ETSI TS 126 223 V13.1.0 (2016-08)

TECHNICAL SPECIFICATION

Universal Mobile Telecommunications System (UMTS);
LTE;
Telepresence using the IP Multimedia Subsystem (IMS);
Media handling and interaction
(3GPP TS 26.223 version 13.1.0 Release 13)

A GLOBAL INITIATIVE

Reference
RTS/TSGS-0426223vd10

Keywords
LTE,UMTS

ETSI

650 Route des Lucioles
F-06921 Sophia Antipolis Cedex - FRANCE

Tel.: +33 4 92 94 42 00 Fax: +33 4 93 65 47 16

Siret N° 348 623 562 00017 - NAF 742 C
Association à but non lucratif enregistrée à la
Sous-Préfecture de Grasse (06) N° 7803/88

Important notice

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

The present document may be made available in electronic versions and/or in print. The content of any electronic and/or print versions of the present document shall not be modified without the prior written authorization of ETSI. In case of any existing or perceived difference in contents between such versions and/or in print, the only prevailing document is the print of the Portable Document Format (PDF) version kept on a specific network drive within ETSI Secretariat.

Users of the present document should be aware that the document may be subject to revision or change of status. Information on the current status of this and other ETSI documents is available at
https://portal.etsi.org/TB/ETSIDeliverableStatus.aspx

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

Intellectual Property Rights

IPRs essential or potentially essential to the present document may have been declared to ETSI. The information pertaining to these essential IPRs, if any, is publicly available for **ETSI members and non-members**, and can be found in ETSI SR 000 314: *"Intellectual Property Rights (IPRs); Essential, or potentially Essential, IPRs notified to ETSI in respect of ETSI standards"*, which is available from the ETSI Secretariat. Latest updates are available on the ETSI Web server (https://ipr.etsi.org/).

Pursuant to the ETSI IPR Policy, no investigation, including IPR searches, has been carried out by ETSI. No guarantee can be given as to the existence of other IPRs not referenced in ETSI SR 000 314 (or the updates on the ETSI Web server) which are, or may be, or may become, essential to the present document.

Foreword

This Technical Specification (TS) has been produced by ETSI 3rd Generation Partnership Project (3GPP).

The present document may refer to technical specifications or reports using their 3GPP identities, UMTS identities or GSM identities. These should be interpreted as being references to the corresponding ETSI deliverables.

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

Modal verbs terminology

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

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

Contents

Foreword

This Technical Specification has been produced by the 3rd Generation Partnership Project (3GPP).

The contents of the present document are subject to continuing work within the TSG and may change following formal TSG approval. Should the TSG modify the contents of the present document, it will be re-released by the TSG with an identifying change of release date and an increase in version number as follows:

Version x.y.z

where:

- x the first digit:
  - 1 presented to TSG for information;
  - 2 presented to TSG for approval;
  - 3 or greater indicates TSG approved document under change control.
- y the second digit is incremented for all changes of substance, i.e. technical enhancements, corrections, updates, etc.
- z the third digit is incremented when editorial only changes have been incorporated in the document.

1 Scope

The present document specifies a client for the IMS-based telepresence service supporting conversational speech, video and text transported over RTP. Telepresence is defined as a conference with interactive audio-visual communications experience between remote locations, where the users enjoy a strong sense of realism and presence between all participants (i.e. as if they are in same location) by optimizing a variety of attributes such as audio and video quality, eye contact, body language, spatial audio, coordinated environments and natural image size. A telepresence system is defined as a set of functions, devices and network elements which are able to capture, deliver, manage and render multiple high quality interactive audio and video signals in a telepresence conference. An appropriate number of devices (e.g. cameras, screens, loudspeakers, microphones, codecs) and environmental characteristics are used to establish telepresence.

The media handling capabilities of a telepresence client (TP UE) are specified in the present document. A TP UE supports Multimedia Telephony Service for IMS (MTSI) UE media handling capabilities [2], but it also supports more advanced media handling capabilities. The media handling aspects of a TP UE within the scope of the present document include media codecs, media configuration and session control, data transport, audio/video parameters, and interworking with MTSI.

2 References

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

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

- For a specific reference, subsequent revisions do not apply.

- For a non-specific reference, the latest version applies. In the case of a reference to a 3GPP document (including a GSM document), a non-specific reference implicitly refers to the latest version of that document *in the same Release as the present document.*

[1] 3GPP TR 21.905: "Vocabulary for 3GPP Specifications".

[2] 3GPP TS 26.114: "IP Multimedia Subsystem (IMS); Multimedia telephony; Media handling and interaction".

[3] 3GPP TS 22.228: "Service requirements for the Internet Protocol (IP) multimedia core network subsystem (IMS); Stage 1".

[4] 3GPP TS 24.229: "IP multimedia call control protocol based on Session Initiation Protocol (SIP) and Session Description Protocol (SDP); Stage 3".

[5] 3GPP TS 24.147: "Conferencing using the IP Multimedia (IM) Core Network (CN) subsystem; Stage 3".

[6] 3GPP TS 24.103: "Telepresence using the IP Multimedia (IM) Core Network (CN) Subsystem (IMS); Stage 3".

[7] draft-ietf-clue-framework-23 (September 2015): "Framework for Telepresence Multi-Streams".

[8] draft-ietf-clue-datachannel-10 (September 2015): "CLUE Protocol Data Channel".

[9] draft-ietf-clue-signaling-06 (August 2015): "CLUE Signaling".

[10] draft-ietf-clue-data-model-schema-10 (June 2015): "An XML Schema for the CLUE data model".

[11] draft-ietf-clue-protocol-06 (October 2015): "CLUE Protocol".

[12] draft-ietf-clue-rtp-mapping-05 (October 2015): "Mapping RTP streams to CLUE media captures".

[13] draft-ietf-mmusic-sctp-sdp-14 (March 2015): "Stream Control Transmission Protocol (SCTP) –Based Media Transport in the Session Description Protocol (SDP)".

[14] draft-ietf-mmusic-data-channel-sdpneg-05 (October 2015): "SDP-based Data Channel Negotiation".

[15] draft-ietf-tsvwg-sctp-dtls-encaps-09 (January 2015): "DTLS Encapsulation of SCTP Packets".

[16] ITU-T Recommendation H.264 (04/2013): "Advanced video coding for generic audiovisual services".

[17] ITU-T Recommendation H.265 (04/2013): "High efficiency video coding".

[18] IETF RFC 6184 (2011): "RTP Payload Format for H.264 Video", Y.-K. Wang, R. Even, T. Kristensen, R. Jesup.

[19] IETF RFC 7798 (2016): "RTP Payload Format for High Efficiency Video Coding (HEVC)", Y.-K. Wang, Y. Sanchez, T. Schierl, S. Wenger, M. M. Hannuksela.

[20] IETF RFC 3264 (2002): "An Offer/Answer Model with the Session Description Protocol (SDP)", J. Rosenberg and H. Schulzrinne.

[21] draft-ietf-mmusic-sdp-simulcast-01 (July 2015): "Using Simulcast in SDP and RTP Session".

[22] Recommendation ITU-T H.241 (02/2012): "Extended video procedures and control signals for ITU-T H.300-series terminals ".

[23] IETF RFC 6236 (2011): "Negotiation of Generic Image Attributes in the Session Description Protocol (SDP) ", I. Johansson and K. Jung.

[24] Recommendation ITU-T H.245 (05/2011): "Control protocol for multimedia communication".

[25] IETF RFC 4566 (2006): "SDP: Session Description Protocol", M. Handley, V. Jacobson and C. Perkins.

[26] IETF RFC 6464: "A Real-time Transport Protocol (RTP) Header Extension for Client-to-Mixer Audio Level Indication".

[27] 3GPP TS 26.441: "Codec for Enhanced Voice Services (EVS); General Overview".

[28] 3GPP TS 26.442: "Codec for Enhanced Voice Services (EVS); ANSI C code (fixed-point)".

[29] 3GPP TS 26.445: "Codec for Enhanced Voice Services (EVS); Detailed Algorithmic Description".

[30] 3GPP TS 26.450: "Codec for Enhanced Voice Services (EVS); Discontinuous Transmission (DTX)".

[31] 3GPP TS 26.171: "Speech codec speech processing functions; Adaptive Multi-Rate - Wideband (AMR-WB) speech codec; General description".

[32] 3GPP TS 26.190: "Speech codec speech processing functions; Adaptive Multi-Rate - Wideband (AMR-WB) speech codec; Transcoding functions".

[33] 3GPP TS 26.173: "ANCI-C code for the Adaptive Multi Rate - Wideband (AMR-WB) speech codec".

[34] 3GPP TS 26.204: "Speech codec speech processing functions; Adaptive Multi-Rate - Wideband (AMR-WB) speech codec; ANSI-C code".

[35] 3GPP TS 26.071: "Mandatory Speech Codec speech processing functions; AMR Speech CODEC; General description".

[36] 3GPP TS 26.090: "Mandatory Speech Codec speech processing functions; Adaptive Multi-Rate (AMR) speech codec; Transcoding functions".

[37] 3GPP TS 26.073: "ANSI C code for the Adaptive Multi Rate (AMR) speech codec".

[38] 3GPP TS 26.104: "ANSI-C code for the floating-point Adaptive Multi Rate (AMR) speech codec".

[39] Recommendation ITU-T F.734 (10/2014): "Definitions, requirements, and use cases for Telepresence Systems".

[40] Recommendation ITU G.1091 (10/2014): "Quality of Experience requirements for telepresence services".

[41] Recommendation ITU-T H.TPS-AV (02/2015): "Audio/video parameters for telepresence systems ".

[42] Recommendation ITU-T H.420 (10/2014): "Telepresence System Architecture".

[43] Recommendation ITU-T H.323 (12/2009): "Packet-based multimedia communication systems".

[44] 3GPP TS 26.093: "Mandatory speech codec speech processing functions; Adaptive Multi-Rate (AMR) speech codec; Source controlled rate operation".

[45] 3GPP TS 26.193: "Speech codec speech processing functions; Adaptive Multi-Rate - Wideband (AMR-WB) speech codec; Source controlled rate operation".

[46] 3GPP TS 26.446: "Codec for Enhanced Voice Services (EVS); AMR-WB Backward Compatible Functions".

[47] 3GPP TS 26.443: "Codec for Enhanced Voice Services (EVS); ANSI C code (floating-point)".

