Telepresence Interoperability Protocol ("TIP")
Evaluation License

Published July 2010

THIS IS A TIME-LIMITED EVALUATION LICENSE THAT PROVIDES YOU WITH THE RIGHT TO EVALUATE THE TIP SPECIFICATION FOR NINETY (90) DAYS FOLLOWING YOUR INITIAL DOWNLOAD OR ACCESS.

This is a legal agreement ("Evaluation License") between the International Multimedia Telecommunications Consortium ("IMTC") and an individual or entity ("You") interested in downloading the Telepresence Interoperability Protocol Specification (the "TIP Specification") solely for purposes of evaluation. Cisco Systems, inc. ("Cisco") is an intended third-party beneficiary of this Evaluation License.

The TIP Specification provides a protocol for interoperability between videoconferencing products, including streaming of audio, video, and data to and from videoconferencing products. IMTC offers this license to encourage interoperability between competitive videoconferencing products that exist as of the date of this license or that may come to exist in the future. The TIP means the specification found at [IMTC Web Site].

Only the TIP Specification published by IMTC will be considered the TIP Specification for purposes of this Evaluation License. Any specification not published by IMTC, or one that incorporates the TIP Specification in part, in whole, or by reference shall not be considered the TIP Specification for purposes of this Evaluation License.

BY ACCESSING OR DOWNLOADING THE TIP SPECIFICATION FOR EVALUATION PURPOSES, YOU AGREE TO BE BOUND BY THE TERMS OF THIS EVALUATION LICENSE.

If You want a license from IMTC to evaluate the TIP Specification, You must agree to these terms. This is an offer to be accepted only on the terms set forth in this Evaluation License. You accept these terms by accessing, downloading, using, or reviewing the evaluation version of the TIP Specification. Once accepted, IMTC grants You the license below, provided that You fully comply with these terms. If You do not agree to any of these terms, then You are not licensed under this Evaluation License. If You attempt to alter or modify any of these terms, then the offer is no longer valid and is revoked.

Time-Limited License
The license granted by this Evaluation License shall extend only for a period of ninety (90) days from the date on which you first download or access the TIP Specification for the purpose of evaluation ("Limited Term").

Immediately upon expiration of the Limited Term, this license shall terminate automatically, without further notice to You. Upon expiration of this license, You agree to discontinue any and all use of the TIP Specification and destroy any copies of the evaluation version of the TIP Specification you have

Bishop Ranch 6 • 2400 Camino Ramon, Suite 375• San Ramon • CA USA• 94583 • phone +1.925.275.6600 • fax +1.925.275.6691

downloaded or otherwise created, and all extracts or summaries of the information embodied in the TIP Specification.

.

Permitted Uses

SUBJECT TO THE TERMS AND CONDITIONS OF THE EVALUATION LICENSE AND ONLY DURING THE LIMITED TERM, IMTC GRANTS YOU A LIMITED, NONEXCLUSIVE, PERSONAL, NONTRANSFERABLE LICENSE TO DOWNLOAD AND USE THE TIP SPECIFICATION SOLELY FOR PURPOSES OF YOUR EVALUATION, AND NOT FOR ANY OTHER PURPOSE, INCLUDING, BUT NOT LIMITED TO, ANY IMPLEMENTATION OF THE TIP SPECIFICATION OR THE DEVELOPMENT, MANUFACTURE, IMPORTATION, OR SALE OF ANY PRODUCTS. THIS EVALUATION LICENSE DOES NOT GRANT A LICENSE TO ANY PATENT ESSENTIAL TO IMPLEMENT THE TIP SPECIFICATION, WHETHER OWNED BY CISCO, IMTC, OR ANY THIRD PERSON. You may make only that number of copies of the TIP Specification as are necessary for Your evaluation. You may not otherwise copy the TIP Specification. You may not distribute, modify, or create derivative works of the TIP Specification. No part of the TIP Specification may be printed out, reproduced, distributed, resold, or transmitted for any other purpose, including, without limitation, commercial purposes, such as selling copies of this TIP Specification or providing paid-for support services.

If you want to use the TIP Specification beyond the Limited Term or want to use TIP Specification to implement or for any purpose other than for Your evaluation purpose, please go to [IMTC web site] to obtain the Implementation License. You acknowledge that any use of the evaluation version of the TIP Specification for any purpose prohibited under this Evaluation License, but permitted under the Implementation License subjects You to all terms of the Implementation License in the form in which it existed as of the date of your download, including, without limitation, the obligation to license patents essential to implement TIP to Cisco and any other implementer of TIP.

Reservation of Rights

IMTC reserves all rights, title, and interest in and to the TIP Specification. IMTC reserves the right to modify the TIP Specification without notice to you. No right or license to any IMTC or Cisco intellectual property is granted under this Evaluation License for any use except the evaluation of the TIP Specification. All rights not expressly granted herein are reserved.

No Warranty

THE TIP SPECIFICATION AND THE LICENSE GRANTED UNDER THIS EVALUATION LICENSE ARE PROVIDED "AS IS", WITHOUT WARRANTY OF ANY KIND, EITHER EXPRESS, IMPLIED, OR STATUTORY, INCLUDING BUT NOT LIMITED TO WARRANTIES OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE, TITLE OR NON-INFRINGEMENT. NONE OF IMTC, ANY IMTC AFFILIATE, CISCO, OR ANY CISCO AFFILIATE SHALL HAVE ANY DUTY TO INDEMNIFY, HOLD HARMLESS OR DEFEND YOU OR ANY THIRD PARTY FROM AND AGAINST ANY LOSS, DAMAGE, LAWSUITS, PROCEEDINGS, CLAIMS OR SIMILAR ACTIONS THAT ARISE OR RESULT FROM YOUR EVALUATION OF THE TIP SPECIFICATION, IN WHOLE OR IN PART.

Limitation Of Liability
NONE OF IMTC, ANY IMTC AFFILIATE, CISCO, OR ANY CISCO AFFILIATE SHALL BE LIABLE UNDER OR IN CONNECTION WITH THIS EVALUATION LICENSE (WHETHER FOR BREACH OF WARRANTY, IN CONTRACT, TORT, NEGLIGENCE, STRICT LIABILITY OR OTHERWISE):

(A) FOR ANY INDIRECT, SPECIAL, CONSEQUENTIAL, PUNITIVE OR INCIDENTAL DAMAGES; OR
(B) FOR ANY DIRECT LOSSES, INCLUDING WITHOUT LIMITATION COSTS OF PROCUREMENT OF SUBSTITUTE GOODS OR SERVICES, LOSS OF PROFITS, LOSS OF REVENUE, LOSS OF USE, LOSS OF OPPORTUNITY, INTERRUPTION OF BUSINESS; OR
(C) FOR ANY LOSSES, DAMAGES, COSTS, CLAIMS OR OTHER LIABILITIES ARISING FROM OR RELATED TO THIS EVALUATION LICENSE WHATSOEVER,

IN EACH CASE HOWEVER CAUSED OR ARISING, PROVIDED THAT NOTHING IN THIS EVALUATION LICENSE SHALL EXCLUDE OR LIMIT THE LIABILITY OF IMTC, OR ANY IMTC AFFILIATE, OR CISCO, OR ANY CISCO AFFILIATE: (1) FOR DEATH OR PERSONAL INJURY RESULTING FROM ITS NEGLIGENCE; OR (2) FOR FRAUD; OR (3) TO THE EXTENT SUCH LIABILITY CANNOT BE EXCLUDED OR LIMITED BY APPLICABLE LAW. YOU ACKNOWLEDGE THAT (A) IMTC WOULD NOT HAVE ENTERED INTO THIS EVALUATION LICENSE ON THE TERMS HEREOF WITHOUT THE LIMITATIONS OF LIABILITY SET FORTH HEREIN; AND (B) THAT CISCO WOULD NOT HAVE ASSIGNED TIP TO IMTC FOR IMTC'S LICENSE UNDER THIS EVALUATION LICENSE WITHOUT THE LIMITATIONS OF LIABILITY SET FORTH HEREIN .

Governing Law
This Evaluation License shall be construed in accordance with, and all disputes hereunder shall be governed by, the laws of England and Wales. You consent to the personal and exclusive jurisdiction of the courts of England and Wales over You and waive any objection to venue in the courts of England and Wales.

General
This Evaluation License is the only agreement between You and IMTC relating to is subject matter, and supersedes any prior agreement, understanding, or communication. This Evaluation License and your rights under this Evaluation License are personal to You and may not be assigned or transferred. Any assignment or transfer is null and void.

Bishop Ranch 6 • 2400 Camino Ramon, Suite 375• San Ramon • CA USA• 94583 • phone +1.925.275.6600 • fax +1.925.275.6691

Cisco TIP Endpoint 1.8 to 1.10 Implementation Profile

(for use with TIP v7 or TIP v8)

Modification History

Revision	Date	Originator	Comments
1.0	12/12/2012	Cisco Systems, Inc.	Initial document

Contents

1 Introduction to TIP

Telepresence Interoperability Protocol (TIP) systems are generally high-end, high-definition video conferencing devices capable of handling multiple audio and video streams. Capabilities that are negotiated in TIP are complimentary to those which are signaled and negotiated using VoIP call setup signaling protocols, such as SIP/SDP.

TIP devices can be endpoints, including single and multi-screen systems participating in point-to-point and multipoint sessions. Media sources are switched when necessary to always present the viewer with the most suitable session participants.

TIP devices can be multipoint devices, such as a multipoint control unit (MCU). In the case of multipoint sessions, endpoints will exchange TIP messaging with MCU that implements TIP. For purposes of this document, "MCU" or "multipoint control device" or "multipoint device" will be used interchangeably, referring to a multipoint session controlling device that may or may not terminate or transcode any of the video or audio media before forwarding on to the rest of the endpoints in a multipoint session.

2 Introduction to Cisco TIP Implementation Profile(s)

This Cisco TIP Implementation Profile document explains what options Cisco TelePresence devices require, can accept and/or do prefer among those defined in the TIP specification. Additional information complementary to TIP needed to achieve interoperability with a Cisco TelePresence installation is also included below, such as what SIP/SDP messaging is required and how to establish encrypted channels.

These profile documents are written with the assumption that an implementer of a video conferencing device will want it to interoperate with the broadest set of Cisco TelePresence products that implement TIP in the broadest set of use cases. This document does not consider interoperability with constrained use cases or interoperability within a limited subset of Cisco TelePresence devices, which may or may not be more relaxed than the requirements or restrictions described below.

These profile documents will be updated as needed for clarity or corrections and new versions will be published as Cisco TelePresence products evolve or as new software releases enable new options, sometimes asynchronous from revisions to the TIP protocol. Information related to interoperating with another company's TIP products is not considered in this document.

The profile information below includes requirements related to SIP signaling, TIP signaling, media encoding constraints, and other general behavior, to help achieve interoperability with a Cisco TelePresence system supporting TIP v7 or later.

2.1 Cisco Implementation Profile 1.8 through 1.10 for TIP v7 or v8 Installations

This Cisco TIP Implementation Profile 1.8 through 1.10 is for helping enable a third-party TIP implementation to fit in to a Cisco TelePresence TIP v7 or TIP v8 installation with software 1.8.x,1.9.x and 1.10 releases on the Cisco Immersive TelePresence System (CTS) endpoints, software release 2.3.x and 3.0.x on the Cisco TelePresence Server. Section 14 has additional information for a Cisco TelePresence TIP v8 installation, only enabled with CTS software release 1.10 installed. Please reference the separate Cisco TIP Endpoint Profile TX 6.0 document [11] for additional information for interoperating in a Cisco TIP v8 installation with devices using the TX 6.0 or later release only available on the newer TX versions of Cisco's TelePresence Systems.

