR-WB 18.25 not allowed  
b9 = ‘1’: AMR-WB 18.25 allowed,

- b10 = ‘0’: AMR-WB 19.85 not allowed  
b10 = ‘1’: AMR-WB 19.85 allowed,

- b11 = ‘0’: AMR-WB 23.05 not allowed  
b11 = ‘1’: AMR-WB 23.05 allowed,

- b12 = ‘0’: AMR-WB 23.85 not allowed  
b12 = ‘1’: AMR-WB 23.85 allowed.

- flags:

- b13 = ‘0’: mode-change-period=1,  
b13 = ’1’: mode-change-period=2,

- b14 = ‘0’: mode-change-neightbor=0,  
b14 = ‘1’: mode-change-neightbor=1.

An MTSI client sending this request shall set at least one of the mode-set bits to ‘1’. An MTSI client receiving a request with all zeroes shall ignore the request.

The mode-set indicated in the EVS Primary mode to EVS AMR-WB IO mode Switching Request can only allow codec modes that have been negotiated in SDP offer-answer. This request cannot be used to allow codec modes that have not been negotiated in SDP offer-answer.

An MTSI client sending this request should also send an RTCP\_APP\_CMR to indicate the codec mode that should be used after switching to EVS AMR-WB IO mode. An MTSI client receiving this request without a request for a codec mode should use the rules for Initial Codec Mode (ICM) defined in clause 7.5.2.1.6 to determine the codec mode that should be used after switching to EVS AMR-WB IO mode.

The last bit (b15) ‘R’ is reserved for future use. An MTSI client sending this request shall set it to ‘0’. An MTSI client receiving this request shall ignore this bit.

10.2.1.11 EVS AMR-WB IO mode to EVS Primary mode Switching Request

**RTCP\_APP\_REQ\_EI2P**: EVS AMR-WB IO mode to EVS Primary mode Switching Request

****

**Figure 10.6f: EVS AMR-WB IO mode to EVS Primary mode Switching request**

Codecs: This request can be used for the EVS codec when operating in EVS AMR-WB IO mode and for the IVAS codec when operating in IVAS immersive mode or EVS AMR-WB IO mode.

The DATA field is a 4-bit field which is reserved for future use. All four bits are set to ‘0’.

The bitrates and bandwidths that can be used after switching to EVS Primary mode are the same as negotiated at session setup or in a preceding session modification.

\*\*\* Next change \*\*\*

#### 10.2.1.x1 IVAS Coded Format Request

**RTCP\_APP\_REQ\_ICF**: IVAS Coded Format Request



Figure 10.6g: IVAS Coded Format request

Codecs: This request can be used for the IVAS codec when operating in IVAS Immersive mode, EVS Primary mode or EVS AMR-WB IO mode.

The DATA field is a 4-bit field and is encoded as described in the table below:

**Table 10.1.c Encoding of the DATA field in the IVAS Coded Format Request.**

|  |  |  |  |
| --- | --- | --- | --- |
| **Data** | **Identifier** | **Coded Format request** | **Subclause in TS 26.253 [x04]** |
| 0000 | Stereo | Stereo Operation | 4.2.3 |
| 0001 | SBA | Scene-based Audio (SBA, Ambisonics) Operation | 4.2.4 |
| 0010 | MASA | Metadata-Assisted Spatial Audio (MASA) Operation | 4.2.5 |
| 0011 | ISM | Objects (Independent Streams with Metadata, ISM) Operation | 4.2.6 |
| 0100 | MC | Multi-Channel (MC) Operation | 4.2.7 |
| 0101 | OMASA | Combined Objects and MASA (OMASA) Operation | 4.2.9 |
| 0110 | OSBA | Combined Objects and SBA (OSBA) Operation | 4.2.8 |
| 0111 | Not used |  |  |
| 1000 | Not used |  |  |
| 1001 | Not used |  |  |
| 1010 | Not used |  |  |
| 1011 | Not used |  |  |
| 1100 | Not used |  |  |
| 1101 | Not used |  |  |
| 1110 | Not used |  |  |
| 1111 | Not used |  |  |

The IVAS Coded Format request may be used for switching between different IVAS Immersive coded format and also requesting switching from IVAS EVS Primary mode or IVAS EVS AMR-WB IO mode to IVAS Immersive mode.