[48] IETF RFC 4574: "The Session Description Protocol (SDP) Label Attribute".

3 Definitions and abbreviations

3.1 Definitions

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

Conference: An IP multimedia session with two or more participants. Each conference has a "conference focus". A conference can be uniquely identified by a user. Examples for a conference could be a Telepresence or a multimedia game, in which the conference focus is located in a game server.

IM session: An IP multimedia (IM) session is a set of multimedia senders and receivers and the data streams flowing from senders to receivers. IP multimedia sessions are supported by the IP multimedia CN Subsystem and are enabled by IP connectivity bearers (e.g. GPRS as a bearer). A user may invoke concurrent IP multimedia sessions.

Telepresence: A conference with interactive audio-visual communications experience between remote locations, where the users enjoy a strong sense of realism and presence between all participants by optimizing a variety of attributes such as audio and video quality, eye contact, body language, spatial audio, coordinated environments and natural image size.

Telepresence System: A set of functions, devices and network elements which are able to capture, deliver, manage and render multiple high quality interactive audio and video signals in a Telepresence conference. An appropriate number of devices (e.g. cameras, screens, loudspeakers, microphones, codecs) and environmental characteristics are used to establish Telepresence.

3.2 Abbreviations

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

AS	Application Server
AVC	Advanced Video Coding
BFCP	Binary Floor Control Protocol
CBP	Constrained Baseline Profile
CHP	Constrained High Profile
CLUE	ControLling mUltiple streams for tElepresence
DTLS	Datagram Transport Layer Security
FIR	Full Intra Request

HEVC	High Efficiency Video Coding
ICE	Interactivity Connectivity Establishment
IDR	Instantaneous Decoding Refresh
IRAP	Intra Random Access Point
MCC	Multiple Content Capture
MIME	Multipurpose Internet Mail Extensions
MRFC	Multimedia Resource Function Controller
MRFP	Multimedia Resource Function Processor
MTSI	Multimedia Telephony Service for IMS
PPS	Picture Parameter Set
RTP	Real-time Transport Protocol
SCTP	Stream Control Transmission Protocol
SDP	Session Description Protocol
SEI	Supplemental Enhancement Information
SPS	Sequence Parameter Set
SSRC	Synchronization Source Identifier
TP	Telepresence
TP UE	TelePresence User Equipment
VPS	Video Parameter Set
XML	EXtensible Markup Language

4 System Description

4.1 Overview

The use cases and requirements on IMS-based telepresence are defined in TS 22.228 [3] to enable telepresence support in IMS applications.

Enabling telepresence support involves updating and enhancing the existing IMS procedures for point-to-point calls as specified in 3GPP TS 24.229 [4] and for multiparty conferences as specified 3GPP TS 24.147 [5]. This has been addressed in 3GPP TS 24.103 [6], which incorporates IETF"s ControLling mUltiple streams for tElepresence (CLUE) framework [7]-[12] with the Session Initiation Protocol (SIP), Session Description Protocol (SDP) and Binary Floor Control Protocol (BFCP) to facilitate controlling multiple spatially related media streams in an IM session supporting telepresence.

In order to provide a "being there" experience for conversational audio and video telepresence sessions between remote locations, where the users enjoy a strong sense of realism and presence, capabilities and preferences need to be co-ordinated and negotiated between local and remote participants such as:

- audio and video spatial composition information; e.g. spatial relationship of two or more objects (audio/video sources) in the same room to allow for accurate reproduction on the receiver side
- capabilities of cameras, screens, microphones and loudspeakers, and their relative spatial relationships
- meeting description, such as view information, language information, participant information, participant type, etc.

CLUE achieves media advertisement and configuration to facilitate controlling and negotiating multiple spatially related media streams in an IMS conference supporting telepresence, taking into account capability information, e.g. screen size, number of screens and cameras, codecs, etc., so that sending system, receiving system, or intermediate system can make decisions about transmitting, selecting, and rendering media streams. With the establishment of the CLUE data channel, the participants have consented to use the CLUE protocol mechanisms to determine the capabilities of the each of the endpoints with respect to multiple streams support, via the exchange of an XML-based data format. The exchange of CLUE messages of each participant's "advertisement" and "configure" is to achieve a common view of media components sent and received in the IM session supporting telepresence.

TS 24.103 [6] specifies procedures to deal with multiple spatially related media streams according to the CLUE framework to support telepresence and to interwork with IM session as below:

1) Initiation of telepresence using IMS, which includes an initial offer/answer exchange establishes a basic media session and a CLUE channel, CLUE exchanges to "advertisement" and "configure" media components used in the session, then followed by an SDP offer/answer in Re-INVITE request to complete the session establishment (see more for the general idea in draft-ietf-clue-framework [7]);

2) Release or leaving of an IM session supporting telepresence, which needs remove the corresponding CLUE channel;

3) Update of an ongoing IM session supporting telepresence, triggered by CLUE exchanges modifying existing CLUE information. For example: a new participant at an endpoint may require the establishment of a specific media stream;

4) Presentation during an IM session supporting telepresence, which may also be initiated by the exchange of CLUE messages and possibly need an updated SDP offer/answer and activation of BFCP for floor control; and

5) Interworking with normal IM session, this is to let the normal IMS users be able to join telepresence using IMS.

4.2 TP UE

A TP UE shall support functional components and user plane protocol stack of an MTSI client as specified in clause 4.2 of TS 26.114 [2].

In addition, a TP UE shall support the protocols required for the data transport channel for the exchange of CLUE messages. This data transport channel runs over DTLS/SCTP (Datagram Transport Layer Security / Stream Control Transmission Protocol) [13] negotiated via the initial SDP offer and answer, and usage of the SDP-based "SCTP over DTLS" data channel negotiation mechanism (see more in draft-ietf-mmusic-data-channel-sdpneg [14]) in order to open the CLUE data channel based on a SCTP stream in each direction. Therefore a TP UE shall support these protocols over the user plane, while the MTSI client protocol stack depicted Figure 4.3 of TS 26.114 currently lacks these capabilities.

The non-media data conveyed over a data channel is handled by using SCTP encapsulated in DTLS. More information about how SCTP is used on the top of the DTLS protocol can be found in draft-ietf-tsvwg-sctp-dtls-encaps [15]. Furthermore, using DTLS over UDP in combination with Interactivity Connectivity Establishment (ICE) enables middle box traversal in IPv4 and IPv6 based networks. This data transport service operates in parallel to the RTP media transport, and all of them can be eventually share a single transport-layer port number.

The layering of protocols for data channel is shown in the following Figure 4.2.1.

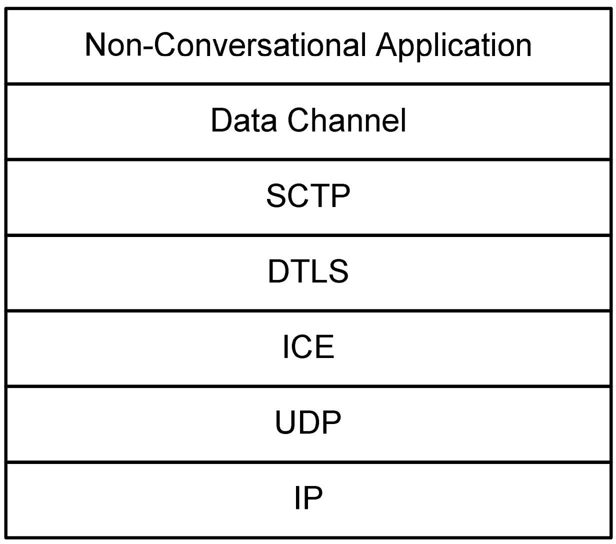

Figure 4.2.1: Protocol stack of data channel for TP UE

A TP UE offers a DTLS/SCTP association together with the media format indicating the use of a data channel in the first SDP offer or subsequent SDP offers. A TP UE can further open the data channel via the SDP-based "SCTP over DTLS" data channel negotiation mechanism to indicate specific non-conversational application (e.g. CLUE protocol) over it.

5 Media Codecs

5.1 Speech

TP clients in terminals offering speech communication shall support super-wideband and full-band audio as per below.

TP clients in terminals offering speech communication shall support,

- super-wideband and full-band through EVS codec (3GPP TS 26.441 [27], 3GPP TS 26.442 [28], 3GPP TS 26.445 [29] and 3GPP TS 26.443 [47]) including functions for discontinuous transmission (3GPP TS 26.450 [30]).

To support transcoder-free interworking with MTSI clients, a TP UE shall additionally offer lower bandwidth (e.g. NB, WB) speech communication as per below.

TP clients in terminals offering speech communication shall support,

- wideband through EVS AMR-WB IO mode (3GPP TS 26.446 [46]) or AMR-WB codec (3GPP TS 26.171 [31], 3GPP TS 26.190 [32], 3GPP TS 26.173 [33] and 3GPP TS 26.204 [34]) including all 9 modes and source controlled rate operation 3GPP TS 26.193 [45].

- narrowband through AMR speech codec (3GPP TS 26.071 [35], 3GPP TS 26.090 [36], 3GPP TS 26.073 [37] and 3GPP TS 26.104 [38]) including all 8 modes and source controlled rate operation 3GPP TS 26.093 [44].

5.2 Video

TP clients in terminals offering video communication shall support:

- H.264 (AVC) [16] Constrained High Profile (CHP), Level 3.1

To support transcoder-free interworking with MTSI clients, TP clients in terminals offering video communication shall also support:

- H.264 (AVC) [16] Constrained Baseline Profile (CBP), Level 1.2

In addition, TP UEs should support:

- H.265 (HEVC) [17] Main Profile, Main Tier, Level 4.1

H.264 (AVC) CHP shall be used, without requirements on output timing conformance (annex C of [16]). Each sequence parameter set of H.264 (AVC) shall contain the vui_parameters syntax structure including the num_reorder_frames syntax element set equal to 0.

H.265 (HEVC) Main Profile shall be used with general_progressive_source_flag equal to 1, general_interlaced_source_flag equal to 0, general_non_packed_constraint_flag equal to 1, general_frame_only_constraint_flag equal to 1, and sps_max_num_reorder_pics[i] equal to 0 for all i in the range of 0 to sps_max_sub_layers_minus1, inclusive, without requirements on output timing conformance (annex C of [17]).