This update for 1.8 through 1.10 installations includes these changes or additions compared to the 1.7 revision of the Cisco Endpoint Profile;

1. Clarification/changes to SDP requirements for call setup and H.264 profile indication (section 3.1 and 3.2) [1]
2. Corrections to the Video line rate (b=TIAS) to resolution mapping (section 3.2.7) [1]
3. Unrestricted H.264 video media support for 720p (section 8.2.12, 13.x and 15.2-15.5)
4. TIP information in SIP Contact Header (section 16)
5. TIP and Native SIP Interoperability Considerations (section 17)
6. Use of TIPv8 "Unrestricted Media Constraints Interoperability" (section 14)

Product Family	TIP v6	TIP v7	TIP v8
Cisco TelePresence Immersive endpoints (500, 1100, 1300, 3000, 9000, etc)	Release 1.6.5 and later releases	Release 1.7.0 and later releases	Release 1.10 plus the TX 6.0 [2] and later releases
Cisco TelePresence Multipoint Switch	Release 1.6.4 and later releases	Release 1.7.0 through 1.9.x	Not supported
Cisco TelePresence Server	Release 2.1.0 and later releases	Release 2.2.x and later releases	Future release

Table 1: Cisco Products with TIP support

[Please check www.cisco.com/go/tip for the latest updates to this table.]

[1] These corrections or clarifications are required to interoperate properly, whereas the other items are optional or recommended only compared to an implementation built around the 1.7 version of this profile.

[2] Please see the Cisco TIP Endpoint Profile TX 6.0 document [11], a supplement to this document

2.2 Interoperating in a typical Cisco TelePresence Deployment

To interoperate with the broadest set of Cisco TelePresence products, an implementer's product will want to operate seamlessly within Cisco's Unified Communications (UC) architecture and products, in particular Cisco's Unified Communications Manager (UCM). To the UCM, your TIP device will appear as UC SIP devices with audio and video capabilities.

A TIP implementer's device using this profile will typically attach to UCM as a as a generic SIP line device using the standard SIP line interface [10], as a SIP trunk device using the UCM SIP trunk interface [17] or as a SIP Extension partner device [18].

Generally, the latest Cisco UCM release is recommended for a Cisco 1.8 through 1.10 profile – based deployment, but please see Cisco's online product compatibility Information [21] for more details on the deployment combinations tested.

2.3 General Device Implementation Requirements for Cisco TIP Deployments

This section provides general requirements for TIP devices to interoperate with Cisco TelePresence systems. More details will be covered in later sections.

Profile	Main HD Video Streams	Simulcast SD Video (legacy video)	Main Shared Video (AUX)	Simulcast Shared Video (AUX)	Main Audio (spatial)	Simulcast mixed Audio	Totals
Triple Screen (TX)	3 x 720p30 -or- 3 x 1080p30	3 x CIF	1 x XGA at 30/1/5 fps	1 x XGA at 30 fps	3 x AAC-LD -or- 4 x AAC-LD	1 x G.711 -or- 1 x G.722	Up to 8 video and 5 audio
Triple Screen (RX)	3 x 720p30 -or- 3 x 1080p30	1 x CIF	1 x XGA at 30/1/5 fps		3 x AAC-LD -or- 4 x AAC-LD		Up to 4 video and 4 audio
Single Screen (TX)	1 x 720p30 -or- 1 x 1080p30	1 x CIF	1 x XGA at 30/1/5 fps	1 x XGA at 30 fps	1 x AAC-LD -or- 2 x AAC-LD	1 x G.711 -or- 1 x G.722	Up to 4 video and 3 audio
Single Screen (RX)	1 x 720p30 -or- 1 x 1080p30	1 x CIF	1 x XGA at 30/1/5 fps		3 x AAC-LD -or- 4 x AAC-LD		Up to 2 video and 4 audio

Table 2: Profile Audio and Video Stream Summary

The general requirements include:

1. SIP signaling for call and media setup
2. H.264 video encoding and decoding at 1080P and/or 720P resolution at 30 fps
3. H.264 video encoding and decoding at XGA resolution for shared stream (AUX)
4. H.264 video encoding and decoding at CIF resolution for 'legacy' video simulcast
5. AAC-LD audio encoding and decoding for all non-legacy audio
6. G.711 audio encoding for recommended "legacy audio" support
7. G.722 audio encoding for recommended "legacy audio" support
8. A recommended auxiliary video encoding device for shared/AUX content (eg, typically used for presentations) and an accompanying ability to encode audio input for shared/AUX content (eg, audio used with the presentation).
9. Ability to additionally mix all audio inputs and encode them using G.711 or G.722 for recommended "legacy audio" support in addition to encoding all audio in multiple AAC-LD streams (see Table 2).
10. Ability to receive and decode "n" number of H.264 encoded streams at 1080P or 720P [for triple screen devices, n=3. For single screen devices, n=1] and optionally CIF resolutions as well as an optional H.264 encoded stream at XGA resolution for the shared/AUX stream.

11. Ability to receive, decode and mix as need for up to four (4) AAC-LD audio streams.
12. RTP/RTCP support using Audio and Video Profile (AVP) profile and optionally Secure AVP (SAVP) profile
13. Optional ability to support SRTP encryption and DTLS-SRTP [19] and EKT [20]

3 SIP signaling

1. TIP devices MUST attach directly or indirectly to the Cisco UCM as a as a generic SIP line device using the standard SIP line interface [10], as a SIP trunk device using the UCM SIP trunk interface [17] or as a SIP Extension partner device [18]. The actual specification to be used depends on the type of SIP device through which the TIP device attaches to the UCM as well as the release of the UCM. Unless otherwise noted, this document refers to the latest UCM release available at the time of its publication.
2. The initial offer SHOULD send the TIP indicators in the SIP Contact Header as defined in section 15. These indictors MUST be sent to enable an automated attachment via TIP to the Cisco TelePresence Server 2.3 and later releases. Conversely, if a device wishes to not start a TIP session, it MUST NOT include these indicators. Note that these TIP indicators in the SIP Contact Header can only be sent or received in a UCM 8.5.x or later deployment.
3. TIP devices SHOULD NOT transmit SIP INFO [14] for codec control when TIP has been successfully negotiated in any subsequent invite.
4. Standard SIP features such as hold/resume MUST be supported as specified in the relevant UCM SIP specification [10, 17, 18].
5. TIP devices SHOULD NOT transmit RTCP packets during HOLD nor depend on receiving them.
6. TIP devices MUST send and receive mid-call INVITEs that specify updated audio and video bit rates following a successful TIP negotiation. More detailed requirements on this behavior are included later in this document.

3.1 SDP Audio Signaling

1. TIP devices MUST specify AAC-LD format before specifying G.711 and G.722 formats in the first SDP audio media line. Only the AAC-LD codec can be negotiated in SIP for Cisco devices to establish a TIP session.
2. It is STRONGLY RECOMMENDED for TIP devices to use 96 as the default AAC-LD dynamic payload type. If another payload type is used, then it MUST support asymmetric payload types and MUST support dynamic changes in payload types while media is active.
3. TIP devices MUST use RFC 3640 [15] and aac-hbr mode in their AAC-LD SDP signaling.
4. The AAC-LD requirements specified in Section 13 MUST take precedence over what is negotiated in the audio SDP ftmp line.
5. TIP devices MUST include a b=TIAS [16] field associated with the audio media line in every SDP offer and answer. The negotiated bit rate value is the minimum of the b=TIAS values in the SDP offer and answer.

    5.1. For the initial call setup and after a resume (i.e., first reINVITE transaction resuming media), the value of the b=TIAS field for audio SHOULD be equal to 128 Kbps for a single screen endpoint and 256 Kbps for a triple screen endpoint.
    5.2. After TIP negotiation has completed, a mid-call INVITE will be needed to make adjustments to the value of the b=TIAS for the audio line to accommodate the needs of the audio stream positions negotiated in TIP (eg, multiple screens, legacy and AUX). Reference section 3.1.7 below.
6. TIP devices MUST NOT initiate any mid-call INVITE transactions after call setup or after resume transactions until the TIP negotiation has completed.
7. All mid-call INVITEs, after TIP negotiation has completed and until the completion of a hold transaction, MUST have the b=TIAS value for audio set to the maximum of the sum of the bitrates for the audio stream positions, 64 kps each, negotiated in the TIP session for transmission by either of the peers, subject to the following rules[1];

    7.1. For the point-to-point calls, legacy audio MUST NOT have their bitrates included in the new b=TIAS value for audio calculation. For non-secure sessions, two single screen endpoints will use 128 Kbps and any combination that includes a triple screen endpoint will use 256 Kbps, neither of which has the legacy audio rates included. Reference section 5 for adding security overhead to the value.
    7.2. For the multipoint calls, MUST includes legacy audio in their bitrates, with a potential maximum of 320 Kbps for non-secure sessions. Reference section 5 for adding security overhead to the value.
8. The negotiated bit rate MUST NOT be changed during the same TIP session.

```
m=audio 21572 RTP/AVP 96 9 0 101^
b=TIAS:128000
a=rtpmap:96 mpeg4-generic/48000
a=fmtp:96 profile-level-id=16;streamtype=5;mode=AAC-
hbr;config=B98C00;sizeLength=13;indexLength=3;indexDel
taLength=3;
constantDuration=480

a=rtpmap:9 G722/8000
a=rtpmap:0 PCMU/8000
a=rtpmap:101 telephone-event/8000
```

Figure 2: Sample audio SDP for an initial TIP endpoint offer

3.2 SDP video signaling

1. TIP devices MUST signal H264 codec in the first video media line.
2. TIP devices MUST support the reception of a profile-level-id value of 4D0028 as well as MUST support the reception profile-level-id values of 42xxxx (720p30 minimum)
3. TIP devices, in the initial SDP invite,
    a. SHOULD indicate profile-level-id value of 42xxxx (minimum of 720p30)
    b. MAY indicate profile-level-id value of 4d0028
    c. MUST indicate a packetization-mode 1
4. It is RECOMMENDED for TIP devices to use 112 as the default dynamic H.264 payload type. If another payload type is used the device MUST support asymmetric payload types and MUST support dynamic changes in payload types while media is active.
5. TIP devices MUST include a b=TIAS [16] field associated with the video media line in every SDP offer and answer.
    a. For the initial call setup and after a resume (i.e., first reINVITE transaction resuming media), the value of the b=TIAS field for video MUST be that of a single HD video stream. Reference section 3.2.7 and Table 1 for the value in non-secure sessions. Reference section 5 for adding security overhead to the value.
    b. After TIP negotiation has completed, a mid-call INVITE will be needed to make adjustments to the value of the b=TIAS for the video line to accommodate the needs of all of the video stream positions negotiated in TIP (eg, multiple screens, legacy and AUX). Reference section 3.8 and Table 2 for the value in non-secure sessions. Reference section 5 for adding security overhead to the value.

6. TIP devices MUST NOT initiate any mid-call INVITE transactions after call setup or after resume transactions until the TIP negotiation has completed.
    a. The b=TIAS value for the initial call setup and after the resume MUST be equal to or larger than .936 Mbps. The negotiated bit rate value is calculated as the minimum of the b=TIAS values in the SDP offer and answer. The negotiated bit rate value for a non-secure session determines the maximum video resolution for the call. A negotiated value equal to or above 3.0 Mbps will result in 1080P resolution mapping, and so on, per the table below.[3] Reference section 5 and 6 for adding security overhead to the value.