The coded format request shall comply with the media type parameters that are negotiated for the session.

A media receiver operating in IVAS EVS Primary mode or IVAS EVS AMR-WB IO mode and sending an IVAS Coded Format request [should] also include an IVAS Bitrate request.

A media sender operating in IVAS EVS Primary mode or IVAS EVS AMR-WB IO mode and receiving an IVAS Coded Format request without an IVAS Bitrate request (subclause 10.2.1.x2) should apply the rules for Initial Codec Mode rules described in subclause 7.5.2.1.x.

#### 10.2.1.x2 IVAS Immersive Bit Rate Request

**RTCP\_APP\_REQ\_IBR**: IVAS Bit Rate Request



Figure 10.6h: IVAS Bit Rate request

Codecs: This request can be used for the IVAS codec when operating in IVAS Immersive mode, EVS Primary mode or EVS AMR-WB IO mode.

The DATA field is a 4-bit field and is encoded as described in the table below:

**Table 10.1.d Encoding of the DATA field in the IVAS Bit Rate Request.**

|  |  |
| --- | --- |
| **Data** | **Bit rate [kbps]** |
| 0000 | 13.2 |
| 0001 | 16.4 |
| 0010 | 24.4 |
| 0011 | 32 |
| 0100 | 48 |
| 0101 | 64 |
| 0110 | 80 |
| 0111 | 96 |
| 1000 | 128 |
| 1001 | 160 |
| 1010 | 192 |
| 1011 | 256 |
| 1100 | 384 |
| 1101 | 512 |
| 1110 | 768 |
| 1111 | NO\_REQ |

NOTE: Split rendering support of IVAS in MTSI is FFS.

The IVAS Bit Rate request may be used for switching between different IVAS Immersive mode bitrates.

The IVAS Bit Rate request shall comply with the media type parameters that are negotiated for the session.

A media receiver operating in IVAS EVS Primary mode or IVAS EVS AMR-WB IO mode and sending an IVAS Bit Rate request [should] also include an IVAS Immersive Coded Format request.

[

#### 10.2.1.x3 IVAS Minor Coded Format Request

**RTCP\_APP\_REQ\_IMCF**: IVAS Minor Coded Format Request



Figure 10.6i: IVAS Minor Coded Format request

Codecs: This request can be used for the IVAS codec when operating in IVAS Immersive mode, EVS Primary mode or EVS AMR-WB IO mode.

The DATA field is a [4-bit] field and is encoded as described in the table below:

**Table 10.1.e Encoding of the DATA field in the IVAS Minor Coded Format Request.**

|  |  |  |  |
| --- | --- | --- | --- |
| **Data** | **Identifier** | **Coded Format request** | **Subclause in TS 26.253 [x04]** |
| 0000 | TBD | TBD | TBD |
| 0001 | TBD | TBD | TBD |
| 0010 | TBD | TBD | TBD |
| 0011 | TBD | TBD | TBD |
| 0100 | TBD | TBD | TBD |
| 0101 | TBD | TBD | TBD |
| 0110 | TBD | TBD | TBD |
| 0111 | TBD | TBD | TBD |
| 1000 | TBD | TBD | TBD |
| 1001 | TBD | TBD | TBD |
| 1010 | TBD | TBD | TBD |
| 1011 | TBD | TBD | TBD |
| 1100 | TBD | TBD | TBD |
| 1101 | TBD | TBD | TBD |
| 1110 | TBD | TBD | TBD |
| 1111 | TBD | TBD | TBD |

NOTE: It is FFS to add the encoding of the DATA field and the description.

The IVAS Minor Coded Format request shall comply with the media type parameters that are negotiated for the session.

]

[

#### 10.2.1.x4 IVAS PI Bit Rate Request

**RTCP\_APP\_REQ\_IPIB**: IVAS Processing Information Bit Rate Request



Figure 10.6j: IVAS PI Bit Rate request

Codecs: This request can be used for the IVAS codec.