For both H.264 (AVC) and H.265 (HEVC), the decoder needs to know the Sequence Parameter Set (SPS) and the Picture Parameter Set (PPS) to be able to decode the received video packets. A compliant H.265 (HEVC) bitstream includes a Video Parameter Set (VPS), although the VPS may be ignored by the decoder in the context of the present specification. When H.264 (AVC) or H.265 (HEVC) is used it is recommended to transmit the parameter sets within the SDP description of a stream, using the relevant MIME/SDP parameters as defined in RFC 6184 [18] for H.264 (AVC) and in [19] for H.265 (HEVC), respectively. Each media source (SSRC) shall transmit the currently used parameter sets at least once in the beginning of the RTP stream before being referenced by the encoded video data to ensure that the parameter sets are available when needed by the receiver. If the video encoding is changed during an ongoing session such that the previously used parameter set(s) are no longer sufficient then the new parameter sets shall be transmitted at least once in the RTP stream prior to being referenced by the encoded video data to ensure that the parameter sets are available when needed by the receiver. When a specific version of a parameter set is sent in the RTP stream for the first time, it should be repeated at least 3 times in separate RTP packets with a single copy per RTP packet and with an interval not exceeding 0,5 seconds to reduce the impact of packet loss. A single copy of the currently active parameter sets shall also be part of the data sent in the RTP stream as a response to FIR. Moreover, it is recommended to avoid using a sequence or picture parameter set identifier value during the same session to signal two or more parameter sets of the same type having different values, such that if a parameter set identifier for a certain type is used more than once in either SDP description or RTP stream, or both, the identifier always indicates the same set of parameter values of that type.

The video decoder in a multimedia TP client in terminal shall either start decoding immediately when it receives data, even if the stream does not start with an IDR/IRAP access unit (IDR access unit for H.264, IRAP access unit for H.265) or alternatively no later than it receives the next IDR/IRAP access unit or the next recovery point Supplemental Enhancement Information (SEI) message, whichever is earlier in decoding order. The decoding process for a stream not starting with an IDR/IRAP access unit shall be the same as for a valid video bit stream. However, the TP client in terminal shall be aware that such a stream may contain references to pictures not available in the decoded picture buffer. The display behaviour of the TP client in terminal is out of scope of the present document.

A TP client in terminal offering video support other than H.264 CHP Level 3.1 shall also offer H.264 CHP Level 3.1.

A TP UE client in terminal offering H.265 (HEVC) shall support negotiation to use a lower Level than the one in the offer, as described in [19] and [20].

To support interworking with MTSI clients, a TP UE shall offer both H.264 CBP Level 1.2 and H.264 CHP Level 3.1 (with preference for the latter) and should also offer H.264 CBP Level 3.1.

In case a codec profile is offered with a Level higher than the required Level, no separate offer for the required Level is needed.

A TP client in terminal offering H.264 (AVC) CHP support at a level higher than Level 3.1 shall support negotiation to use a lower Level as described in [18] and [20].

A TP client in terminal offering H.264 (AVC) CBP support at a level higher than Level 1.2 shall support negotiation to use a lower Level as described in [18] and [20].

A TP client in terminal offering H.264 (AVC) CBP support at a level higher than Level 3.1 shall support negotiation to use a lower Level as described in [18] and [20].

NOTE 1: All levels are minimum requirements. Higher levels may be supported and used for negotiation.

NOTE 2: TP clients in terminals may use full-frame freeze and full-frame freeze release SEI messages of H.264 (AVC) to control the display process. For H.265 (HEVC), MTSI clients may set the value of pic_output_flag in the slice segment headers to either 0 or 1 to control the display process.

NOTE 3: An H.264 (AVC) encoder should code redundant slices only if it knows that the far-end decoder makes use of this feature (which is signalled with the redundant-pic-cap MIME/SDP parameter as specified in RFC 6184 [18]). H.264 (AVC) encoders should also pay attention to the potential implications on end-to-end delay. The redundant slice header is not supported in H.265 (HEVC).

NOTE 4: If a codec is supported at a certain level, it implies that on the receiving side, the decoder is required to support the decoding of bitstreams up to the maximum capability of this level. On the sending side, the support of a particular level does not imply that the encoder will produce a bitstream up to the maximum capability of the level. This method can be used to set up an asymmetric video stream. For H.264 (AVC), another method is to use the SDP parameters "level-asymmetry-allowed" and "max-recv-level" that are defined in the H.264 payload format specification, [18]. For H.265 (HEVC) it is possible to use the SDP parameter "max-recv-level-id" defined in the H.265 payload format specification, [19], to indicate a higher level in the receiving direction than in the sending direction.

NOTE 5: For H.264 (AVC) and H.265 (HEVC), respectively, multiple sequence and picture parameter sets can be defined, as long as they have unique parameter set identifiers, but only one sequence and picture parameter set can be active between two consecutive IDRs and IRAPs, respectively.

5.3 Real-time Text

The real-time text requirements for MTSI clients in terminals specified in TS 26.114 [2], clause 5.2.3, also apply for TP UEs.

6 Media Configuration

The media configuration requirements for MTSI clients in terminals specified in TS 26.114 [2], clause 6, also apply for TP UEs.

To enable devices to participate in an IMS-based telepresence call, selecting the sources they wish to view, receiving those media sources and displaying them in an optimal fashion, CLUE involves two principal and inter-related protocol negotiations. SDP, conveyed via SIP, is used to negotiate the specific media capabilities that can be delivered to specific addresses on the TP UE. Meanwhile, a CLUE protocol [11], transported via a CLUE data channel [8], is used to negotiate the Capture Sources available, their attributes and any constraints in their use, along with which Captures the far end provides a TP UE wishes to receive.

The CLUE data channel [8] is a bidirectional SCTP over DTLS channel used for the transport of CLUE messages. This channel shall be established before CLUE protocol messages can be exchanged and CLUE-controlled media can be sent.

Beyond negotiating the CLUE channel, SDP is also used to negotiate the details of supported media streams and the maximum capability of each of those streams. As the CLUE Framework [7] defines a manner in which the Media Provider expresses their maximum encoding capabilities, SDP is also used to express the encoding limits for each potential Encoding.

Backwards-compatibility with MTSI is an important consideration and it is vital that a CLUE-capable TP UE contacting a terminal that does not support CLUE is able to fall back to a fully functional non-CLUE call governed by the requirements on MTSI in 3GPP TS 26.114 [2].

CLUE support shall be offered in the first SDP offer, as follows. At the beginning of a CLUE-based telepresence session over IMS, the support for CLUE shall be negotiated via the SDP between two TP UEs. The CLUE extension shall be indicated using an SDP session-level 'group' attribute. Each SDP media "m" line that is included in this group, using SDP media-level "mid" attributes, is CLUE-controlled, by a CLUE data channel also included in this CLUE group. A CLUE group should include the "mid" of the "m" line for one (and only one) CLUE data channel. For interoperability with non-CLUE devices, a CLUE-capable device sending an initial SDP offer shall not include any "m" line for CLUE-controlled media beyond the "m" line for the CLUE data channel (this is unless the remote terminal"s CLUE support was already indicated at the SIP level using the "sip.clue" media feature tag), and includes at least one non-CLUE-controlled media "m" line.

For audio and video, the first SDP offer shall also contain the basic (i.e. non-CLUE-controlled) media streams with the set of mandatory codecs for TP UEs, i.e. namely EVS-SWB and H.264/AVC CHP Level 3.1. For each of the offered basic (i.e. non-CLUE-controlled) media streams indicated in an "m=" line, one mandatory codec of the same media type that is specified in 3GPP TS 26.114 [2] shall also be included toward ensuring interworking with MTSI terminals. The preference in the codec order should favour the mandatory codecs for TP UEs, i.e. namely EVS-SWB and H.264/AVC CHP Level 3.1.

If the CLUE negotiation is successful and the remote terminal is also a CLUE-capable TP UE, then the subsequent offers for all media streams, including basic streams and CLUE-controlled media, shall contain the mandatory codecs for TP UEs, namely EVS-SWB and H.264/AVC CHP Level 3.1, and shall contain at least one RTP payload type of the corresponding codec for each media line. Accordingly, the SDP answers from TP UEs shall also accept these codecs and contain the corresponding RTP payload types, and also shall conform to the requirements established in Tables 6.3a and 6.3b of 3GPP TS 26.114 [2].

A TP UE receiving an SDP offer from an MTSI UE or a non-3GPP access client that does not support CLUE capabilities shall fall back to operate as an MTSI client and answer the SDP offer as per the requirements established in 3GPP TS 26.114 [2].

TP UEs may be involved in media sessions where CLUE could be enabled or disabled during an ongoing call. If, in an ongoing non-CLUE call, an SDP offer/answer exchange completes with both sides having included a data channel "m" line in their SDP and with the "mid" for that channel in corresponding CLUE groups then the call is now CLUE-enabled; negotiation of the data channel and subsequently the CLUE protocol begin. If, in an ongoing CLUE-enabled call, an SDP offer-answer negotiation completes in a fashion in which either the CLUE data channel was not successfully negotiated or one side did not include the data channel in a matching CLUE group then CLUE for this channel is disabled. In the event that this occurs, CLUE is no longer enabled and sending of all CLUE-controlled media associated with the corresponding CLUE group shall stop.

A TP UE offering a media session should generate an SDP offer according to the examples in Annex A.

7 Data Transport

7.1 Introduction

The data transport requirements for MTSI clients in terminals specified in TS 26.114 [2], clause 7, also apply for TP UEs.

7.2 RTP Payload Formats for TP UEs

The requirements on RTP payload formats for MTSI clients as specified in clause 7.4 of TS 26.114 [2] also apply for TP UEs.

NOTE: Further requirements on data transport aspects with regards to the usage of RTP / RTCP protocols, e.g. the use of RTP multiplexing and mapping of RTP streams to CLUE media captures, are FFS.

8 Audio/Video Parameters

8.1 Overview

The audio/video parameters provided in clauses 8.2 and 8.3 should be supported by TP UEs as part of CLUE-based signaling in IMS-based telepresence sessions both at session initiation and during a session.

Collectively, these audio/video parameters and their associated values can be expected to provide a high quality telepresence experience for 3GPP's IMS-based telepresence services from a media handling point of view.

Clause 8.2 describes the set of the capture-related audio/video parameters for 3GPP IMS-based telepresence services, while clause 8.3 describes the audio/video parameters on the telepresence system environment. Furthermore, guidance is provided in these clauses on the need for signalling these parameters at session initiation and during a session. While most of the parameters are already part of the CLUE framework, some of them are not and further references on the suitable signalling options for such parameters are also described. Some of these parameters are not signalled neither during session initiation nor during a session, but are still recommended to be supported in TP UEs for the purposes of quality monitoring.