Negotiated video line bit rate (Mbps)	Corresponding video resolution mapped
b=TIAS >= 3.0	**1080p**
.936 < b=TIAS < 3.0	**720p**
b=TIAS < .936	**Call Drop**

Table 1: Bit rate (b=TIAS) to video resolution mapping

7. All mid-call INVITEs, after TIP negotiation has completed and until the completion of a hold transaction, MUST have the b=TIAS value for video set to the sum of the bit rates of all video streams negotiated in TIP, including all main screen video streams, any legacy video streams as well as the shared presentation (AUX) streams.
    a. The bit rate for each legacy video stream MUST be **704 Kbps** per TIP v7 spec.
    b. The presentation bit rate is calculated based on the total number of shared (AUX) streams negotiated in the current TIP session and is based on the first negotiated frame rate. Thus, the first bit rate for AUX negotiated is always used for the bit rate calculation irrespective of whether the presentation video is active or not and irrespective of whether a lower frame rate is being used. Below is the mapping between number of shared streams, initial frame rates negotiated and the associated total bit rates for use in calculating the video line bit rate value.
        i. Example b=TIAS value for video in a mid-call reINVITE; 3 video screens at 1080p (3000 Kbps or more each) + 3 legacy streams (704 Kbps each) and 1 AUX stream at 5 fps (500 Kbps) = 11612 Kbps (or more) [3]

[3] The instructions do not take in to consideration any changes any B2BUA may make to an INVITEs media line bitrates for a given session level bit rate or for any other reason.

Number of shared (AUX) streams	Initial presentation frame rate	Total bit rate (Mbps)
0	-	0
1	30	4
1	5	.5
1	1	.1
2	any combo	4.5

Table 2: AUX quantity, frame rates to bit rate mappings

8. TIP devices MUST send a mid-call INVITE once TIP negotiation has completed on both video and audio sessions, if its video MUXCTRL packet has a lower NTP timestamp value than that of the remote peer's and the remote peer has not specified itself as a multi-point device in the TIP negotiation.
9. The negotiated bit rate MAY be changed mid-call by sending a mid-call INVITE with a new b=TIAS value

10. TIP devices MUST recalculate the negotiated bit rate for a single video stream for each mid-call INVITE transaction and adjust its video encoders accordingly.

```
m=video 26618 RTP/AVP 112
b=TIAS:4000000
a=rtpmap:112 H264/90000
a=fmtp:112 profile-level-
id=4d0028;sprop-parameter-
sets=R00AKEWUgDwBDyA=,SErjyA==;packeti
zation-mode=1
```

4 Sample non-secure SIP call flows

4.1 TIP call setup

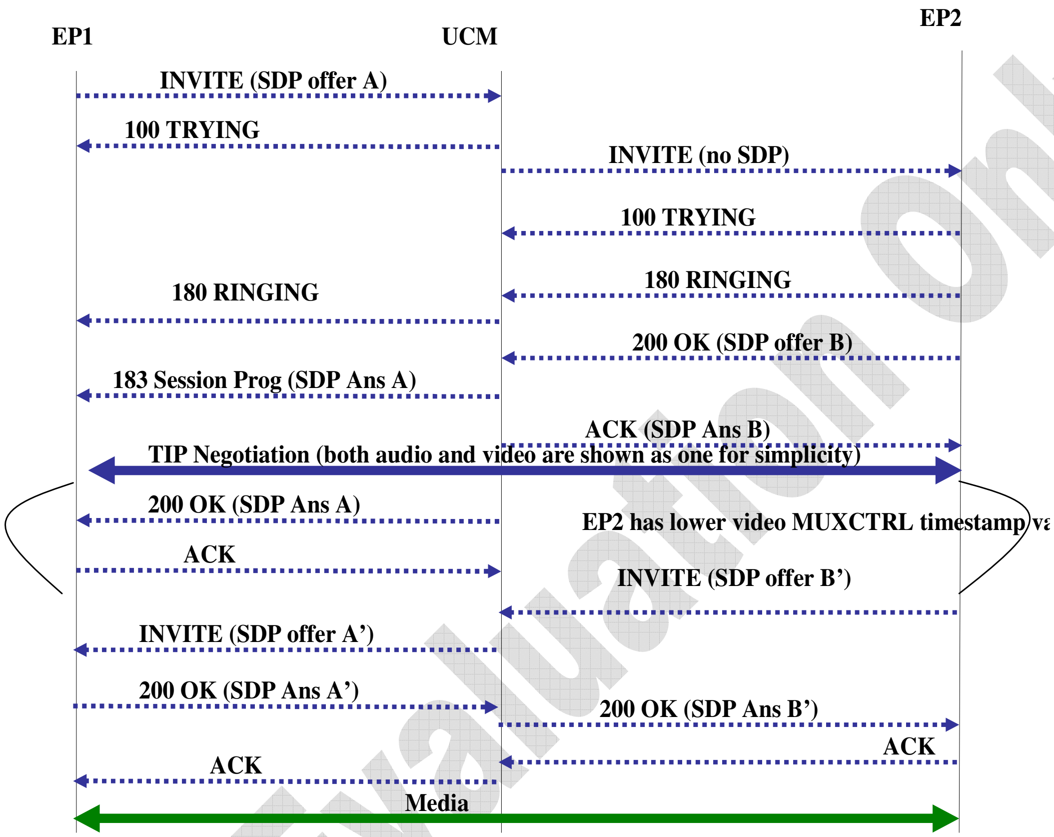

Figure 4: TIP endpoint calling another endpoint

Figure 4 shows a call setup between two TIP endpoints. The first INVITE transaction negotiates a single audio and video media line with H.264 and AAC. The video resolution is determined by the bit rates specified by the b=TIAS parameter. As stated in section 3, the video bit rate in this offer/answer exchange represents a single HD video stream.

Once each device has the media information of the other side, at the end of the first SDP offer/answer exchange, TIP negotiation MUST start on both the audio and the video channels. Once TIP has completed on both the audio and video channels (see 5 for definition of TIP negotiation completion), the endpoint that has a lower NTP timestamp in its last transmitted video MUXCTRL packet compared to the last video MUXCTRL packet received from the remote side MUST send a mid-call INVITE which updates the video and audio bit rates according to the capabilities negotiated in TIP (see 3 for more details on bit rate negotiation). Note that the mid-call bit rate negotiation could end up downgrading the video resolution and/or bit rate from that negotiated in the initial call setup, due to policies enforced by UCM. Once the mid-call SIP INVITE updating the bit rates have been completed, each endpoint SHOULD start transmitting their media streams.

4.2 TIP multi-point call setup

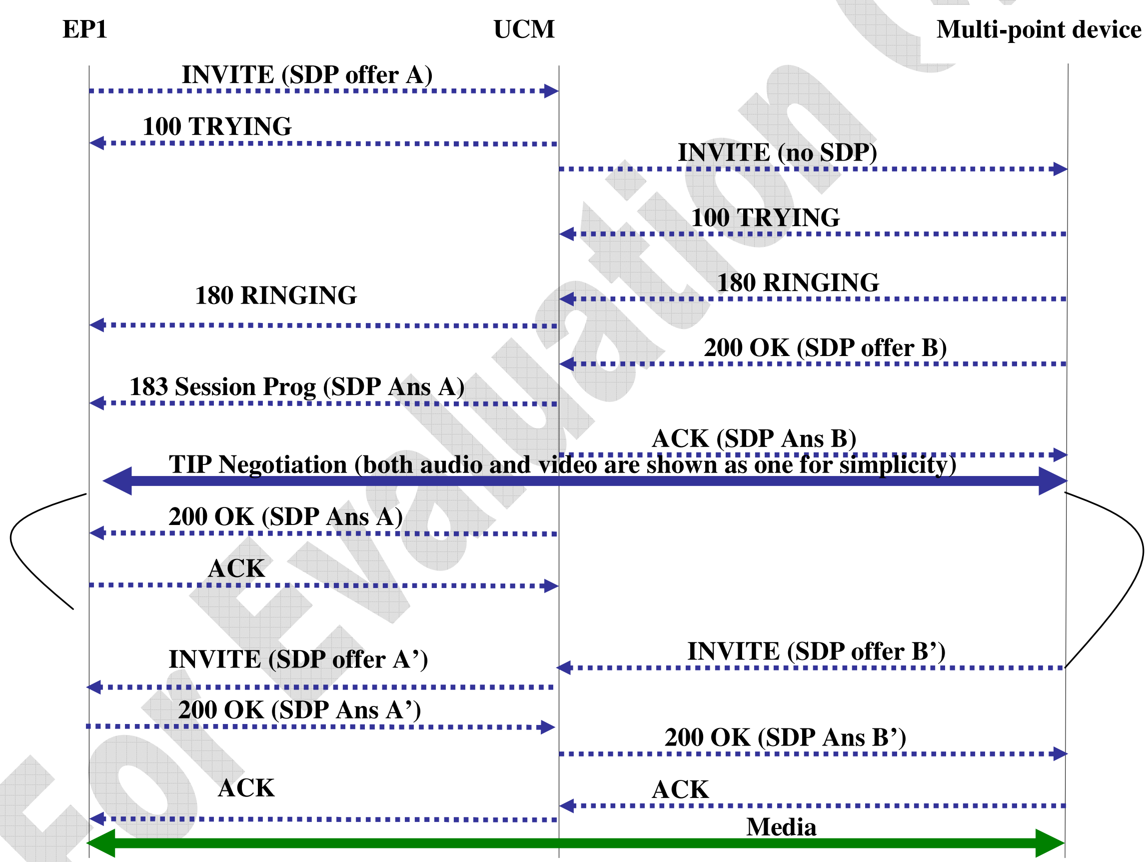

Figure 5: TIP endpoint calling a TIP multi-point device

Figure 5 shows a TIP endpoint calling a TIP multi-point device to join an ongoing multi-point session. This multi-point call flow is similar to the point to point call flow except that it is the multi-point device that MUST initiate the mid-call INVITE to update the media bit rates, once the TIP negotiation has completed on the video and audio channels.

4.3 First participant in a TIP multi-point session

Figure 6 shows a TIP endpoint dialing into a TIP multi-point device as a first participant. This flow is similar to the previous flow except that once TIP negotiation finishes, the multi-point device will send a mid-call INVITE that puts the call on hold, stopping media transmission. The call will be resumed once a second participant joins the multi-point session. One thing to note is that the SDP hold message has the media bit rate in the b=TIAS parameters representing the values negotiated in the TIP audio and video sessions. The same is true for the endpoint's SDP answer. Also note that for some conferences, the first participant may not be resumed until after the "conference host" endpoint joins, which might be well after second participant joins.

4.4 Resuming a TIP call after hold

This call flow is identical to Figure 6. At resume time the TIP negotiation MUST BE repeated and the same strategy used at call setup for bit rate negotiation is repeated. No decisions are made at resume time based on what was negotiated before the hold. A call is identified as on hold if the SDP at the audio media line level has connection information (c=) specifying a 0.0.0.0 IP address or has a send-only or inactive attribute. The same information specified at the session level while omitted at the audio media line level will also indicate hold.

A call is resumed when a non-hold SDP is received following a hold SDP.