The DATA field is a 4-bit field and is encoded as described in the table below:

**Table 10.1.f Encoding of the DATA field in the IVAS PI Bit Rate Request.**

|  |  |
| --- | --- |
| **Data** | **Requested PI bit rate in relation to negotiated PI bitrate** |
| 0000 | 0% (disable PI) |
| 0001 | 25% |
| 0010 | 50% |
| 0011 | 75% |
| 0100 | 100% |
| 0101 | Not used |
| 0110 | Not used |
| 0111 | Not used |
| 1000 | Not used |
| 1001 | Not used |
| 1010 | Not used |
| 1011 | Not used |
| 1100 | Not used |
| 1101 | Not used |
| 1110 | Not used |
| 1111 | Not used |

The IVAS PI Bit Rate request may be used for increasing or decreasing the amount of PI data section of the payload used by the media sender. The requested PI bit rate is in relation to the PI bitrate agreed in the session negotiation.

The IVAS PI Bit Rate request shall comply with the media type parameters that are negotiated for the session.

]

[

#### 10.2.1.x5 IVAS PI Redundancy Request

**RTCP\_APP\_REQ\_IPIR**: IVAS Processing Information Bit Rate Request



Figure 10.6k: IVAS PI Redundancy request

Codecs: This request can be used for the IVAS codec.

The DATA field is a 4-bit field and is encoded as described in the table below:

**Table 10.1.g Encoding of the DATA field in the IVAS PI Redundancy Request.**

|  |  |
| --- | --- |
| **Data** | **Requested level for Processing Information** |
| 0000 | 0% (no redundancy) |
| 0001 | 100% |
| 0010 | 200% |
| 0011 | 300% |
| 0100 | 400% |
| 0101 | 500% |
| 0110 | Not used |
| 0111 | Not used |
| 1000 | Not used |
| 1001 | Not used |
| 1010 | Not used |
| 1011 | Not used |
| 1100 | Not used |
| 1101 | Not used |
| 1110 | Not used |
| 1111 | Not used |

The IVAS Redundancy request may be used for increasing or decreasing the amount of redundancy used for the Processing Information.

The IVAS Redundancy request shall comply with the media type parameters that are negotiated for the session.

]

[

#### 10.2.1.x8 IVAS Fragmentation Level Request

**RTCP\_APP\_REQ\_IFLR**: IVAS Fragmentation Level Request



Figure 10.6l: IVAS Fragmentation Level request

Codecs: This request can be used for the IVAS codec.

The DATA field is a 1-byte field and is encoded as described in the table below:

**Table 10.1.h Encoding of the DATA field in the IVAS Fragmentation Request.**

|  |  |
| --- | --- |
| **Data** | **Fragmentation level request** |
| 0000 | No fragmentation of payload[, soft limit] |
| 0001 | Max 1 fragmentation of payload[, soft limit] |
| 0010 | Max 2 fragmentation of payload[, soft limit] |
| 0011 | Max 3 fragmentation of payload[, soft limit] |
| 0100 | Max 4 fragmentation of payload[, soft limit] |
| 0101 | Not used |
| 0110 | Not used |
| 0111 | Not used |
| 1000 | No fragmentation of payload[, hard limit] |
| 1001 | Max 1 fragmentation of payload[, hard limit] |
| 1010 | Max 2 fragmentation of payload[, hard limit] |
| 1011 | Max 3 fragmentation of payload[, hard limit] |
| 1100 | Max 4 fragmentation of payload[, hard limit] |
| 1101 | Not used |
| 1110 | Not used |
| 1111 | Not used |

The length is set to 1 (‘0000 0001’)

The IVAS Fragmentation request may be used for requesting certain fragmentation levels, either with a soft limit or a hard limit. A request for no fragmentation means that the RTP payload shall/should not be fragmented into several RTP packets. A request for 1, 2, 3, or 4 fragments means that the RTP payload may be fragmented into 2, 3, 4 or 5 RTP packets, respectively.