8.2 Capture-Related Parameters

8.2.1 General Parameters

Table 8.2.1.1: General parameters

Parameter	Need for signalling at session initiation	Need for signalling during session	Remarks
mediaType	Y	Y	See the "MediaCapture" attributes in IETF CLUE data model schema [10].
captureScene description	Y	Y	See the "Description" attribute in clause 7.3.1 of IETF CLUE framework [7] and clause 10.24 of IETF CLUE data model schema [10].
sceneView description	Y	Y	See the "View" attribute in clause 7.3.2 of IETF CLUE framework [7] and clause 10.25 of IETF CLUE data model schema [10].
Lang	Y	N	See "Language" in IETF CLUE framework [7] and IETF CLUE data model schema [10].
Priority	Y	Y	See "Priority" in IETF CLUE framework [7] and IETF CLUE data model schema [10].
Embeddedtext	Y	Y	See "Embedded Text" in IETF CLUE framework [7] and IETF CLUE data model schema [10].
relatedTo	Y	Y	See "Related to" in IETF CLUE framework [7] and IETF CLUE data model schema [10].
Presentation	Y	Y	See "Presentation" in IETF CLUE framework [7] and IETF CLUE data model schema [10].
personInfo	Y	Y	As per clause 7.1.1.10 in IETF CLUE framework [7] and clause 10.29 of IETF CLUE data model schema [10].
personType	Y	Y	As per clause 7.1.1.11 in IETF CLUE framework [7] and clause 10.29 of IETF CLUE data model schema [10].
sceneInformation	Y	Y	As per clause 7.3.1.1 in IETF CLUE framework [7] and clause 10.24 of IETF CLUE data model schema [10].
mediaCapture description	Y	Y	See the "Description" attribute in clause 7.1.1 of IETF CLUE framework [7] and clause 10.13 of IETF CLUE data model schema [10].
captureScene scale	Y	N	See "Scale information" in clause 7.3.1 of IETF CLUE framework [7] and clause 10.24 of IETF CLUE data model schema [10].
mediaCapture mobility	Y	N	See "Mobility of capture" in clause 7.1.1.4 of IETF CLUE framework [7] and clause 10.16 of IETF CLUE data model schema [10].
mediaCapture view	Y	Y	See "View" in clause 7.1.1.8 of IETF CLUE framework [7] and clause 10.18 of IETF CLUE data model schema [10].
maxGroupBandwidth	Y	N	See "maxGroupBandwidth" in IETF CLUE framework [7] and IETF CLUE data model schema [10].
Simulcast	Y	Y	Telepresence systems may provide multiple encodings for the one capture through a technique known as simulcast. For example, this may be achieved by sending multiple video coding streams with different characteristics to allow a receiving endpoint to choose the stream that meets its needs. Mechanisms for accomplishing simulcast in RTP and how to signal it in SDP are provided in [21].

8.2.2 Visual Parameters

Table 8.2.2.1: Visual parameters

Parameter	Need for signalling at session initiation	Need for signalling during session	Remarks
colorGamut	Y	N	This parameter indicates the Colour Gamut used in a Telepresence Video Stream. Signalled as part of the codec information, e.g. in H.264 and H.265 SEI [16]-[17].
lumaBitDepth	Y	N	This parameter indicates the bit depth of the luma samples in a digital picture. Signalled as part of the codec information, e.g. in H.264 and H.265 SEI [16]-[17].
chromaBitDepth	Y	N	This parameter indicates the bit depth of the chroma samples in a digital picture. Signalled as part of the codec information, e.g. in H.264 and H.265 SEI [16]-[17].
effectiveResolution	N	N	This parameter indicates effective resolution of a rendered video stream as perceived by the viewer, as defined by ITU-T H.TPS-AV [41]. Not signalled.
captureArea	Y	Y	See "Area of Capture" in clause 7.1.1.3 of IETF CLUE framework [7] and clause 10.5.2 of IETF CLUE data model schema [10].
capturePoint	Y	Y	See "Point of Capture" and "captureOrigin" in clause 7.1.1.1 in IETF CLUE framework [7] and clause 10.5.1 of IETF CLUE data model schema [10].
lineOfCapturePoint	Y	Y	See the "Point on line of Capture" attribute in clause 7.1.1.2 of IETF CLUE framework [7] and clause 10.5.1 of IETF CLUE data model schema [10].
maxVideoBitrate	Y	Y	This parameter indicates the maximum number of bits per second relating to a single video encoding and is signalled in the SDP. See "max-mbps" in IETF RFC 6184 [18] and "CustomMaxMBPS" in ITU-T H.241 [22].
maxWidth	Y	N	This parameter indicates the maximum video resolution width in pixels and is signalled in the SDP. See "horizontal image size" in IETF RFC 6236 [23] and "CustomPictureFormat" in ITU-T H.245 [24].
maxHeight	Y	N	This parameter indicates the maximum video resolution height in pixels and is signalled in the SDP. See "vertical image size" in IETF RFC 6236 [23] and "CustomPictureFormat" in ITU-T H.245 [24].
maxFramerate	Y	N	This parameter indicates the maximum video framerate and is signalled in the SDP. See "framerate" in IETF RFC 4566 [25] and "MaxFPS" in ITU-T H.241 [22].

8.2.3 Audio Parameters

Table 8.2.3.1: Audio parameters

Parameter	Need for signalling at session initiation	Need for signalling during session	Remarks
Audio capturePoint	Y	Y	See "Point of Capture" and "captureOrigin" in clause 7.1.1.1 in IETF CLUE framework [7] and clause 10.5.1 of IETF CLUE data model schema [10].
Audio lineOfCapturePoint	Y	Y	See the "Point on line of Capture" attribute in clause 7.1.1.2 of IETF CLUE framework [7] and clause 10.5.1 of IETF CLUE data model schema [10].
Audio sensitivityPattern	Y	Y	See the "Audio Capture Sensitivity Pattern" attribute in clause 7.1.1.5 of IETF CLUE framework [7] and clause 10.20.1 of IETF CLUE data model schema [10].
maxAudioBitrate	Y	Y	This parameter indicates the maximum number of bits per second relating to a single audio encoding and signalled in the SDP. See "bandwidth" in IETF RFC 4566 [25] and "maxBitRate" in ITU-T H.245 [24].
nominalAudio Level	Y	Y	This parameter indicates the nominal audio level sent in the Telepresence audio stream. See ITU-T H.245 [24] and clause 7.1.3.3 of ITU-T H.TPS-AV [41].
dynamicAudioLevel	N	Y	This parameter indicates the actual audio level sent in the Telepresence audio stream as it varies as a function of time, and may be signalled in the RTP header extension. See IETF RFC 6464 [26] and clause 7.1.3.4 of ITU-T H.TPS-AV [41].

8.2.4 Delay Parameters

Table 8.2.4.1: Delay parameters

Parameter	Need for signalling at session initiation	Need for signalling during session	Remarks
endToEndVideoDelay	N	N	This parameter indicates the one-way end to end delay (camera lens to video display) of the video media sent between two Telepresence terminals. In order to provide a high QoE telepresence experience to end-users, telepresence systems, it is desirable for the end to end video delay to be less than 320 milliseconds [39]-[41]. Not signalled.
endToEndAudioDelay	N	N	This parameter indicates the one-way end to end delay (mouth to ear) of the audio media sent between two Telepresence terminals. In order to provide a high QoE telepresence experience to end-users, telepresence systems, it is desirable for the end to end audio delay to be less than 280 milliseconds [39]-[41]. Not signalled.
audioVideoSynchronization	N	N	This parameter indicates the synchronization between an audio and the corresponding video media stream (EndtoEndVideoDelay-EndtoEndAudioDelay). In order to provide high QoE telepresence services to end-users, telepresence systems should maintain synchronization within 40 and -60 milliseconds (i.e. synchronization error is less than 40 ms if the audio stream is ahead of the video stream and less than 60 ms if the video stream is ahead of the audio stream) [39]-[41]. Not signalled.

NOTE: Delay numbers are based on ITU-T references [39]-[41] and 3GPP-specific modifications are FFS.

8.2.5 Multiple Source Capture Parameters

Table 8.2.5.1: Multiple Source Capture parameters

Parameter	Need for signalling at session initiation	Need for signalling during session	Remarks
multiContentCapture	Y	Y	See the "Multiple content capture" in clause 7.2 of IETF CLUE framework [7].
MCC sources	Y	Y	See the "Multiple content capture" in clause 7.2 of IETF CLUE framework [7].
MCC maxCaptures	Y	Y	See the "Maximum Number of Captures within a MCC" MCC attribute in clause 7.2 of IETF CLUE framework [7].
MCC policy	Y	Y	See the "Policy" MCC attribute in clause 7.2 of IETF CLUE framework [7].
MCC synchronizationID	Y	Y	See the "Synchronization Identity" MCC attribute in clause 7.2 of IETF CLUE framework [7].
multiContentCapture	Y	Y	See the "Multiple content capture" in clause 7.2 of IETF CLUE framework [7].
MCC sources	Y	Y	See the "Multiple content capture" in clause 7.2 of IETF CLUE framework [7].
MCC maxCaptures	Y	Y	See the "Maximum Number of Captures within a MCC" MCC attribute in clause 7.2 of IETF CLUE framework [7].
MCC policy	Y	Y	See the "Policy" MCC attribute in clause 7.2 of IETF CLUE framework [7].
MCC synchronizationID	Y	Y	See the "Synchronization Identity" MCC attribute in clause 7.2 of IETF CLUE framework [7].

8.3 Telepresence System Environment Parameters

Table 8.3.1: Telepresence System Environment parameters

Parameter	Need for signalling at session initiation	Need for signalling during session	Remarks
illuminantType	Y	Y	This parameter describes the profile of the visible light at a telepresence endpoint. May need to be signalled if lighting changes during session. Signalling is based on Annex E of ITU-T H.264 [16] and Annex E of ITU-T H.265 [17].
illuminantCRI Index	Y	Y	This parameter describes the colour rendering index (CRI) of the visible (ambient) light at the telepresence endpoint. Signalling is based on Annex E of ITU-T H.264 [16] and Annex E of ITU-T H.265 [17].
illuminantColourTemperature	Y	Y	This parameter describes the correlated colour temperature (CCT) of the visible (ambient) light at the telepresence endpoint. Signalling is based on Annex E of ITU-T H.264 [16] and Annex E of ITU-T H.265 [17].