4.5 Reducing the video bit rate of a multi-point call

In this example, a TIP endpoint has joined a multi-point conference and the negotiated bit rate for a single video stream was 4 Mbps. At some point the TIP multi-point device needs to reduce the video bit rate of video streams in the conference to 2.5 Mbps, because a new endpoint, which can only handle 2.5 Mbps per video stream, has joined the conference (this example is used for illustrative purposes only as there may be other reasons to reduce the bit rate during a session). The multi-point device will send a mid-call INVITE to the TIP endpoint, similar to the mid-call INVITE in Figure 4, specifying a lower bit rate than the one negotiated in the previous INVITE transaction.

Note that the new video bit rate will be the desired bit rate per video stream multiplied by the number of HD video streams negotiated in TIP, plus the bit rate associated with the legacy video streams multiplied by the number of negotiated legacy video streams, plus the presentation bit rate. As before, the presentation bit rate will be added irrespective of whether or not the presentation is active at this time. Also, the presentation bit rate will be calculated based on the number of shared streams and the frame rate first negotiated in the TIP session irrespective of whether this frame rate has been downgraded later or not.

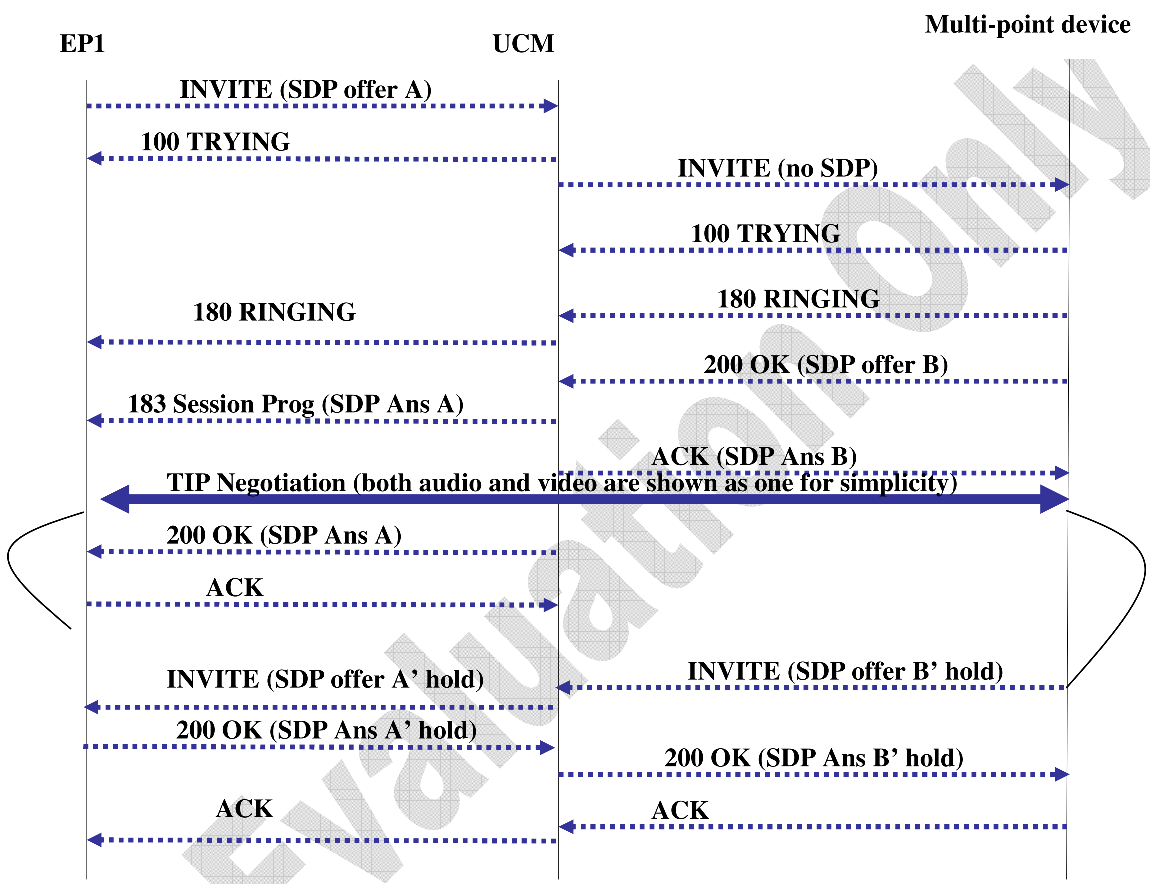

Figure 6: TIP endpoint joining a multi-point session as first participant

5 Security Requirements

5.1 SIP, SDP and DTLS-SRTP requirements

Secure TIP devices implementing this profile MUST adhere to all the requirements specified in the TIP v7 specification related to security [11] and:

1. MUST support SDES [3] or DTLS-SRTP [19] using the opportunistic approach or SDP negotiated approach as specified in section 5 of this document
2. SHOULD support DTLS-SRTP negotiation in SDP and fingerprint security context [19]
   2.1. If signaling security mechanisms in SDP, device MUST support requirements specified in the appropriate UCM SIP line specifications [10, 17, 22]
3. If supporting SDES, it is MUST offer `SRTP_AES128_CM_HMAC_SHA1_32` as one of the SRTP profiles for best interoperability with the installed base of Cisco TIP devices.
4. If supporting DTLS-SRTP, MUST support `SRTP_AES128_CM_HMAC_SHA1_80` SRTP/SRTCP protection profiles:
5. MUST support RSA, but it is RECOMMENDED to support both RSA and Diffe-Hellman, with Diffe-Hellman preferred, by listing first, in the DTLS handshake
6. MUST support remote peers using self-signed certificates in their DTLS exchange
7. MUST add security overhead to all bitrates negotiated in SDP when communicating using SRTP as specified in section 3 and 5 of this document.
8. As indicated in the TIP v7 specifications, TIP devices MUST NOT encrypt the voice activity metric or the video refresh flag when using these features in a secure TIP session
9. Secure TIP devices MUST be able to handle both encrypted and unencrypted SRTCP packets transmitted over a secure TIP session. The SRTCP E bit will be used as the indicator for whether the packet is encrypted or not. This behavior is irrespective of what has been negotiated in SDP or DTLS-SRTP.

5.2 EKT requirements

Secure TIP devices implementing this profile and supporting EKT MUST adhere to all the requirements specified in the TIP v7 specification related to security [11] and:

1. MUST indicate the ability to receive group EKT security parameters and MUST NOT indicate the ability to transmit group EKT security parameters in MediaOptions packet.

2. When support for group EKT security parameters has been negotiated in TIP, secure TIP devices MUST NOT start media transmission until an SPIMAP packet is received, authenticated and correctly decrypted.
3. Once a valid SPIMAP is received, TIP devices MUST transition to EKT mode.
4. All SRTP and SRTCP packets transmitted by a TIP device operating in EKT mode MUST include an EKT authentication tag [20].
5. TIP devices operating in EKT mode MUST use the parameters specified in the SPIMAP packet to encrypt and authenticate SRTP packets and generate EKT authentication tags.
6. TIP devices operating in EKT mode MUST use the parameters specified in the SPIMAP packet to authenticate and optionally encrypt SRTCP packets that do not use the local MUX-SSRC as their source SSRC.
7. TIP devices operating in EKT mode MUST use the keys and security parameters negotiated in signaling to authenticate and optionally encrypt SRTCP packets that use the local MUX-SSRC as their source SSRC.
8. All SRTCP packets transmitted using the MUX-SSRC as the source SSRC, MUST only include an abbreviated rather than a full EKT authentication tag.
9. All SRTP and SRTCP packets received by a TIP device operating in EKT mode MUST be processed using the rules specified in the EKT specifications [20]
10. TIP devices operating in EKT mode MUST use the parameters specified in the SPIMAP packet to validate the authentication and decrypt SRTP packets
11. TIP devices operating in EKT mode MUST use the parameters specified in the SPIMAP packet to validate the authentication of SRTCP packets that do not use the remote MUX-SSRC as their source SSRC
12. TIP devices operating in EKT mode MUST use the security parameters negotiated in signaling to validate the authentication of SRTCP packets that use the remote MUX-SSRC as their source SSRC and include an abbreviated EKT authentication tag
13. For every RTP SSRC a TIP device operating in EKT mode MUST regularly transmit packets with a full EKT authentication tag that includes the SRTP encrypted master key.
14. TIP device operating in EKT mode SHOULD include a full EKT authentication tag for all standard RTCP packets such as SR and RR
15. TIP devices operating in EKT mode SHOULD include a full EKT authentication tag in the first packet of every video frame
16. For audio streams, TIP devices operating in EKT mode SHOULD transmit a packet with a full EKT authentication tag every 100 milliseconds

6 Secure point to point TIP Session Establishment

Optionally, TIP endpoints can securely communicate using this subset of Cisco TIP profile. To participate in a secure point to point session:

- TIP devices MAY connect to the UCM using secure TLS connection. The availability of TLS as an option depends on the security configuration for the TIP device on the UCM and the version of the UCM used. TIP devices should consult the relevant UCM documentation for the availability and the requirements for TLS connection, including the UCM Security for Generic Secure Video SIP Endpoints information [22].
- At call setup TIP devices MAY signal support for the SAVP profile in their SDP offer for both the audio and video media lines. For each media line signaling the SAVP profile, an SDP crypto attribute MUST be specified similar to TLS, the option and requirements for signaling secure media depends security configuration for the TIP device on the UCM and the version of the UCM. TIP devices signaling SAVP MUST adhere to the requirements specified in the UCM SIP interface [10, 17, 22], relevant to the SIP device type used.
- TIP devices supporting the signaling of SAVP in SDP SHOULD support security fallback as specified in the relevant Cisco UCM SIP interface documentation [10, 22]. Security fallback enables interoperability between a secure TIP device and a non-secure TIP device. Note that if an administrator disables security fallback on Cisco UCM, non-secure devices will be dropped from calls comprised of secure devices.
- Once secure TIP negotiation starts all requirements applying to non-secure TIP sessions MUST apply to secure TIP sessions.
- A secure TIP device MUST NOT initiate re-keying within the same TIP session.
- All mid-call INVITEs following the completion of a TIP session and until the completion of a SIP Hold transaction MUST have the same RTP profile as the one negotiated during call setup irrespective of whether a secure session succeeded or not.
- Regardless of SAVP or AVP negotiated in the original invite, all mid-call INVITEs following the successful completion of a secure TIP session and until the completion of a SIP Hold transaction MUST add 10% security overhead to the audio bit rate in the SDP b=TIAS line operating in SRTP mode and 5% security overhead to the video bit rate in the SDP b=TIAS line when in SRTP mode.
- Latest releases of Cisco TIP devices support negotiating DTLS-SRTP roles and/or SDES in SDP. Depending on the UCM version deployed, it may then be possible to establish a secure call without relying on opportunistic DTLS-SRTP as previously mandated. For the best secure session interoperability results with the installed base, it is RECOMMENDED that a secure TIP device always start opportunistic DTLS-SRTP session establishment [19] once call setup completes, per section 6.2, if DTLS-SRTP roles were not negotiated in SDP in the delayed offer from the UCM. Even if SDES was negotiated in SDP, most Cisco TIP devices will attempt opportunistic DTLS-SRTP per the flow in figure 7.