]

NOTE: RTCP-APP for PLAYBACK\_DEVICE\_ORIENTATION, HEAD\_ORIENTATION, LISTENER\_POSITION, DISABLE\_DEVICE\_ORIENTATION\_COMPENSATION, ENABLE\_DEVICE\_ORIENTATION\_COMPENSATION is FFS.

\*\*\* Next change \*\*\*

10.2.3 SDP negotiation for RTCP-APP

RTCP-APP request messages that can be used are negotiated with SDP using the ‘3gpp\_mtsi\_app\_adapt’ attribute. The syntax for the 3GPP MTSI RTCP-APP adaptation attribute is:

a=3gpp\_mtsi\_app\_adapt:<reqNames>

where:

<reqNames> is a comma-separated list identifying the different request messages (see below).

The ABNF for the RTCP-APP adaptation messages negotiation attribute is the following:

adaptation attribute = "a" "=" "3gpp\_mtsi\_app\_adapt" ":" reqName \*("," reqName)

reqName = "RedReq" / "FrameAggReq" / "AmrCmr" / "EvsRateReq" / "EvsBandwidthReq" / "EvsParRedReq" / "EvsIoModeReq" / "EvsPrimaryModeReq" / "IvasCodedFrameReq" / "IvasImmersiveBitRateReq" / ["IvasMinorCodedFormatReq"] / ["IvasPiBitRateReq"] / ["IvasPiRedReq"] / ["IvasFragmentationLevelReq"]

The name denotes the RTCP APP packet types the SDP sender supportes to receive. The meaning of the values is as follows:

RedReq: Redundancy Request, clause 10.2.1.3

FrameAggReq: Frame Aggregation Request, clause 10.2.1.4

AmrCmr: Codec Mode Request for AMR and AMR-WB, clause 10.2.1.5

EvsRateReq: EVS Primary Rate Request, clause 10.2.1.7

EvsBandwidthReq: EVS Bandwidth Request, clause 10.2.1.8

EvsParRedReq: EVS Partial Redundancy Request, clause 10.2.1.9

EvsIoModeReq: EVS Primary mode to EVS AMR-WB IO mode Switching Request, clause 10.2.1.10

EvsPrimaryModeReq: EVS AMR-WB IO mode to EVS Primary mode Switching Request, clause 10.2.1.11

IvasCodedFrameReq: IVAS Coded Format Request, clause 10.2.1.x1

IvasImmersiveBitRateReq: IVAS Bit Rate Request, clause 10.2.1.x2

[IvasMinorCodedFormatReq: IVAS Minor Coded Format Request, clause 10.2.1.x3]

[IvasPiBitRateReq: IVAS Processing Information Bit Rate Request, clause 10.2.1.x4]

[IvasPiRedReq: IVAS Processing Information Bit Rate Request, clause 10.2.1.x5]

[IvasFragmentationLevelReq: IVAS Fragmentation Level Request, clause 10.2.1.x8]

An MTSI client supporting the reception of any RTCP APP packets defined in the present specification shall indicate the supported RTCP APP packet types in an initial SDP offer or answer it sends using the SDP "a=3gpp\_mtsi\_app\_adapt" attribute. If the answerer receives an "a=3gpp\_mtsi\_app\_adapt" attribute in the SDP offer, it may send the indicated RTCP APP packet types towards the offerer. The answerer shall indicate its capabilties with the "a=3gpp\_mtsi\_app\_adapt" attribute irrespective if an "a=3gpp\_mtsi\_app\_adapt" attribute was received and the capabilities within. If the offerer receives an "a=3gpp\_mtsi\_app\_adapt" attribute in the SDP answer, it may send the indicated RTCP APP packet types towards the answerer.

An MTSI client supporting only AMR and AMR-WB therefore may for instance include the following in the SDP offer:

a=3gpp\_mtsi\_app\_adapt: RedReq,FrameAggReq,AmrCmr

An MTSI client supporting only AMR, AMR-WB and EVS may for instance include the following in the SDP offer:

a=3gpp\_mtsi\_app\_adapt: RedReq,FrameAggReq,AmrCmr,EvsRateReq,EvsBandwidthReq,EvsParRedReq,EvsIoModeReq,EvsPrimaryModeReq

The attribute shall only be used on media level.