9 Interworking

The requirements to enable transcoder-free interworking between TP UEs and MTSI UEs are specified in clause 6.

The interworking requirements for MTSI clients in terminals specified in TS 26.114 [2], clause 12, also apply for TP UEs, for both the basic media streams (non-CLUE-controlled) as well as the CLUE-controlled media streams, latter being applicable when the remote terminal is also CLUE-capable.

When the TP UE falls back to MTSI UE (e.g. upon failure of CLUE capability negotiation during the initial SDP offer-answer exchange), TP UE shall conform to the interworking requirements established for MTSI UEs in clause 12 of TS 26.114 [2].

The interworking requirements for MTSI clients in terminals using fixed access specified in TS 26.114 [2], clause 18, also apply for TP UEs in terminals using fixed access, for both the basic media streams (non-CLUE-controlled) as well as the CLUE-controlled media streams, latter being applicable when the remote terminal is also CLUE-capable.

If the TP UE is in terminal using fixed access, and it falls back to MTSI UE using fixed access (e.g. upon failure of CLUE capability negotiation during the initial SDP offer-answer exchange), TP UE shall conform to the interworking requirements established for MTSI UEs in clause 18 of TS 26.114 [2].

Interworking with non-3GPP telepresence systems, e.g. those based on ITU-T Telepresence [39]-[43], [22], [24] is FFS.

10 Jitter Buffer Management

The jitter buffer management requirements for MTSI clients in terminals specified in TS 26.114 [2], clause 8, also apply for TP UEs.

Annex A (informative): SDP and CLUE Examples

A.1 TP Session Setup with CLUE-Controlled Video Support

The following example demonstrates the SDP offer for negotiation of a CLUE data channel – the assumption here is that TP UE1 has three cameras and three screens and TP UE2 has two cameras and two screens.

Table A.1.1: Example SDP offer for Establishment of CLUE Data Channel

SDP offer
a=group CLUE 3
m=audio 49152 RTP/AVP 96 97 98 99 100
a=tcap:1 RTP/AVPF
a=pcfg:1 t=1
b=AS:89
b=RS:0
b=RR:4000
a=rtpmap:96 EVS/16000/1
a=fmtp:96 br=13.2-64; bw=swb; max-red=220
a=rtpmap:97 AMR-WB/16000/1
a=fmtp:97 mode-change-capability=2; max-red=220
a=rtpmap:98 AMR-WB/16000/1
a=fmtp:98 mode-change-capability=2; max-red=220; octet-align=1
a=rtpmap:99 AMR/8000/1
a=fmtp:99 mode-change-capability=2; max-red=220
a=rtpmap:100 AMR/8000/1
a=fmtp:100 mode-change-capability=2; max-red=220; octet-align=1
a=ptime:20
a=maxptime:240
a=sendrecv
a=mid:1
m=video 49154 RTP/AVP 99 100
a=tcap:1 RTP/AVPF
a=pcfg:1 t=1
b=AS:1060
b=RS:0
b=RR:5000
a=rtpmap:99 H264/90000
a=fmtp:99 packetization-mode=0; profile-level-id=640c1f; \ sprop-parameter-sets=Z2QMH5WgFAFugH9Q,aM46gA==
a=rtpmap:100 H264/90000
a=fmtp:100 packetization-mode=0; profile-level-id=42e00c; \ sprop-parameter-sets=J0LgDJWgUH6Af1A=,KM46gA==
a=rtcp-fb:* trr-int 5000
a=rtcp-fb:* nack
a=rtcp-fb:* nack pli
a=rtcp-fb:* ccm fir
a=rtcp-fb:* ccm tmmbr
a=sendrecv
a=mid:2

```
m=application 6100 UDP/DTLS/SCTP webrtc-datachannel
a=sctp-port: 5000
a=dcmap:2 subprotocol="CLUE"; ordered=true
a=mid:3
```

In the above SDP example and remaining SDP examples below, boldface font is used to highlight the key lines demonstrating the described SDP offer answer procedures in the context of TP sessions controlled by CLUE.

The SDP offer from TP UE1 in Table A.1.1 contains basic media streams (non-CLUE controlled) and an establishment request for a DTLS/SCTP association used to realize a CLUE data channel. A CLUE group is included and the data channel is shown in group (3).

The initial SDP offer message negotiates the port and transport information for setting up the DTLS/SCTP association, via a separate SDP "m=" line with a UDP/DTLS/SCTP or TCP/DTLS/SCTP proto value, together with an SDP "sctp-port" attribute, and an SDP "dcmap" attribute to indicate "CLUE" as the application protocol running over the data channel. The procedures for establishment of the DTLS/SCTP association via SDP can be found in [13] and [8].

For the basic media streams, the offer contains the AMR narrowband and AMR-WB wideband codecs for audio and H.264/AVC Constrained Baseline Profile Level 1.2 (in addition to the mandatory codecs for TP UEs, namely EVS-SWB and H.264/AVC CHP Level 3.1).

Table A.1.2: Example SDP answer for Establishment of CLUE Data Channel

SDP answer

```
a=group CLUE 100
m=audio 49152 RTP/AVPF 96
a=acfg:1 t=1
b=AS:89
b=RS:0
b=RR:4000
a=rtpmap:96 EVS/16000/1
a=fmtp:96 br=13.2-64; bw=swb; max-red=220
a=ptime:20
a=maxptime:240
a=sendrecv
a=mid:9
m=video 49154 RTP/AVPF 99
a=acfg:1 t=1
b=AS:1060
b=RS:0
b=RR:5000
a=rtpmap:99 H264/90000
a=fmtp:99 packetization-mode=0; profile-level-id=640c1f; \
     sprop-parameter-sets=Z2QMH5WgFAFugH9Q,aM46gA==
a=rtcp-fb:* trr-int 5000
a=rtcp-fb:* nack
a=rtcp-fb:* nack pli
a=rtcp-fb:* ccm fir
a=rtcp-fb:* ccm tmmbr
a=sendrecv
a=mid:10
m=application 6100 UDP/DTLS/SCTP webrtc-datachannel
a=sctp-port: 5000
a=dcmap:2 subprotocol="CLUE"; ordered=true
a=mid:100
```

The answer from TP UE2 in Table A.1.2 indicates that the CLUE data channel is accepted. Now CLUE and configure messages can be exchanged in order to negotiate on the capture and rendering capabilities for the telepresence session. Moreover, TP UE2 wishes to receive initial media, and so includes corresponding non-CLUE-controlled audio and video lines.

With the successful initial SDP offer-answer, TP UEs negotiate the CLUE channel via the exchange of CLUE ADVERTISEMENT and CLUE CONFIGURE messages. TP UE1 advertises three static Captures representing its three cameras. It also includes switched Captures suitable for two- and one-screen systems. TP UE1 also includes an Encoding Group with three Encoding IDs: "enc1", "enc2" and "enc3", which will appear in the subsequent SDP offer TP UE1 intends to send. TP UE2 also sends its CLUE ADVERTISEMENT message, where it advertises two static Captures representing its two cameras. It also includes a single composed Capture for single-screen systems, in which it will composite the two camera views into a single video stream. TP UE2 also includes a single Encoding Group with two Encoding IDs: "foo" and "bar".

Following the exchange of these CLUE messages, further SDP offer-answer negotiations can occur that include CLUE-controlled encodings. Every "m" line representing a CLUE Encoding contains a "label" attribute as defined in [48] to identify the Encoding by the sender in CLUE Advertisement messages and by the receiver in CLUE Configure messages, e.g. "enc1", "enc2", "enc3" for TP UE1 and "foo" and "bar" for TP UE2.

Table A.1.3: Example SDP offer for Negotiating CLUE-controlled Media

SDP offer
a=group CLUE 3 4 5 6
m=audio 49152 RTP/AVPF 96
b=AS:89
b=RS:0
b=RR:4000
a=rtpmap:96 EVS/16000/1
a=fmtp:96 br=13.2-64; bw=swb; max-red=220
a=ptime:20
a=maxptime:240
a=sendrecv
a=mid:1
m=video 49154 RTP/AVPF 99
b=AS:1060
b=RS:0
b=RR:5000
a=rtpmap:99 H264/90000
a=fmtp:99 packetization-mode=0; profile-level-id=640c1f; \   sprop-parameter-sets=Z2QMH5WgFAFugH9Q,aM46gA==
a=rtcp-fb:* trr-int 5000
a=rtcp-fb:* nack
a=rtcp-fb:* nack pli
a=rtcp-fb:* ccm fir
a=rtcp-fb:* ccm tmmbr
a=sendrecv
a=mid:2
m=application 6100 UDP/DTLS/SCTP webrtc-datachannel
a=sctp-port: 5000
a=dcmap:2 subprotocol="CLUE"; ordered=true
a=mid:3
m=video 49156 RTP/AVP 99
a=tcap:1 RTP/AVPF
a=pcfg:1 t=1
b=AS:1060
b=RS:0
b=RR:5000
a=rtpmap:99 H264/90000
a=fmtp:99 packetization-mode=0; profile-level-id=640c1f; \

```
     sprop-parameter-sets=Z2QMH5WgFAFugH9Q,aM46gA==
a=rtcp-fb:* trr-int 5000
a=rtcp-fb:* nack
a=rtcp-fb:* nack pli
a=rtcp-fb:* ccm fir
a=rtcp-fb:* ccm tmmbr
a=sendonly
a=mid:4
a=label:enc1
m=video 49158 RTP/AVP 99
a=tcap:1 RTP/AVPF
a=pcfg:1 t=1
b=AS:1060
b=RS:0
b=RR:5000
a=rtpmap:99 H264/90000
a=fmtp:99 packetization-mode=0; profile-level-id=640c1f; \
     sprop-parameter-sets=Z2QMH5WgFAFugH9Q,aM46gA==
a=rtcp-fb:* trr-int 5000
a=rtcp-fb:* nack
a=rtcp-fb:* nack pli
a=rtcp-fb:* ccm fir
a=rtcp-fb:* ccm tmmbr
a=sendonly
a=mid:5
a=label:enc2
m=video 49160 RTP/AVP 99
a=tcap:1 RTP/AVPF
a=pcfg:1 t=1
b=AS:1060
b=RS:0
b=RR:5000
a=rtpmap:99 H264/90000
a=fmtp:99 packetization-mode=0; profile-level-id=640c1f; \
     sprop-parameter-sets=Z2QMH5WgFAFugH9Q,aM46gA==
a=rtcp-fb:* trr-int 5000
a=rtcp-fb:* nack
a=rtcp-fb:* nack pli
a=rtcp-fb:* ccm fir
a=rtcp-fb:* ccm tmmbr
a=sendonly
a=mid:6
a=label:enc3
```

The second SDP offer in Table A.1.3 from TP UE1 maintains the "sendrecv" audio, video and includes three additional "sendonly" m-lines representing the three CLUE-controlled encodings for video. Each of these m-lines contains a "label" corresponding to one of the encoding IDs from the CLUE advertisement from TP UE1. Each also has its "mid" added to the grouping attribute to show that they are controlled by the CLUE channel.