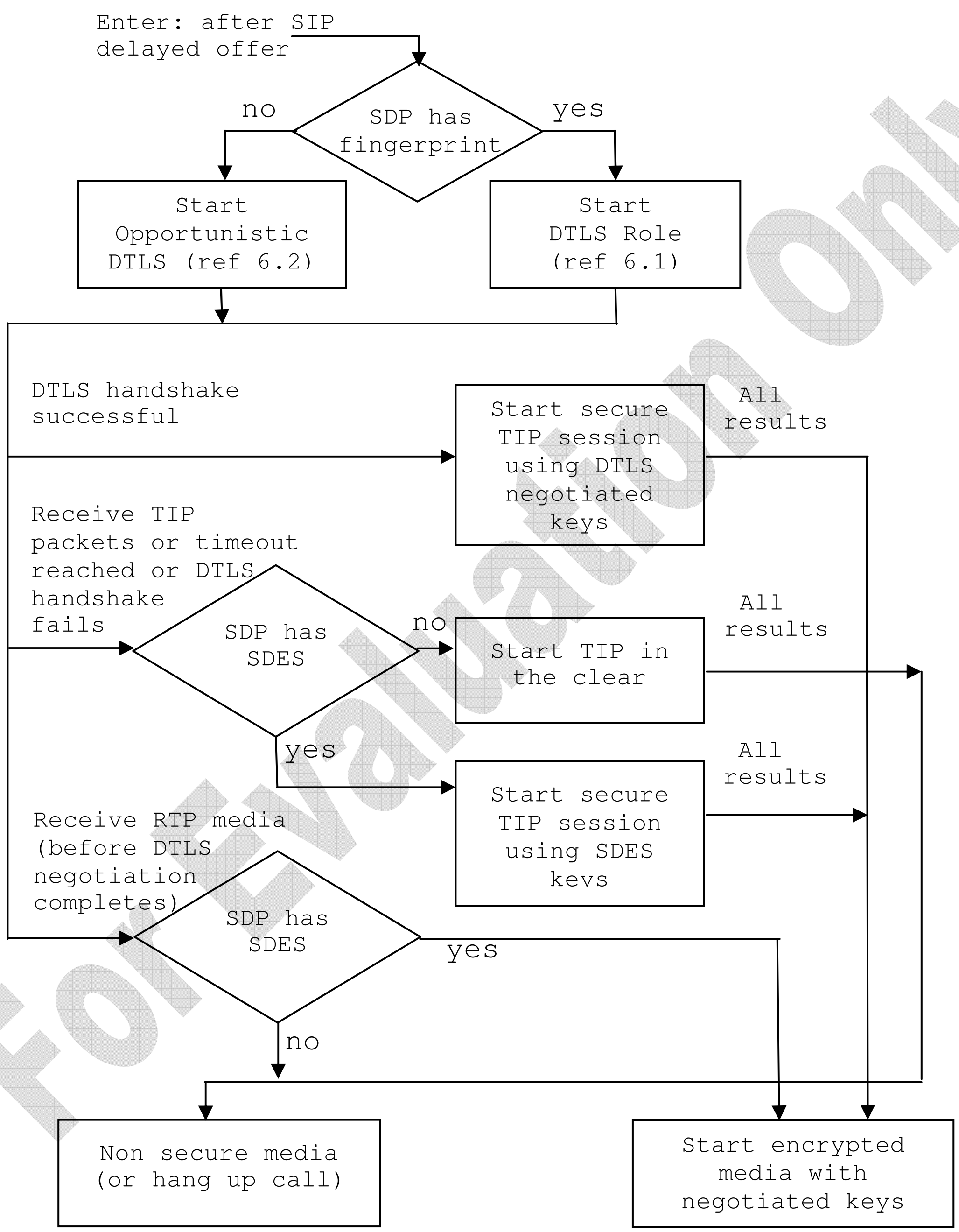

Figure 7: Cisco Device Secure Point-to-point Session Set-up Flow

6.1 SDP Negotiated DTLS-SRTP

- In addition to crypto attributes in the offer, the TIP device MAY include DTLS role and fingerprint attributes per media line. If both parties offer both the crypto and DTLS attributes, then later revisions of the UCM will choose either the crypto attribute or the DTLS attribute. Please refer to UCM documentation for details of negotiation of crypto and DTLS attributes [10, 22].
- In order to successfully negotiate DTLS-SRTP session in SDP while preserving the option of falling back to SDES, a secure TIP device MUST
  a. Include fingerprint attributes [19] of the certificate that will be sent in the DTLS handshake. The fingerprint received from the remote peer in SDP MUST match the fingerprint of the certificate received from the remote peer in the DTLS handshake.
  b. Include crypto attributes [19], also reference section 7 for related requirements.
  c. Signal RTP as the transport protocol for both DTLS-SRTP and SDES (an exception to the IETF specification [19]).

Figure 8: Sample video SDP Offer

```
Example DTLS-SRTP or SDES offer:
v=0
o=- 1181923068 1181923196 IN IP4 ua1.example.com
s=example1
c=IN IP4 ua1.example.com
t=0 0
m=video 26618 RTP/SAVP 112
b=TIAS:4000000
a=rtpmap:112 H264/90000
a=fmtp:112 profile-level-id=420028;...
a=sendrecv
a=crypto:xxxxxx
a=setup:actpass
a=fingerprint: SHA-1 \
4A:AD:B9:B1:3F:82:18:3B:54:02:12:DF:3E:5D:49:6B:19:E5:7C:AB
```

d. The SDP ANSWER from the UCM will indicate either SRTP with SDES, DTLS-SRTP or RTP (then opportunistic DTLS-SRTP MUST be attempted for secure calls).

```
Example SDP answer with SRTP:
v=0
o=- 1181923068 1181923196 IN IP4 ua2.example.com
s=example1
c=IN IP4 ua2.example.com
t=0 0
m=video 26618 RTP/SAVP 112
b=TIAS:4000000
a=rtpmap:112 H264/90000
a=fmtp:112 profile-level-id=420028;...
a=sendrecv
a=crypto:xxxxxx
```

```
Example SDP answer with DTLS-SRTP:
v=0
o=- 1181923068 1181923196 IN IP4 ua2.example.com
s=example1
c=IN IP4 ua2.example.com
t=0 0
m=video 26618 RTP/SAVP 112
b=TIAS:4000000
a=rtpmap:112 H264/90000
a=fmtp:112 profile-level-id=420028;...
a=sendrecv
a=setup:active
a=fingerprint: SHA-1 \
FF:AD:FF:B1:3F:82:18:3B:54:02:12:DF:3E:5D:49:6B:19:E5:7C:AB
```

- If the DTLS role has been negotiated, a secure TIP device MUST consider the session to be an SRTP session AFTER the DTLS handshake is completed using the negotiated role. Then the TIP device MUST:
  a. Derive the SRTP and SRTCP master keys used for the encryption and authentication of all transmitted SRTP and SRTCP packets using the write key associated with the DTLS-SRTP session.
  b. Derive the SRTP and SRTCP master keys used for the decryption and authentication of all transmitted SRTP and SRTCP packets using the read key associated with the DTLS-SRTP session.
  c. Use the crypto algorithm negotiated in the server session for SRTP and SRTCP packet authentication and decryption.
  d. Then, start TIP protocol negotiation.

6.2 Opportunistic DTLS-SRTP session Establishment

- When using opportunistic DTLS-SRTP establishment (hence, DTLS roles were not negotiated in SDP during call set up), a secure TIP device MUST start two DTLS-SRTP [19] sessions [a TIP device MUST act as a DTLS client for one of the DTLS-SRTP sessions while acting as a DTLS server for the other] for each negotiated media line irrespective of whether the AVP or the SAVP profiles were negotiated during call setup. Note that this is unlike a SDP negotiated DTLS-SRTP, where only a single session per media-line is established.
- The reception of an RTCP APP packet with subtype 1 on an RTP or SRTP session prior to the reception of a valid DTLS-SRTP communication over the same session MUST result in the termination of all DTLS-SRTP sessions associated with the (S)RTP session. Following the DTLS-SRTP termination, TIP negotiation MUST start using RTP/AVP or RTP/SAVP based on what was negotiated during call setup.
- Once BOTH of the DTLS-SRTP sessions associated with an (S)RTP session has succeeded, the TIP device MUST consider the session to be an SRTP session and MUST:
    - Derive the SRTP and SRTCP master keys used for the encryption and authentication of all transmitted SRTP and SRTCP packets using the write key associated with the DTLS-SRTP client session.
    - Use the crypto algorithm negotiated in the client session for SRTP and SRTCP packet authentication and encryption
    - Derive the SRTP and SRTCP master keys used for the authentication and decryption of all received SRTP and SRTCP packets using the read key associated with the DTLS-SRTP server session.
    - Use the crypto algorithm negotiated in the server session for SRTP and SRTCP packet authentication and decryption
    - Start TIP protocol negotiation.
    - If one or both of the DTLS-SRTP sessions associated with an (S)RTP session fails, the TIP device MUST start its TIP negotiation and operate in an RTP/AVP or RTP/SAVP mode based on what was negotiated during call setup.
- Following a hold/resume transaction, a secure TIP device MUST re-start new opportunistic DTLS-SRTP sessions prior to establishing the new secure TIP session.

6.3 Security downgrade

A secure TIP session could be downgraded into non-secure. This is typically a scenario involving a multi-point session where a non-secure device joins the secure conference turning it non-secure, assuming the session has not been configured to be a secure-only conference. Two mechanisms can be used to turn a TIP session into a non-secure session.

1. The first mechanism involves putting the SIP session on hold and immediately resuming it as a non-secure SIP session. The downgrading device does not start any DTLS sessions but rather start its TIP sessions following call setup causing the remote device DTLS sessions to be aborted and non-secure TIP sessions to be established
2. The second mechanism involves continuing to use SRTP and SRTCP on the secure TIP session. However a TIP Notification message indicating the change in the security state of the session is sent by the downgrading device causing the remote device to treat the TIP session as non-secure in all aspects except its media transport (see 5.4 for more details on TIP security state notification)

6.4 Security State Notification

Secure TIP devices MAY transmit a TIP NOTIFY message following the establishment of a secure TIP session to indicate that despite the use of SRTP, the TIP session MUST be considered as non-secure by the remote peer.

Secure TIP devices MAY transmit a TIP NOTIFY message to revert an SRTP TIP session security state from non-secure and back to secure. Secure TIP devices MUST NOT transmit a TIP NOTIFY message over a non-secure TIP session.

Secure TIP devices MUST be prepared to receive a TIP NOTIFY message and MUST inform the user of the security state indicated by the NOTIFY message.

TIP devices receiving a TIP NOTIFY message over a non-secure TIP session MUST ignore those.