When interworking with pre-Rel-12 clients or non-MTSI clients, it may happen that they support the RTCP-APP signalling but not the SDP negotiation for AMR and AMR-WB. An MTSI client failing to negotiate RTCP-APP as described may still try to use the RTCP-APP signalling when requesting adaptation, but the MTSI client shall then also monitor the received media in order to determine if some or all of the adaptation requests included in the RTCP-APP were partially or fully followed or not followed at all. If none of the adaptation requests is followed, not even partially, then this is an indication that the remote client does not support the RTCP-APP signalling. The MTSI client should then try to use other means for triggering the adaptation, for example CMR in the AMR/AMR-WB payload or RTCP Sender Reports/Receiver Reports.

\*\*\* Next change \*\*\*

11.1 General

Terminals used for MTSI shall conform to the minimum performance requirements on the acoustic characteristics of 3G terminals specified in TS 26.131 [35]. The codec modes and source control rate operation (DTX) settings shall be as specified in TS 26.132 [36].

Furthermore, the test point (Point-of-Interconnect (POI)) specified in [35] shall be a reference terminal capable of receiving digital speech data at the send side and producing a digital output of the received signal (see figure 11.1). During the testing, the radio conditions should be error free and the jitter and packet loss in the IP transport shall be kept to a minimum.

****

**Figure 11.1: Interface for testing acoustic properties of a terminal used for MTSI**

An MTSI client in terminal supporting the IVAS codec shall conform to the minimum performance requirements on acoustic characteristics specified in TS 26.261 [x09].

\*\*\* Next change \*\*\*

#### 18.2.2.1 Speech codecs

MTSI clients in terminal using fixed access supporting AMR, AMR-WB, EVS or IVAS shall follow clause 5.2.1.

NOTE: Split rendering support of IVAS in MTSI is FFS.

An MTSI client in terminal using fixed access supporting G.711 [77] shall support either A-law PCM or -law PCM and should support both.

MTSI client in terminal using fixed access supporting G.722 shall use the mode operation 1 at 64 kbps as specified in ITU-T Recommendation G.722 [78] when G.722 is used. The bitstream ordering shall be in chronological order with Most Significant Bit (MSB) first.

MTSI client in terminal using fixed access supporting EVRC, EVRC-B, and /or EVRC-WB shall follow 3GPP2 C.S0014-E v1.0 [99] when any of these codecs are used.

Encoding of DTMF is described in Annex G.

\*\*\* Next change \*\*\*

18.2.2.3 Source controlled rate operation

An MTSI client in terminal using fixed access supporting AMR, AMR-WB, EVS or IVAS shall support source controlled rate operation in accordance with clause 5.2.1.

For an MTSI client in terminal using fixed access supporting other codecs than AMR, AMR-WB, EVS or IVAS the following recommendations apply:

- Source controlled rate operation for G.729 should be supported according to Annex B of ITU-T G.729 [100].

- Source controlled rate operation for G.729.1 should be supported according to Annex C and Annex F of ITU-T G.729.1 [101]. Annex C specifies a discontinuous transmission (DTX) and comfort noise generation for G.729.1. Annex F specified the voice activity detector (VAD) to be used together with the DTX/CNG scheme of Annex C to provide the complete functionality of the discontinuous transmission system.

- Source controlled rate operation for G.711 should be supported according to Appendix 2 of ITU-T G.711 [77].

- No source controlled rate operation has been standardized for G.722.

NOTE 1: Use of source controlled rate operation is optional. Source controlled rate operation is known to degrade the speech quality, especially in noisy environments or with background music, and is not needed when both MTSI client in terminals are using fixed access and when the bandwidth is sufficient to ensure best possible voice quality.

NOTE 2: Apart from source controlled rate operation (VAD/DTX) specified in clause 4.19 of 3GPP2 C.S0014-E [99] and in 3GPP2 C.S0076 v1.0 [102], EVRC, EVRC-B, and EVRC-WB can dynamically vary the source coding bit-rate for active speech to achieve a targeted active speech average data rate as specified in 3GPP2 C.S0014-E.