Since it is now clear that the remote endpoint is a CLUE-capable TP UE, the offer for the basic streams contains the mandatory audio and video codecs for TP UEs, namely the EVS codec as well as the H.264/AVC Constrained High Profile Level 3.1.

TP UE2 now has all the information he needs to decide which streams to configure. As such he now sends its CLUE CONFIGURE message. This requests the pair of switched Captures that represent TP UE1's scene, and it configures them with encoder ids "enc1" and "enc2".

TP UE1 receives the CLUE CONFIGURE from TP UE2 and sends a CLUE RESPONSE message to acknowledge its receptions.

Table A.1.4: Example SDP answer for Negotiating CLUE-controlled Media

SDP answer
a=group CLUE 11 12 100
m=audio 49152 RTP/AVPF 96
b=AS:89
b=RS:0
b=RR:4000
a=rtpmap:96 EVS/16000/1
a=fmtp:96 br=13.2-64; bw=swb; max-red=220
a=ptime:20
a=maxptime:240
a=sendrecv
a=mid:9
m=video 49154 RTP/AVPF 99
b=AS:1060
b=RS:0 b=RR:2500
a=rtpmap:99 H264/90000
a=fmtp:99 packetization-mode=0; profile-level-id=640c1f; \\ sprop-parameter-sets=Z2QMH5WgFAFugH9Q,aM46gA==
a=rtcp-fb:* trr-int 5000
a=rtcp-fb:* nack
a=rtcp-fb:* nack pli
a=rtcp-fb:* ccm fir
a=rtcp-fb:* ccm tmmbr
a=sendrecv **a=mid:10**
m=application 6100 UDP/DTLS/SCTP webrtc-datachannel
a=sctp-port: 5000
a=dcmap:2 subprotocol="CLUE"; ordered=true
a=mid:100
m=video 49156 RTP/AVPF 99
a=acfg:1 t=1
b=AS:1060
b=RS:0 b=RR:5000
a=rtpmap:99 H264/90000
a=fmtp:99 packetization-mode=0; profile-level-id=640c1f; \\ sprop-parameter-sets=Z2QMH5WgFAFugH9Q,aM46gA==
a=rtcp-fb:* trr-int 5000
a=rtcp-fb:* nack
a=rtcp-fb:* nack pli
a=rtcp-fb:* ccm fir
a=rtcp-fb:* ccm tmmbr
a=recvonly **a=mid:11**
m=video 49158 RTP/AVPF 99
a=acfg:1 t=1
b=AS:1060
b=RS:0 b=RR:5000
a=rtpmap:99 H264/90000
a=fmtp:99 packetization-mode=0; profile-level-id=640c1f; \\ sprop-parameter-sets=Z2QMH5WgFAFugH9Q,aM46gA==
a=rtcp-fb:* trr-int 5000

```
a=rtcp-fb:* nack
a=rtcp-fb:* nack pli
a=rtcp-fb:* ccm fir
a=rtcp-fb:* ccm tmmbr
a=recvonly
a=mid:12
m=video 0 RTP/AVP 99
```

TP UE2 now sends its SDP answer, shown in Table 4. Alongside the original audio, video and CLUE m-lines, it includes two active "recvonly" m-lines and a zeroed m-line for the third, indicating that only two of the offered CLUE-controlled encodings are accepted. It adds their "mid" values to the grouping attribute to show they are controlled by the CLUE channel.

The constraints of offer/answer meant that TP UE2 could not include its encoder information as new m-lines in the SDP answer. As such TP UE2 now generates a third SDP offer. Along with all the accepted streams from the second offer-answer exchange, TP UE2 also includes two new "sendonly" streams. Each stream has a "label" corresponding to the Encoding IDs in the CLUE ADVERTISEMENT message from TP UE2. It also adds their "mid" values to the grouping attribute to show they are controlled by the CLUE channel. The resulting SDP offer is shown in Table A.1.5.

Table A.1.5: Example SDP offer for Negotiating CLUE-controlled Media

SDP offer

```
a=group CLUE 11 12 13 14 100
m=audio 49152 RTP/AVPF 96
b=AS:89
b=RS:0
b=RR:4000
a=rtpmap:96 EVS/16000/1
a=fmtp:96 br=13.2-64; bw=swb; max-red=220
a=ptime:20
a=maxptime:240
a=sendrecv
a=mid:9
m=video 49154 RTP/AVPF 99
b=AS:1060
b=RS:0
b=RR:5000
a=rtpmap:99 H264/90000
a=fmtp:99 packetization-mode=0; profile-level-id=640c1f; \
     sprop-parameter-sets=Z2QMH5WgFAFugH9Q,aM46gA==
a=rtcp-fb:* trr-int 5000
a=rtcp-fb:* nack
a=rtcp-fb:* nack pli
a=rtcp-fb:* ccm fir
a=rtcp-fb:* ccm tmmbr
a=sendrecv
a=mid:10
m=application 6100 UDP/DTLS/SCTP webrtc-datachannel
a=sctp-port: 5000
a=dcmap:2 subprotocol="CLUE"; ordered=true
a=mid:100
m=video 49156 RTP/AVPF 99
b=AS:1060
b=RS:0
b=RR:5000
a=rtpmap:99 H264/90000
a=fmtp:99 packetization-mode=0; profile-level-id=640c1f; \
```

```
     sprop-parameter-sets=Z2QMH5WgFAFugH9Q,aM46gA==
a=rtcp-fb:* trr-int 5000
a=rtcp-fb:* nack
a=rtcp-fb:* nack pli
a=rtcp-fb:* ccm fir
a=rtcp-fb:* ccm tmmbr
a=recvonly
a=mid:11
m=video 49158 RTP/AVPF 99
b=AS:1060
b=RS:0
b=RR:5000
a=rtpmap:99 H264/90000
a=fmtp:99 packetization-mode=0; profile-level-id=640c1f; \
     sprop-parameter-sets=Z2QMH5WgFAFugH9Q,aM46gA==
a=rtcp-fb:* trr-int 5000
a=rtcp-fb:* nack
a=rtcp-fb:* nack pli
a=rtcp-fb:* ccm fir
a=rtcp-fb:* ccm tmmbr
a=recvonly
a=mid:12
m=video 0 RTP/AVP 99
m=video 49160 RTP/AVP 99
a=tcap:1 RTP/AVPF
a=pcfg:1 t=1
b=AS:1060
b=RS:0
b=RR:5000
a=rtpmap:99 H264/90000
a=fmtp:99 packetization-mode=0; profile-level-id=640c1f; \
     sprop-parameter-sets=Z2QMH5WgFAFugH9Q,aM46gA==
a=rtcp-fb:* trr-int 5000
a=rtcp-fb:* nack
a=rtcp-fb:* nack pli
a=rtcp-fb:* ccm fir
a=rtcp-fb:* ccm tmmbr
a=sendonly
a=mid:13
a=label:foo
m=video 49162 RTP/AVP 99
a=tcap:1 RTP/AVPF
a=pcfg:1 t=1
b=AS:1060
b=RS:0
b=RR:5000
a=rtpmap:99 H264/90000
a=fmtp:99 packetization-mode=0; profile-level-id=640c1f; \
     sprop-parameter-sets=Z2QMH5WgFAFugH9Q,aM46gA==
a=rtcp-fb:* trr-int 5000
a=rtcp-fb:* nack
a=rtcp-fb:* nack pli
a=rtcp-fb:* ccm fir
a=rtcp-fb:* ccm tmmbr
a=sendonly
a=mid:14
a=label:bar
```

Having received this TP UE1 now has all the information it needs to send a new CLUE CONFIGURE message. It requests the two static Captures from TP UE2, to be sent on Encodings "foo" and "bar". TP UE2 receives the CLUE CONFIGURE message from TP UE1 and sends a CLUE RESPONSE message to acknowledge its receptions.

TP UE1 now sends an SDP answer matching two "recvonly" m-lines to TP UE2's new "sendonly" lines. It includes their "mid" values in the grouping attribute to show they are controlled by the CLUE channel. This is shown in Table A.1.6.