6.5 Sample secure TIP setup call flow (Opportunistic DTLS-SRTP)

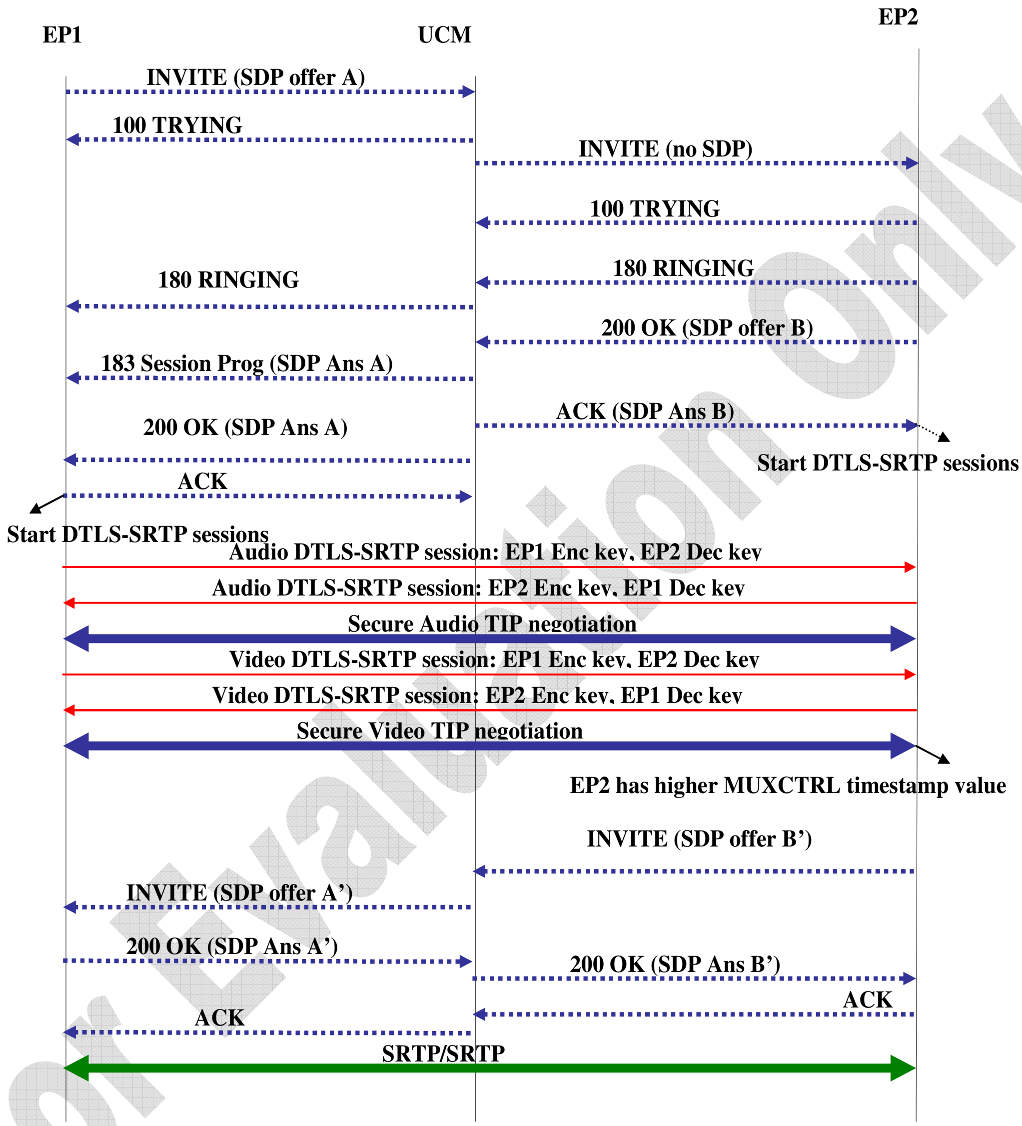

Figure 9: Secure TIP Session Set-up

Figure 9 shows a call setup between two secure TIP endpoints. The first INVITE transaction negotiates a single audio and video media line with H264 and AAC. The negotiated media line might have an AVP or an SAVP RTP profile.

Once the initial call setup completes, and if DTLS roles were not negotiated in SDP, then each secure TIP device starts two DTLS-SRTP sessions for each media line irrespective of whether the AVP or the SAVP profile was negotiated for the media line.

1. If DTLS roles were negotiated in SDP, then the encryption and decryption key is derived from the same session.
2. If DTLS roles were NOT negotiated in SDP, the DTLS-SRTP session for which the TIP device acts as a client will be used by the device to generate its SRTP and SRTCP master keys used for encryption. The crypto algorithm negotiated in the session will also be used for transmitted packet authentication and encryption
3. The DTLS-SRTP session for which the TIP device acts as a server will be used by the device to generate the SRTP and SRTCP master keys used for decryption. The crypto algorithm negotiated in the session will also be used for received packets authentication and decryption

Once both DTLS-SRTP sessions have completed successfully for a specific (S)RTP session, the TIP device starts its secure TIP negotiation.

From now on the call flow should follow the same pattern as described in 4.1, except that SRTP and SRTCP packets are transmitted and received rather than RTP and RTCP packets.

6.6 Sample secure to non-secure TIP setup call flow

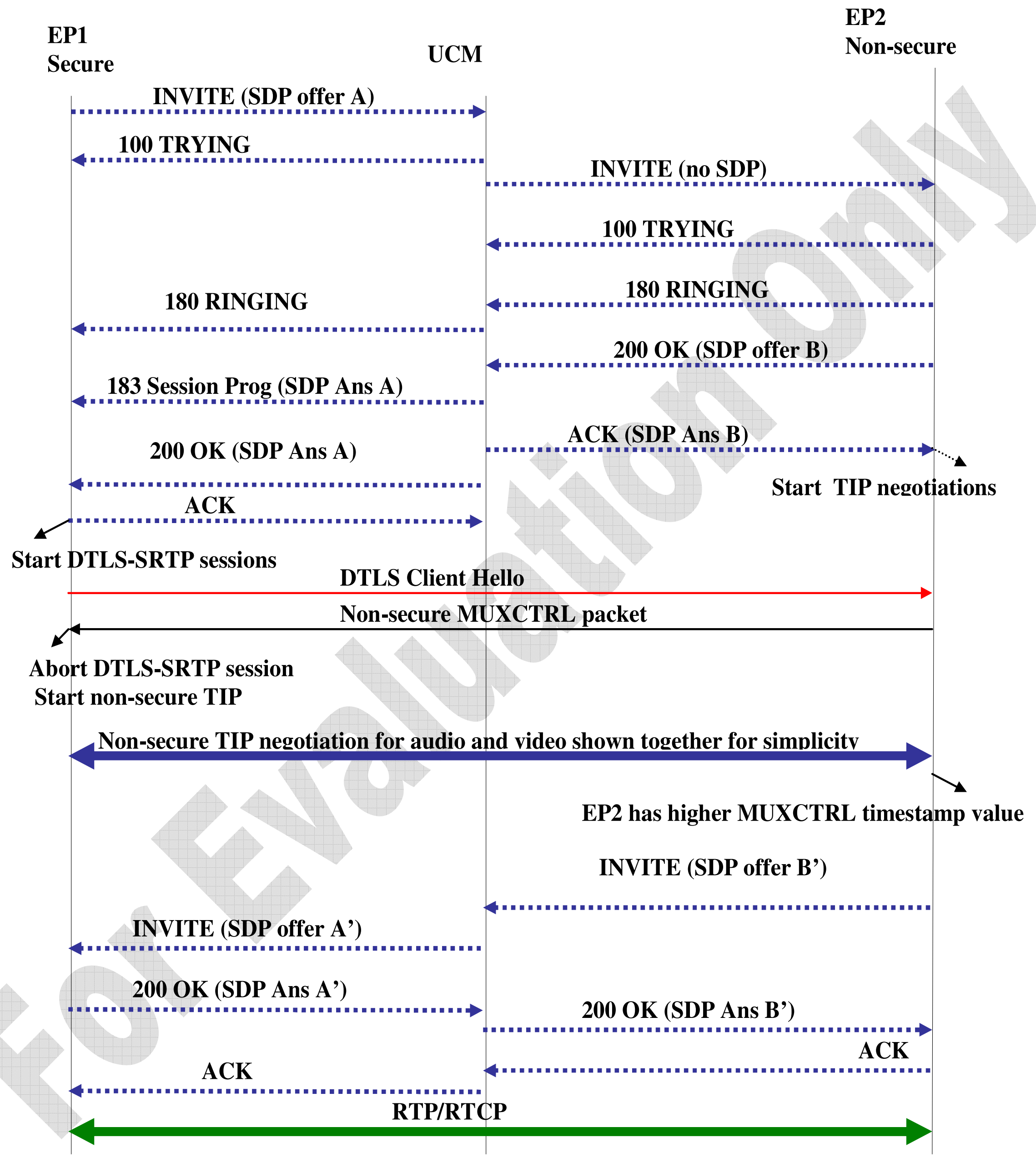

Figure 10: Secure TIP session setup

Figure 10 shows a call setup between a secure TIP endpoint and a non-secure TIP endpoint. The first INVITE transaction negotiates a single audio and video media line with H.264 and AAC. Note that if the secure endpoint has offered SAVP RTP profile in its SDP then it MUST support SRTP fallback as specified by the UCM SIP device interface specification appropriate for the SIP device type of the TIP endpoint.

Once the initial call setup completes, the secure TIP device starts two DTLS-SRTP sessions for each media line. Conversely, if an endpoint is not going to support security, it will start a non-secure TIP negotiation.

Upon reception of the MUXCTRL RTCP APP packet coming from the non-secure endpoint, the secure TIP endpoint MUST abort its DTLS-SRTP sessions and start its TIP negotiation.

From now on the call flow should follow the same pattern as described in 4.1.

7 Secure switched multi-point TIP Session establishment

TIP endpoints can using this profile to communicate securely in a switched multi-point TIP session, as in a multipoint session where encryption is not terminated by the multipoint controlling device and group key information is required. To participate in such session:

1. TIP devices MUST support all requirements for establishing secure point to point TIP sessions as specified in section 5
2. During the establishment of the secure TIP session, a TIP device MUST indicate the ability to receive group EKT security parameters on both the audio and video sessions
3. Once the secure TIP session is established and the remote multi-point device indicates support for transmission of group EKT parameters, the TIP device MUST wait till the TIP SPIMAP packet is received on the secure TIP session prior to starting any media transmission on that session.
4. Once the TIP SPIMAP packet is received, the TIP device MUST operate in an EKT mode. See section 7.2 for EKT requirements