\*\*\* Next change \*\*\*

#### 18.3.2.2 Speech

If an MTSI client in terminal using fixed access supports AMR and/or AMR-WB, EVS and/or IVAS, then clause 6.2.2 applies for session set up.

NOTE: Split rendering support of IVAS in MTSI is FFS.

An MTSI client in terminal using fixed access shall support RTP/AVP. When at least one multi-rate codec is supported (AMR, AMR-WB, EVS or G.729.1) then RTP/AVPF should be supported to allow for end-to-end rate adaptation.

If an MTSI client in terminal using fixed access supports AMR and/or AMR-WB, EVS, or IVAS, then clause 6.2.2.2 applies for generating SDP offers for AMR-NB, AMR-WB and EVS.

An MTSI client in terminal using fixed access supporting both A-law PCM and -law PCM shall offer both variants when sending an SDP offer for G.711.

When an MTSI client in terminal using fixed access supports EVRC-B or EVRC-WB, then clauses 14-18 of RFC 5188 [103] apply when generating SDP offers and answers for EVRC-B and EVRC-WB.

If an MTSI client in terminal using fixed access supports G.729.1 then it also supports G.729 and should offer both G.729.1 and G.729 when sending an SDP offer.

An MTSI client in terminal offering G.729 with source controlled rate operation shall use the parameter "annexb" according to RFC 4855 [107].

The following codec preference order applies for the SDP offer in the session negotiation:

- If AMR-WB is offered it shall be listed first in the codec list (in order of preference, the first codec being preferred).

- If both narrowband codecs and wideband codecs are offered, wideband codecs shall be listed first in the codec list.

When sending the SDP answer, if a wide-band speech session is possible, then selection of narrow-band speech should be avoided whenever possible, unless another preference order is indicated in the SDP offer.

Session setup for sessions including speech and DTMF events is described in Annex G.

\*\*\* Next change \*\*\*

# A.xx SDP offers and answers for speech sessions with IVAS

These examples show SDP offers and answers for speech sessions where IVAS is negotiated. These SDP offer and answer examples are designed to highlight the respective area that is being described and should therefore not be considered as complete SDP offers and answers.

## A.xx.1 SDP offers initiated by MTSI client in terminal

The SDP offers below can be used by MTSI client in terminal, depending on the access technology or the number of audio channels.

### A.xx.1.1 Unknown access technology

When the access technology is unknown to MTSI client in terminal, the SDP offer below can be used to initiate a speech session. In this example, RTP Payload Type 96 is defined for IVAS, RTP Payload Type 97 is defined for EVS, and two sets of RTP Payload Types, 98 and 99, and 100 and 101 are defined for AMR-WB and AMR respectively.

Table A.xx.1: SDP example

|  |
| --- |
| SDP offer |
| m=audio 49152 RTP/AVP 96 97 98 99 100 101  a=tcap:1 RTP/AVPF  a=pcfg:1 t=1  b=AS:556  b=RS:0  b=RR:2000  a=rtpmap:96 IVAS/16000/1  a=fmtp:96 cf=OSBA,OMASA,MC,ISM,SBA,Stereo; ibr=512; [pi-types=xxx; pi-br=20;] max-red=220  a=rtpmap:97 EVS/16000/1  a=fmtp:97 max-red=220  a=rtpmap:98 AMR-WB/16000/1  a=fmtp:98 mode-change-capability=2; max-red=220  a=rtpmap:99 AMR-WB/16000/1  a=fmtp:99 mode-change-capability=2; max-red=220; octet-align=1  a=rtpmap:100 AMR/8000/1  a=fmtp:100 mode-change-capability=2; max-red=220  a=rtpmap:101 AMR/8000/1  a=fmtp:101 mode-change-capability=2; max-red=220; octet-align=1  a=ptime:20  a=maxptime:240 |

NOTE: It is FFS to add the list of PI data types offered for the session.