Table A.1.6: Example SDP answer for Negotiating CLUE-controlled Media

SDP answer
a=group CLUE 3 4 5 6 7 8
m=audio 49152 RTP/AVPF 96
b=AS:89
b=RS:0
b=RR:4000
a=rtpmap:96 EVS/16000/1
a=fmtp:96 br=13.2-64; bw=swb; max-red=220
a=ptime:20
a=maxptime:240
a=sendrecv
a=mid:1
m=video 49154 RTP/AVPF 99
b=AS:1060
b=RS:0 b=RR:5000
a=rtpmap:99 H264/90000
a=fmtp:99 packetization-mode=0; profile-level-id=640c1f; \ sprop-parameter-sets=Z2QMH5WgFAFugH9Q,aM46gA==
a=rtcp-fb:* trr-int 5000
a=rtcp-fb:* nack
a=rtcp-fb:* nack pli
a=rtcp-fb:* ccm fir
a=rtcp-fb:* ccm tmmbr
a=sendrecv
a=mid:2
m=application 6100 UDP/DTLS/SCTP webrtc-datachannel
a=sctp-port: 5000
a=dcmap:2 subprotocol="CLUE"; ordered=true
a=mid:3
m=video 49156 RTP/AVPF 99
b=AS:1060
b=RS:0
b=RR:5000
a=rtpmap:99 H264/90000
a=fmtp:99 packetization-mode=0; profile-level-id=640c1f; \ sprop-parameter-sets=Z2QMH5WgFAFugH9Q,aM46gA==
a=rtcp-fb:* trr-int 5000
a=rtcp-fb:* nack
a=rtcp-fb:* nack pli
a=rtcp-fb:* ccm fir
a=rtcp-fb:* ccm tmmbr
a=sendonly
a=mid:4 **a=label:enc1**
m=video 49158 RTP/AVPF 99

```
b=AS:1060
b=RS:0
b=RR:5000
a=rtpmap:99 H264/90000
a=fmtp:99 packetization-mode=0; profile-level-id=640c1f; \
     sprop-parameter-sets=Z2QMH5WgFAFugH9Q,aM46gA==
a=rtcp-fb:* trr-int 5000
a=rtcp-fb:* nack
a=rtcp-fb:* nack pli
a=rtcp-fb:* ccm fir
a=rtcp-fb:* ccm tmmbr
a=sendonly
a=mid:5
a=label:enc2
m=video 0 RTP/AVPF 99
m=video 49160 RTP/AVPF 99
a=acfg:1 t=1
b=AS:1060
b=RS:0
b=RR:5000
a=rtpmap:99 H264/90000
a=fmtp:99 packetization-mode=0; profile-level-id=640c1f; \
     sprop-parameter-sets=Z2QMH5WgFAFugH9Q,aM46gA==
a=rtcp-fb:* trr-int 5000
a=rtcp-fb:* nack
a=rtcp-fb:* nack pli
a=rtcp-fb:* ccm fir
a=rtcp-fb:* ccm tmmbr
a=recvonly
a=mid:7
m=video 49162 RTP/AVPF 99
a=acfg:1 t=1
b=AS:1060
b=RS:0
b=RR:5000
a=rtpmap:99 H264/90000
a=fmtp:99 packetization-mode=0; profile-level-id=640c1f; \
     sprop-parameter-sets=Z2QMH5WgFAFugH9Q,aM46gA==
a=rtcp-fb:* trr-int 5000
a=rtcp-fb:* nack
a=rtcp-fb:* nack pli
a=rtcp-fb:* ccm fir
a=rtcp-fb:* ccm tmmbr
a=recvonly
a=mid:8
```

Both sides of the call are now sending multiple video streams with their sources defined via CLUE negotiation. As the call progresses either side can send new CLUE Advertisement or Configure message or new SDP negotiation to add, remove or change what they have available or want to receive.

A.2 TP Session Setup with CLUE-Controlled Audio Support

The following example is similar to the example in clause A.1, but demonstrates CLUE-controlled audio support channel – the assumption here is that TP UE1 has three microphones and three speakers and TP UE2 has two microphones and two speakers. No video communication is assumed in this example.

Table A.2.1: Example SDP offer for Establishment of CLUE Data Channel

SDP offer
a=group CLUE 3
m=audio 49150 RTP/AVP 96 97 98 99 100
a=tcap:1 RTP/AVPF
a=pcfg:1 t=1
b=AS:89
b=RS:0
b=RR:4000
a=rtpmap:96 EVS/16000/1
a=fmtp:96 br=13.2-64; bw=swb; max-red=220
a=rtpmap:97 AMR-WB/16000/1
a=fmtp:97 mode-change-capability=2; max-red=220
a=rtpmap:98 AMR-WB/16000/1
a=fmtp:98 mode-change-capability=2; max-red=220; octet-align=1
a=rtpmap:99 AMR/8000/1
a=fmtp:99 mode-change-capability=2; max-red=220
a=rtpmap:100 AMR/8000/1
a=fmtp:100 mode-change-capability=2; max-red=220; octet-align=1
a=ptime:20
a=maxptime:240
a=sendrecv
a=mid:1
m=application 6100 UDP/DTLS/SCTP webrtc-datachannel
a=sctp-port: 5000
a=dcmap:2 subprotocol="CLUE"; ordered=true
a=mid:3

The SDP offer from TP UE1 in Table A.2.1 contains the basic audio stream (non-CLUE controlled) and an establishment request for a DTLS/SCTP association used to realize a CLUE data channel. A CLUE group is included and the data channel is shown in group (3).

For the basic media streams, the offer contains the AMR narrowband, AMR-WB wideband and EVS codecs, with the former two included for the purposes of interworking with MTSI clients.

Table A.2.2: Example SDP answer for Establishment of CLUE Data Channel

SDP answer
a=group CLUE 100
m=audio 49152 RTP/AVPF 96
a=acfg:1 t=1
b=AS:89
b=RS:0
b=RR:4000
a=rtpmap:96 EVS/16000/1
a=fmtp:96 br=13.2-64; bw=swb; max-red=220
a=ptime:20

```
a=maxptime:240
a=sendrecv
a=mid:9
m=application 6100 UDP/DTLS/SCTP webrtc-datachannel
a=sctp-port: 5000
a=dcmap:2 subprotocol="CLUE"; ordered=true
a=mid:100
```

The answer from TP UE2 in Table A.2.2 indicates that the CLUE data channel is accepted. Now CLUE and configure messages can be exchanged in order to negotiate on the capture and rendering capabilities for the telepresence session. Moreover, TP UE2 wishes to receive initial media, and so includes corresponding non-CLUE-controlled audio lines, accepting the use of the EVS codec.

With the successful initial SDP offer-answer, TP UEs negotiate the CLUE channel via the exchange of CLUE ADVERTISEMENT and CLUE CONFIGURE messages. TP UE1 advertises three static Captures representing its three microphones. It also includes switched Captures suitable for two- and one-speaker systems. TP UE1 also includes an Encoding Group with three Encoding IDs: "enc1", "enc2" and "enc3", which will appear in the subsequent SDP offer TP UE1 intends to send. TP UE2 also sends its CLUE ADVERTISEMENT message, where it advertises two static Captures representing its two microphones. It also includes a single composed Capture for single-speaker systems, in which it will composite the two microphone views into a single audio stream. TP UE2 also includes a single Encoding Group with two Encoding IDs: "foo" and "bar".

Following the exchange of these CLUE messages, further SDP offer-answer negotiations can occur that include CLUE-controlled encodings. Every "m" line representing a CLUE Encoding contains a "label" attribute as defined in [48] to identify the Encoding by the sender in CLUE Advertisement messages and by the receiver in CLUE Configure messages, e.g. "enc1", "enc2", "enc3" for TP UE1 and "foo" and "bar" for TP UE2.

Table A.2.3: Example SDP offer for Negotiating CLUE-controlled Media

SDP offer

```
a=group CLUE 3 4 5 6
m=audio 49150 RTP/AVPF 96
b=AS:89
b=RS:0
b=RR:4000
a=rtpmap:96 EVS/16000/1
a=fmtp:96 br=13.2-64; bw=swb; max-red=220
a=ptime:20
a=maxptime:240
a=sendrecv
a=mid:1
m=application 6100 UDP/DTLS/SCTP webrtc-datachannel
a=sctp-port: 5000
a=dcmap:2 subprotocol="CLUE"; ordered=true
a=mid:3
m=audio 49152 RTP/AVP 96
a=pcfg:1 t=1
b=AS:89
b=RS:0
b=RR:4000
a=rtpmap:96 EVS/16000/1
a=fmtp:96 br=13.2-64; bw=swb; max-red=220
a=ptime:20
a=maxptime:240
a=sendonly
a=mid:4
a=label:enc1
```

```
m=audio 49154 RTP/AVP 96
a=pcfg:1 t=1
b=AS:89
b=RS:0
b=RR:4000
a=rtpmap:96 EVS/16000/1
a=fmtp:96 br=13.2-64; bw=swb; max-red=220
a=ptime:20
a=maxptime:240
a=sendonly
a=mid:5
a=label:enc2
m=audio 49156 RTP/AVP 96
a=pcfg:1 t=1
b=AS:89
b=RS:0
b=RR:4000
a=rtpmap:96 EVS/16000/1
a=fmtp:96 br=13.2-64; bw=swb; max-red=220
a=ptime:20
a=maxptime:240
a=sendonly
a=mid:6
a=label:enc3
```

The second SDP offer in Table A.2.3 from TP UE1 maintains the "sendrecv" audio and includes three additional "sendonly" m-lines representing the three CLUE-controlled encodings for audio. Each of these m-lines contains a "label" corresponding to one of the encoding IDs from the CLUE advertisement from TP UE1. Each also has its "mid" added to the grouping attribute to show that they are controlled by the CLUE channel.

TP UE2 now has all the information it needs to decide which streams to configure. As such it now sends its CLUE CONFIGURE message. This requests the pair of switched Captures, and it configures them with encoder ids "enc1" and "enc2".

TP UE1 receives the CLUE CONFIGURE from TP UE2 and sends a CLUE RESPONSE message to acknowledge its receptions.

Table A.2.4: Example SDP answer for Negotiating CLUE-controlled Media

SDP answer
a=group CLUE 11 12 100 m=audio 49152 RTP/AVPF 96 b=AS:89 b=RS:0 b=RR:4000 a=rtpmap:96 EVS/16000/1 a=fmtp:96 br=13.2-64; bw=swb; max-red=220 a=ptime:20 a=maxptime:240 **a=sendrecv** **a=mid:9** **m=application 6100 UDP/DTLS/SCTP webrtc-datachannel** **a=sctp-port: 5000** **a=dcmap:2 subprotocol="CLUE"; ordered=true** **a=mid:100** m=audio 49154 RTP/AVPF 96

```
a=acfg:1 t=1
b=AS:89
b=RS:0
b=RR:4000
a=rtpmap:96 EVS/16000/1
a=fmtp:96 br=13.2-64; bw=swb; max-red=220
a=ptime:20
a=maxptime:240
a=recvonly
a=mid:11
m=audio 49156 RTP/AVPF 96
a=acfg:1 t=1
b=AS:89
b=RS:0
b=RR:4000
a=rtpmap:96 EVS/16000/1
a=fmtp:96 br=13.2-64; bw=swb; max-red=220
a=ptime:20
a=maxptime:240
a=recvonly
a=mid:12
m=audio 0 RTP/AVP 96
```

TP UE2 now sends its SDP answer, shown in Table A.2.4. Alongside the original audio and CLUE m-lines, it includes two active "recvonly" m-lines and a zeroed m-line for the third, indicating that only two of the offered CLUE-controlled encodings are accepted. It adds their "mid" values to the grouping attribute to show they are controlled by the CLUE channel.