7.1 Sample multi-point session secure TIP setup call flow

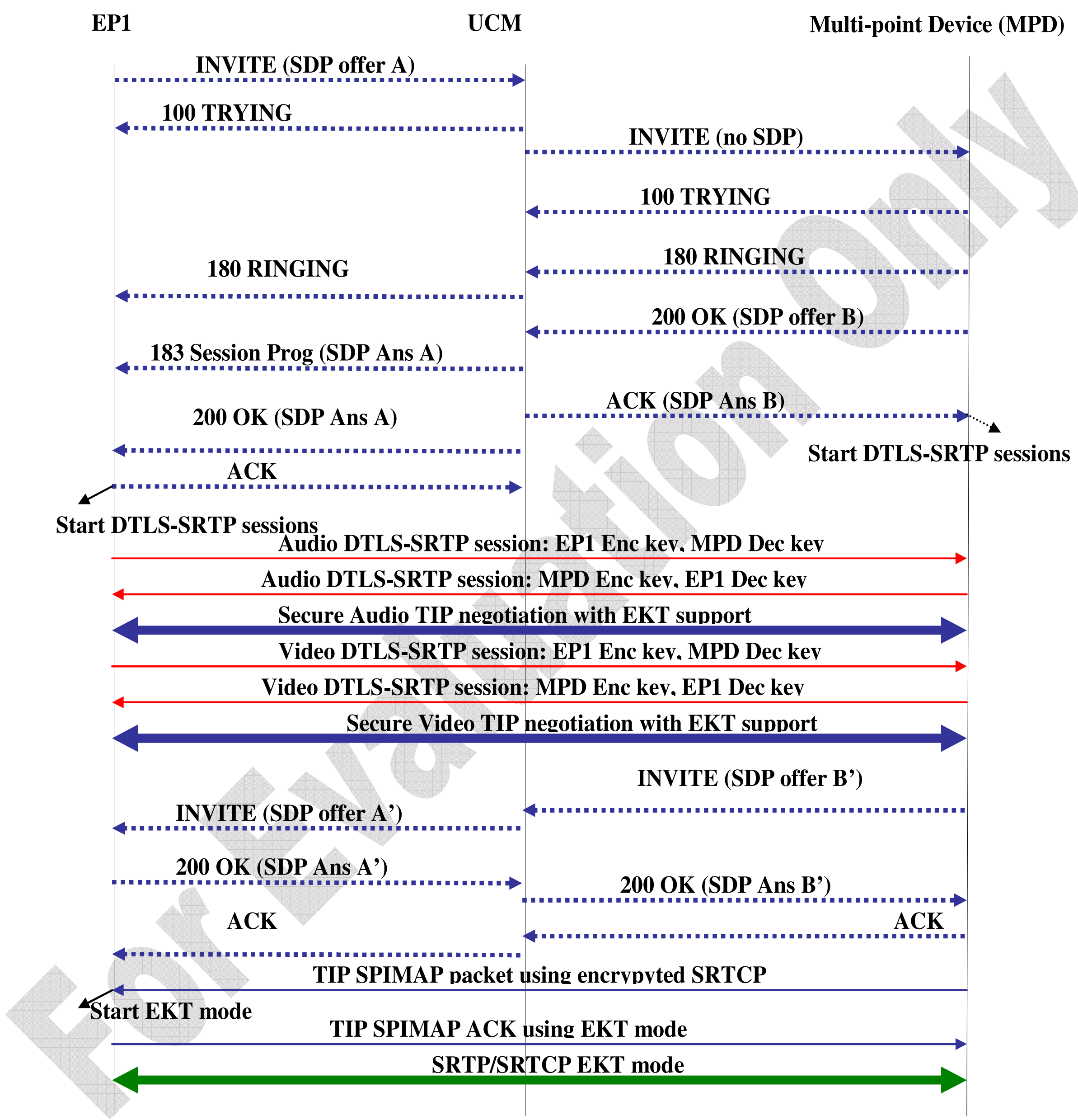

Figure 11: Secure TIP Session setup

Figure 11 shows a call setup between a secure TIP device that supports EKT and a secure multipoint server that supports EKT. The initial call setup flow and the establishment of opportunistic DTLS-SRTP sessions are exactly the same as the case of a point to point secure session. However in this flow the ability to support group EKT security parameters are established during the secure TIP session establishment phase. Unlike the secure point to point session establishment call flow, the TIP device will not start its secure media transmission after the bit rate update reINVITE is finished. The secure media transmission MUST wait until it receives the SPIMAP packet from the multi-point device prior to starting its secure media transmission in EKT mode.

8 Initial TIP signaling

TIP devices MUST adhere to all of the MUST requirements specified in the TIP document [11] as well as the MUST requirements in this document.

8.1 MUXCTRL Packet

1. In a video session, the AVPF profile MUST be supported and signaled in the MUXCTRL packet as specified in the TIP specs [11] regardless if this is a non-secure or secure session. This in turn mandates support for the transmission and reception of the RTCP video feedback packets as specified in the TIP specs.
2. In an audio-only session, the profile MUST be AVP, regardless if this is a non-secure or secure session.
3. The options field of the MUXCTRL packet MUST be set to 0x00.
4. Triple screen devices MUST set the number of transmitted streams to 6 for a video session supporting legacy video streams and to 3 for a video session not supporting legacy streams.
5. Single screen devices MUST set the number of transmitted streams to 2 for a video session supporting legacy streams and to 1 for a video session not supporting legacy.
6. Devices MUST set the number of shared streams for the video session to zero if presentation is not supported, to 1 if a single presentation stream is supported and to 2 if dual stream presentation is supported.
7. Devices MUST set the number of shared streams for audio sessions to zero.
8. Triple screen devices MUST set the number of transmitted audio streams to 5 when both legacy and presentation are supported, to 4 when no legacy streams are supported and to 3 when no presentation and no legacy streams are supported.
9. Single screen devices MUST set the number of transmitted audio streams to 3 when both legacy and presentation are supported, to 2 when no legacy streams are supported and to 1 when no presentation and no legacy streams are supported.

10. The number of received video streams MUST be set to 3 for triple screen devices and to 1 for single screen devices in MUXCTRL.
11. The number of received audio streams MUST be set to 5 for devices supporting legacy streams and set to 4 for devices not supporting legacy streams in MUXCTRL.
12. For triple screen devices, the transmit video positions MUST only contain: center, left, right, legacy center, legacy left and legacy right, when legacy is supported and MUST only contain: center, left and right when legacy is not supported.
13. For single screen devices, the transmitting video positions MUST only contain: center, and legacy center, when legacy is supported and MUST only contain: center when legacy is not supported.
14. The table 3 below shows how the shared video positions MUST be set based on the level of auxiliary (AUX), or presentation, video support required for the session.

Supported AUX Video fps	Number of Shared Video Streams	Shared Video MUX Position
none	0	0
1 fps only	1	1/5
1 or 5 fps only	1	1/5
1 or 5 or 30 fps only	1	1/5 and 30
1 or 5 and 30 fps	2	1/5 and 30

Table 3: Video AUX capability to shared video positions mapping

15. For triple screen all transmit audio positions (left, right, center, legacy mix and auxiliary) MUST be set when legacy and presentation is supported. The legacy mix MUST be unset when legacy is not supported and the auxiliary position MUST be unset when presentation is not supported
16. For single screen all transmit audio positions (center, legacy mix and auxiliary) MUST be set when legacy and presentation is supported. The legacy mix MUST be unset when legacy is not supported and the auxiliary position MUST be unset when presentation is not supported
17. The receive video positions MUST contain center, left, right for tripe screen devices and MUST contain center for single screen devices

18. For both single and triple screen devices, the receive audio positions MUST contain center, left, right, legacy and auxiliary when legacy streams are supported and MUST contain center, left, right and auxiliary when legacy streams are not supported
19. TIP endpoints SHOULD set the conference identifier field, or if the identifier is not known, set the field to 0.

8.2 MediaOpts Packet

1. TIP endpoints MUST indicate support for transmission of Audio Activity Metric and MUST NOT indicate support for reception of Audio Activity Metric.
2. TIP endpoints MUST support the reception of Audio Dynamic Output Channel.
3. TIP endpoints MUST NOT indicate support for transmission of the Audio Dynamic Output Channel and MUST set output positions in all TIP-CSRCs to zero.
4. TIP endpoints MAY indicate support and use of G722 encoding for legacy audio streams.
5. TIP endpoints MUST indicate support for the transmission of the video Refresh flag but MUST NOT indicate support for the reception of the flag.
6. TIP endpoints MUST indicate support for In-band Parameter Sets (SPS/PPS) both for transmission and reception.
7. TIP endpoints MAY signal support for CABAC (Please see exception note in 11.1)
8. TIP endpoints MAY signal support for LTRP and GDR.
9. TIP endpoints MUST signal support for the maximum auxiliary frame rate that they can support
10. TIP endpoints are RECOMMENDED to support 30 fps as their maximum video auxiliary frame rate
11. TIP endpoints MAY signal support for 8x8 H264 video transforms (eg, HiP) for Restricted Media negotiated resolutions only.
12. TIP endpoints MUST signal support for the transmission of Restricted or Unrestricted media for 1080P, 720P, XGA 1/5 fps and XGA 30 fps, regardless of what was signaled in SDP. Please reference sections 13 and 14 for guidelines about Restricted or Unrestricted video media interoperability strategies.
13. TIP endpoints MAY signal a Satellite deployment Transmitter profile if it is known that this endpoint is connected using a satellite link.

8.3 TIP Negotiation and general mid-session signaling rules

1. A TIP endpoint MUST NOT start media transmission until the TIP negotiation has

finished. TIP negotiation is finished once the following has completed:

a. Local MUXCTRL packet has been sent and a remote ACK was received.

b. Remote MUXCTRL has been received and an ACK was sent.

c. Local MediaOpts packet has been sent and an ACK was received.

d. Remote MediaOpts packet has been received and ACK was sent.

Note that, TIP negotiation can, and will, complete at one peer prior to the other peer(s). Implementations MUST be able to handle such scenarios.

2. TIP endpoints SHOULD start media transmission after the reINVITE transaction updating the video and audio bitrates have completed.
3. Secure TIP endpoints operating in EKT mode MUST NOT start media transmission till after the SPIMAP is received and authenticated
4. The number of transmitted streams MUST NOT exceed the number of streams the remote peer is capable of receiving
5. Only positions indicated as supported by the remote peer can be used in the CSRC received position
6. Typically the number of transmitted streams is the minimum of what the local endpoint can transmit and what the remote peer can receive. The exception is for legacy streams, which MUST NOT be transmitted, except if the remote peer indicates its capability to receive the legacy streams AND the remote peer indicates that it is a multi-point device by signaling the isfocus parameter in its MUXCTRL packet.
7. As an example, the maximum number of streams that will be sent and received when a triple screen TIP endpoint supporting a single presentation/AUX (ie, no simulcast of AUX video) stream communicates with a remote TIP peer are as follows:
    a. Remote peer is a triple screen TIP endpoint
        i. Transmitted audio streams: 4
        ii. Received audio streams: 4
        iii. Transmitted video streams: 4
        iv. Received video streams: 4
    b. Remote peer is a single screen TIP endpoint
        i. Transmitted audio streams: 4
        ii. Received audio streams: 2
        iii. Transmitted video streams: 2
        iv. Received video streams: 2
    c. Remote peer is a multi-point device requiring legacy stream support:
        i. Transmitted audio streams: 5

ii. Received audio streams: 5
iii. Transmitted video streams: 7
iv. Received video streams: 4

d. Remote peer is a multi-point device requiring no legacy stream support:
   i. Transmitted audio streams: 4
   ii. Received audio streams: 4
   iii. Transmitted video streams: 4
   iv. Received video streams: 4

8. As an example, the maximum number of streams that will be sent and received when a single TIP endpoint supporting a single presentation/AUX (ie, no simulcast of AUX video) communicates with a remote TIP peer are as follows:
   a. Remote peer is a triple screen TIP endpoint
      i. Transmitted audio streams: 2
      ii. Received audio streams: 4
      iii. Transmitted video streams: 2
      iv. Received video streams: 2
   b. Remote peer is a single screen TIP endpoint
      i. Transmitted audio streams: 2
      ii. Received audio streams: 2
      iii. Transmitted video streams: 2
      iv. Received video streams: 2
   c. Remote peer is a multi-point device requiring legacy stream support:
      i. Transmitted audio streams: 3
      ii. Received audio streams: 5
      iii. Transmitted video streams: 3
      iv. Received video streams: 2
   d. Remote peer is a multi-point device requiring no legacy stream support:
      i. Transmitted audio streams: 2
      ii. Received audio streams: 4
      iii. Transmitted video streams: 2
      iv. Received video streams: 2
9. When a triple screen endpoint communicates with a single screen endpoint, the sender

has more positions and streams to transmit than the remote peer can receive. In this case, the triple screen endpoint SHOULD switch between the signals coming from its three main video inputs, based on criteria such as the associated audio input level.