**Comments:**

The MTSI client in terminal IVAS with up to 512 kbps and all EVS codecs modes, for both sending and receiving directions. For IVAS, all audio bandwidths from wideband to fullband are offered but no parameter is needed since this is default when the ibw parameter is not included. PI date is also offered for both directions with up to 20 kbps. All audio bandwidths are allowed for both IVAS and EVS. For the EVS mode in IVAS, all EVS configuration parameters use their default values.

The clock rate of IVAS is set to 16 kHz.

The media level bandwidth (b=AS) is calculated for the highest offered bitrate of IVAS, 512 kbps, and including 20 kbps for PI data, and then adding 24 kbps for IPv6 overhead, resulting in 556 kbps.

\*\*\* Next change \*\*\*

## A.xx.2 SDP answers from MTSI client in terminal

The SDP answers below can be used by MTSI client in terminal, depending on access technology or service policy. It is assumed that an SDP offer such as described in Table A.xx.1 is received.

### A.xx.2.1 SDP answer from MTSI client in terminal when IVAS is negotiated

In this example, the MTSI client in terminal includes only the IVAS codec in the SDP answer.

Table A.xx.2: SDP example

|  |
| --- |
| SDP answer |
| m=audio 49152 RTP/AVPF 96  a=acfg:1 t=1  b=AS:172  b=RS:0  b=RR:2000  a=rtpmap:96 IVAS/16000/1  a=fmtp:96 cf=Stereo; ibr=128; [pi-types=xxx; pi-br=20;] max-red=220  a=ptime:20  a=maxptime:240 |

NOTE: It is FFS to add the list of PI data types offered for the session.

**Comments:**

For IVAS, stereo at 128 kbps is selected for the session, while all other the configuration parameters are the same as in the received SDP offer.

The media level bandwidth (b=AS) is calculated by adding 128 kbps for IVAS, 20 kbps for PI data and adding 24 kbps for IPv6.

\*\*\* Next change \*\*\*

# M.4 3gpp\_mtsi\_app\_adapt

Contact name, email address, and telephone number:

3GPP Specifications Manager

3gppContact@etsi.org

+33 (0)492944200

Attribute Name (as it will appear in SDP)

3gpp\_mtsi\_app\_adapt

Long-form Attribute Name in English:

3GPP MTSI RTCP-APP Adaptation attribute

Type of Attribute

Media level

Is Attribute Value subject to the Charset Attribute?

This Attribute is not dependent on charset.

Purpose of the attribute:

This attribute is used to negotiate which RTCP-APP request messages that can be used in a session.

Appropriate Attribute Values for this Attribute:

The attribute is a value attribute. The defined values are: "RedReq", "FrameAggReq", "AmrCmr", "EvsRateReq", "EvsBandwidthReq", "EvsParRedReq", "EvsIoModeReq", "EvsPrimaryModeReq", "IvasCodedFrameReq", "IvasImmersiveBitRateReq", ["IvasMinorCodedFormatReq"], ["IvasPiBitRateReq"], ["IvasPiRedReq"], ["IvasFragmentationLevelReq"].

MUX Category for this Attribute:

IDENTICAL-PER-PT

\*\*\* Next change \*\*\*

Annex YY (informative):  
Computation of b=AS for IVAS

# YY.1 General

This annex contains examples of computing b=AS for the IVAS codec when ptime=20. In these examples, it is assumed that no extra bandwidth is allocated for redundancy.

# YY.2 Procedure for computing the bandwidth

The bandwidth is calculated using the following procedure when no extra bandwidth is allocated for redundancy:

1) Use the highest negotiated bitrate for the IVAS codec included in the SDP. Use ibr or ibr-recv parameters, if specified.

2) Add bandwidth needed for PI data. Use [pi-br or pi-br-recv parameters], if specified.

3) Add bandwidth needed for IP, UDP and RTP headers assuming 50 frames per second: 16 kbps for IPv4 and 24 kbps for IPv6.

4) Add bandwidth needed for RTCP.

5) The b=AS bandwidth is the sum of the above listed bitrates after rounding up to nearest integer kbps.

If the SDP includes multiple codecs and/or configurations, the bandwidth is calculated for each configuration and the b=AS bandwidth is set to the highest of the bandwidths.

\*\*\* Last change \*\*\*