The constraints of offer/answer meant that TP UE2 could not include its encoder information as new m-lines in the SDP answer. As such, TP UE2 now generates a third SDP offer. Along with all the accepted streams from the second offer-answer exchange, TP UE2 also includes two new "sendonly" streams. Each stream has a "label" corresponding to the Encoding IDs in the CLUE ADVERTISEMENT message from TP UE2. It also adds their "mid" values to the grouping attribute to show they are controlled by the CLUE channel. The resulting SDP offer is shown in Table A.2.5.

Table A.2.5: Example SDP offer for Negotiating CLUE-controlled Media

SDP offer

```
a=group CLUE 11 12 13 14 100
m=audio 49152 RTP/AVPF 96
b=AS:89
b=RS:0
b=RR:4000
a=rtpmap:96 EVS/16000/1
a=fmtp:96 br=13.2-64; bw=swb; max-red=220
a=ptime:20
a=maxptime:240
a=sendrecv
a=mid:9
m=application 6100 UDP/DTLS/SCTP webrtc-datachannel
a=sctp-port: 5000
a=dcmap:2 subprotocol="CLUE"; ordered=true
a=mid:100
m=audio 49154 RTP/AVPF 96
b=AS:89
b=RS:0
```

```
b=RR:4000
a=rtpmap:96 EVS/16000/1
a=fmtp:96 br=13.2-64; bw=swb; max-red=220
a=ptime:20
a=maxptime:240
a=recvonly
a=mid:11
m=audio 49156 RTP/AVPF 96
b=AS:89
b=RS:0
b=RR:4000
a=rtpmap:96 EVS/16000/1
a=fmtp:96 br=13.2-64; bw=swb; max-red=220
a=ptime:20
a=maxptime:240
a=recvonly
a=mid:12
m=audio 0 RTP/AVP 96
m=audio 49158 RTP/AVP 96
a=pcfg:1 t=1
b=AS:89
b=RS:0
b=RR:4000
a=rtpmap:96 EVS/16000/1
a=fmtp:96 br=13.2-64; bw=swb; max-red=220
a=ptime:20
a=maxptime:240
a=sendonly
a=mid:13
a=label:foo
m=audio 49160 RTP/AVP 96
a=pcfg:1 t=1
b=AS:89
b=RS:0
b=RR:4000
a=rtpmap:96 EVS/16000/1
a=fmtp:96 br=13.2-64; bw=swb; max-red=220
a=ptime:20
a=maxptime:240
a=sendonly
a=mid:14
a=label:bar
```

Having received this TP UE1 now has all the information it needs to send a new CLUE CONFIGURE message. It requests the two static Captures from TP UE2, to be sent on Encodings "foo" and "bar". TP UE2 receives the CLUE CONFIGURE message from TP UE1 and sends a CLUE RESPONSE message to acknowledge its receptions.

TP UE1 now sends an SDP answer matching two "recvonly" m-lines to TP UE2's new "sendonly" lines. It includes their "mid" values in the grouping attribute to show they are controlled by the CLUE channel. This is shown in Table A.2.6.

Table A.2.6: Example SDP answer for Negotiating CLUE-controlled Media

SDP answer
a=group CLUE 3 4 5 6 7 8
m=audio 49152 RTP/AVPF 96

```
b=AS:89
b=RS:0
b=RR:4000
a=rtpmap:96 EVS/16000/1
a=fmtp:96 br=13.2-64; bw=swb; max-red=220
a=ptime:20
a=maxptime:240
a=sendrecv
a=mid:1
m=application 6100 UDP/DTLS/SCTP webrtc-datachannel
a=sctp-port: 5000
a=dcmap:2 subprotocol="CLUE"; ordered=true
a=mid:3
m=audio 49154 RTP/AVPF 96
b=AS:89
b=RS:0
b=RR:4000
a=rtpmap:96 EVS/16000/1
a=fmtp:96 br=13.2-64; bw=swb; max-red=220
a=ptime:20
a=maxptime:240
a=sendonly
a=mid:4
a=label:enc1
m=audio 49156 RTP/AVPF 96
b=AS:89
b=RS:0
b=RR:4000
a=rtpmap:96 EVS/16000/1
a=fmtp:96 br=13.2-64; bw=swb; max-red=220
a=ptime:20
a=maxptime:240
a=sendonly
a=mid:5
a=label:enc2
m=audio 0 RTP/AVP 96
m=audio 49158 RTP/AVPF 96
a=acfg:1 t=1
b=AS:89
b=RS:0
b=RR:4000
a=rtpmap:96 EVS/16000/1
a=fmtp:96 br=13.2-64; bw=swb; max-red=220
a=ptime:20
a=maxptime:240
a=recvonly
a=mid:7
m=audio 49160 RTP/AVPF 96
a=acfg:1 t=1
b=AS:89
b=RS:0
b=RR:4000
a=rtpmap:96 EVS/16000/1
a=fmtp:96 br=13.2-64; bw=swb; max-red=220
a=ptime:20
```

```
a=maxptime:240
a=recvonly
a=mid:8
```

Both sides of the call are now sending multiple audio streams with their sources defined via CLUE negotiation. As the call progresses either side can send new CLUE Advertisement or Configure message or new SDP negotiation to add, remove or change what they have available or want to receive.

A.3 Interworking between TP UE and MTSI UE

The following example is similar to the example in clause A.1, but demonstrates the case of interworking between TP UE and MTSI UE.

Table A.3.1: Example SDP offer for Establishment of CLUE Data Channel

SDP offer

```
a=group CLUE 3
m=audio 49152 RTP/AVP 96 97 98 99 100
a=tcap:1 RTP/AVPF
a=pcfg:1 t=1
b=AS:34
b=RS:0
b=RR:4000
a=rtpmap:96 EVS/16000/1
a=fmtp:96 br=13.2-24.4; bw=swb; max-red=220
a=rtpmap:97 AMR-WB/16000/1
a=fmtp:97 mode-change-capability=2; max-red=220
a=rtpmap:98 AMR-WB/16000/1
a=fmtp:98 mode-change-capability=2; max-red=220; octet-align=1
a=rtpmap:99 AMR/8000/1
a=fmtp:99 mode-change-capability=2; max-red=220
a=rtpmap:100 AMR/8000/1
a=fmtp:100 mode-change-capability=2; max-red=220; octet-align=1
a=ptime:20
a=maxptime:240
a=sendrecv
a=mid:1
m=video 49154 RTP/AVP 100 99
a=tcap:1 RTP/AVPF
a=pcfg:1 t=1
b=AS:1060
b=RS:0
b=RR:5000
a=rtpmap:99 H264/90000
a=fmtp:99 packetization-mode=0; profile-level-id=42e00c; \
     sprop-parameter-sets=J0LgDJWgUH6Af1A=,KM46gA==
a=rtpmap:100 H264/90000
a=fmtp:100 packetization-mode=0; profile-level-id=640c1f; \
     sprop-parameter-sets=Z2QMH5WgFAFugH9Q,aM46gA==
a=rtcp-fb:* trr-int 5000
a=rtcp-fb:* nack
a=rtcp-fb:* nack pli
a=rtcp-fb:* ccm fir
a=rtcp-fb:* ccm tmmbr
```

```
a=sendrecv
a=mid:2
m=application 6100 UDP/DTLS/SCTP webrtc-datachannel
a=sctp-port: 5000
a=dcmap:2 subprotocol="CLUE"; ordered=true
a=mid:3
```

The SDP offer from TP UE in Table A.3.1 contains basic media streams (non-CLUE controlled) and an establishment request for a DTLS/SCTP association used to realize a CLUE data channel. A CLUE group is included and the data channel is shown in group (3).

For the basic video stream, the offer contains H.264/AVC Constrained Baseline Profile Level 1.2 (in addition to the mandatory video codec for TP UEs, namely H.264/AVC CHP Level 3.1) for the purposes of interworking with MTSI clients.

For the basic audio stream, the offer contains the mandatory EVS-SWB codec with bit rate range [13.2 to 24.4 kbps] for TP session media establishment and AMR/AMR-WB in bandwidth efficient and octet-aligned formats for enabling transcoder-free interworking with potential legacy MTSI client that does not support EVS-SWB in the TP session.

Table A.3.2: Example SDP answer for Establishment of CLUE Data Channel

SDP answer
m=audio 49152 RTP/AVPF 96 a=acfg:1 t=1 b=AS:30 b=RS:0 b=RR:4000 a=rtpmap:96 EVS/16000/1 a=fmtp:96 br=13.2; bw=swb; max-red=220 a=ptime:20 a=maxptime:240 **a=sendrecv** m=video 49154 RTP/AVPF 100 a=acfg:1 t=1 b=AS:1060 b=RS:0 b=RR:5000 a=rtpmap:100 H264/90000 a=fmtp:100 packetization-mode=0; profile-level-id=640c1f; \   sprop-parameter-sets=Z2QMH5WgFAFugH9Q,aM46gA== a=rtcp-fb:* trr-int 5000 a=rtcp-fb:* nack a=rtcp-fb:* nack pli a=rtcp-fb:* ccm fir a=rtcp-fb:* ccm tmmbr **a=sendrecv** **m=application 0 UDP/DTLS/SCTP webrtc-datachannel**

The answer from MTSI UE shown in Table A.3.2. Since the MTSI UE is not CLUE capable, it does not recognize the CLUE semantic for grouping attribute nor does it support the CLUE data channel. Accordingly, the SDP answer indicates that the CLUE data channel is not accepted.

In the meantime, the MTSI UE accepts the offered basic audio and video streams that are based on RTP payloads that it supports, in this case based on EVS-SWB and H.264/AVC CHP Level 3.1. Consequently, the TP UE can fall back to the non-CLUE operation governed by MTSI client capabilities and exchange the basic media streams with the MTSI UE.

Annex B (informative): Change history

Change history								
Date	**TSG #**	**SA Doc.**	**CR**	**Rev**	**Cat**	**Subject/Comment**	**Old**	**New**
12-2015	SA#70	SP-150646				Presented to TSG SA#70 (for approval)		1.0.0
12-2015	SA#70					Approved at TSG SA#70	1.0.0	13.0.0

Change history							
Date	**Meeting**	**TDoc**	**CR**	**Rev**	**Cat**	**Subject/Comment**	**New version**
2016-06	SA#72	SP-160259	0001	1	F	RFC 7798: RTP Payload Format for HEVC for Telepresence	13.1.0

History

Document history		
V13.0.0	January 2016	Publication
V13.1.0	August 2016	Publication