10. TIP endpoints MUST NOT generate a second MUXCTRL packet within the same TIP session.
11. TIP endpoints MUST NOT generate a second MediaOpts packet within the same TIP session[4] in which there have been changes to any media capability other than CABAC, GDR, LTRP, auxiliary (AUX) frame rate and High Profile (HiP). Transmission of mid-session MediaOpts packets is typically expected from MCU devices following this endpoint profile.
12. TIP endpoints MUST support the ability to receive MediaOpts packets that change the CABAC, GDR, and LTRP, auxiliary (AUX) frame rate or High Profile (HiP) indicators. Any other capability changes communicated in mid-session MediaOpts packet MUST be discarded.
13. TIP endpoints MUST silently discard any received TIP packet with an unrecognized RTCP APP packet subtype. Secure TIP endpoints MAY use such packets to abort opportunistic DTLS-SRTP sessions.
14. TIP endpoints MUST be prepared to receive compound RTCP packets with multiple RTCP APP packets. TIP endpoints MUST process all RTCP packets in a compound packet
15. TIP endpoints MUST silently discard any received TIP packet with an RTCP APP name other than “xcts” . Secure TIP endpoints MAY use such packets to abort opportunistic DTLS-SRTP sessions.
16. TIP endpoints MUST respond to ECHO requests as specified in the TIP specs.
17. TIP endpoints MUST support the reception of the media flow control request and process it according to the TIP specs [11].
18. TIP endpoints MUST NOT generate media flow control requests if the remote peer is a multi-point device.
19. TIP endpoints MUST support the reception of the video refresh request and process it according to the TIP specs.
20. TIP endpoints MUST be able to receive a video refresh request and process it according to the TIP specs [11].
21. TIP endpoints MAY generate video refresh requests though it is NOT RECOMMENDED to use this mechanism when packet losses are encountered. Alternatively the video feedback mechanism should be used for reporting loss.

[4] A TIP session is initiated at call setup and after a hold or resume transaction. It is identified by a unique TIP SSRC value.

9 Presentation Signaling

Permission to present is requested and released using the REQTOSEND TIP request on the video TIP session. The REQTOSEND ACK is used to grant or deny the request to send.

REQTOSEND can be used to request permission to transmit dual presentation video streams, one stream at 30 fps and the other at 5 or 1 fps. The REQTOSEND ACK can grant one of the two video streams or both. Note that the REQTOSEND ACK packet is not the standard TIP ACK packet [10].

TIP endpoints MUST NOT transmit REQTOSEND messages over the TIP audio session.

9.1 Transmitting and receiving permission-to-present requests in a multi-point session

To request permission to present, a TIP endpoint MUST send a REQTOSEND packet on the video session setting the flag parameter to indicate a "request to transmit" and setting a single or dual auxiliary (aka, AUX/Presentation) video positions in the video position parameter.

If the request is granted, the multi-point device will send a REQTOSEND ACK packet on the video session setting the video positions that are granted. The endpoint SHOULD start its transmission once it has received the REQTOSEND ACK packet.

If the multi-point device denies the request due to conflict with another endpoint that already has permission to transmit, the multi-point device will deny the request by sending a REQTOSEND ACK packet setting video position to 0.

A TIP endpoint receiving a REQTOSEND MUST grant the request as long as it includes a video position satisfying the current TIP session's negotiated capability. The request is granted via ACK-ing it and starting the presentation receivers. If the endpoint was already transmitting presentation prior to receiving the REQTOSEND, it MUST stop its presentation transmission on both audio and video channels, ACK the request and start its presentation receivers.

TIP endpoints SHOULD only deny a REQTOSEND packet if the REQTOSEND does not include a video position that satisfies the current TIP session's negotiated capabilities.

A presentation (AUX) conflict can occur when a REQTOSEND is received from the multipoint device requesting permission to present at the same time that a local REQTOSEND has been sent from the endpoint. In that case, the TIP endpoint should assume it is the winner of the conflict and wait for the REQTOSEND ACK to come from the remote peer before sending media on the presentation/AUX steam.

9.2 Transmitting and receiving permission-to-present requests in a point-to-point session

To request permission to present, a TIP endpoint MUST send a REQTOSEND packet on the video session setting the flag parameter to indicate a “request to transmit” and setting a single or dual auxiliary (aka AUX/Presentation) video positions in the video positions parameter.

The request is granted once the remote endpoint ACKs the request setting the granted video positions. Once the ACK packet is received, the endpoint can start transmitting its presentation video and audio.

A presentation conflict can occur when a REQTOSEND is received from the remote peer requesting permission to present at the same time that a local REQTOSEND has been sent. An endpoint should resolve this conflict by assuming it is the winner if it has the lower NTP timestamp as part of its REQTOSEND. In the case where the NTP timestamp of the two endpoint’s REQTOSEND is the same, the endpoint with the lower TIP video session’s SSRC should assume it is the winner of the conflict. To achieve such a result the following behavior is a MUST in the case of a presentation conflict:

1. Before it considers itself the new presenter, the endpoint that assumes it is the winner of the conflict MUST wait for the REQTOSEND ACK to come from the remote peer.
2. The endpoint that lost the conflict MUST send an ACK for the remote REQTOSEND it has received.

A TIP endpoint receiving a REQTOSEND MUST grant the request as long as it includes a video position satisfying the current TIP session’s negotiated capability and as long as it is does not cause a conflict. The request is granted via ACK-ing it and starting the presentation receivers. If the endpoint was already transmitting presentation prior to receiving the REQTOSEND request, it MUST stop its presentation transmission of both audio and video and ACK the request.

TIP endpoints SHOULD only deny a REQTOSEND if it does not include a video position that satisfies the current TIP session’s negotiated capabilities.

9.3 Transmitting and receiving permission-to-present release requests

To release permission to present, the endpoint MUST send a REQTOSEND packet on the video session, setting the flag parameter to indicate stop transmission indication and setting the video positions to the positions that were included in the REQTOSEND used to request permission.

Note that in a multi-point session, an endpoint which has previously sent a permission-to-present that was not granted MUST still release the permission-to-present when it no longer transmits the presentation.

In point-to-point scenarios when the endpoint receives permission-to-present release request, it MUST ACK the request. If the endpoint is interested in presentation transmission then it MUST send a new REQTOSEND requesting permission to present.

9.4 Receiving a request-to-present grant in a multi-point session

A multi-point device can grant permission to present to an endpoint after it has initially denied the request. This is the case when a conflict occurs between multiple endpoints requesting to present at the same time, which results in the multi-point device granting only one request while denying others. Once the winning endpoint releases its permission to present, the multi-point might grant the permission to present to the endpoint whose request was previously denied. In this case, the endpoint will receive a REQTOSEND indicating a permission-to-present release from the multi-point device. After ACK-ing the release request, if the endpoint is still interested in presentation transmission, it MUST send a REQTOSEND indicating permission-to-present.

10 Presentation frame rate negotiation

10.1 Initial frame rate negotiation

As specified in the TIP specifications [11], the MediaOpts packet is used for frame rate negotiation. It is RECOMMENDED that TIP endpoints support and signal a 30 fps capability in the MediaOpts packet. Any endpoint signaling 30 fps in its MediaOptions MUST also include both the 5/1 fps and the 30 fps auxiliary (aka AUX/Presentation) video positions in the shared positions of the MUXCTRL packet sent over the video TIP session.

Once a TIP endpoint has signaled its initial presentation frame rate capability, it MUST NOT change it during the same TIP session. The negotiated initial frame rate is calculated as the minimum of what was signaled by both peers.

Note that the negotiated initial frame rate is what will be used to determine the presentation bit rate included in the SDP video b=TIAS [16] when the number of shared streams negotiated in the MUXCTRL is 1. This bit rate value will remain the same irrespective of whether or not the frame rate has changed later within the TIP session.

It is important to note that when an endpoint signals a frame rate capability, such as 30 fps, this endpoint MUST also support all lower frame rates such as 5 and 1 fps. However, an endpoint transmitting a single presentation stream with a negotiated frame rate of 30 fps MUST NOT transmit presentation using a different frame rate such as 5 fps without first receiving a MediaOpts packet negotiating the new frame rate and getting an ACK for a REQTOSEND specifying a position that satisfies the new frame rate.

10.2 Signaling a new frame rate

TIP endpoints MUST NOT dynamically change their signaled frame rate. However TIP endpoints MUST support receiving mid-session MediaOpts packets that can lower or increase the negotiated frame rate. Note that the negotiated frame rate MUST NOT be higher than the initially negotiated frame rate in the TIP session. Once the endpoint ACKs the MediaOpts packet, any new presentation transmitted MUST use the new frame rate. Any future REQTOSEND packet sent MUST at least have one position that satisfies the last frame rate signaled. A transmitter receiving a new frame rate that is not satisfied by its latest REQTOSEND request MUST initiate a new REQTOSEND prior to starting transmission using the new rate.

To illustrate, below are four examples for an endpoint that supports a single presentation stream and for an initially negotiated frame rate of 30 fps.

In the first example, the endpoint starts presentation transmission at 30 fps and then gets downgraded to 5 fps. Note that the endpoint needed to transmit a new REQTOSEND following the reception of the MediaOpts as the previous REQTOSEND did not include a position that satisfies the new frame rate.

In the second example, the endpoint starts transmission at 5 fps and then gets downgraded to 1 fps. Note that the endpoint switched to the new frame rate once it has received the MediaOpts packet as the last REQTOSEND request it sent had a position that satisfies the new frame rate.

In the third example, the endpoint starts receiving at 30 fps and then gets downgraded to 5 fps. Note that no MediaOpts was sent to the endpoint to indicate the downgrade as a frame rate of 30 fps translates to 30 fps on the 30 fps auxiliary position and 5 fps on the 5/1 fps auxiliary position. Hence this downgrade took place by using a REQTOSEND.

In the last example, the endpoint starts receiving 30 fps and then gets downgraded to 1 fps. Note that unlike the previous example a MediaOpts packet is needed to downgrade the frame rate to 1 fps on the 5/1 fps auxiliary position. The MediaOpts is then followed by a REQTOSEND that negotiated the use of the 5/1 position.

30 fps to 5 fps transmitter downgrade

The endpoint sends a REQTOSEND indicating 30 fps auxiliary position
The endpoint receives a REQTOSEND ACK
The endpoint starts transmission of a 30 fps presentation using the 30 fps auxiliary position
The endpoint receives a MediaOpts signaling a new frame rate of 5 fps
The endpoint stops its transmission and ACKs the MediaOpts request
The endpoint sends a REQTOSEND specifying the 1/5 fps auxiliary position
The endpoint receives a REQTOSEND ACK
The endpoint starts transmission of 5 fps presentation using the 1/5 fps auxiliary position

5 fps to 1 fps transmitter downgrade

The endpoint sends a REQTOSEND indicating 5 fps auxiliary position
The endpoint receives a REQTOSEND ACK

The endpoint starts transmission of a 5 fps presentation using the 1/5 fps auxiliary position
The endpoint receives a MediaOpts signaling a new frame rate of 1 fps
The endpoint stops its transmission and ACKs the MediaOpts request
The endpoint starts transmission of 1 fps presentation using the 1/5 fps auxiliary position

30 fps to 5 fps receiver downgrade

The endpoint receives a REQTOSEND indicating 30fps auxiliary position
The endpoint transmits a REQTOSEND ACK
The endpoint starts receiving a 30 fps stream transmitted using the 30 fps auxiliary position
The endpoint receives a new REQTOSEND specifying the 1/5 fps auxiliary position
The endpoint transmits a REQTOSEND ACK
The endpoint starts receiving of 5 fps stream transmitted using the 1/5 fps auxiliary position

30 fps to 1 fps receiver downgrade

The endpoint receives a REQTOSEND indicating 30 fps auxiliary position
The endpoint transmits a REQTOSEND ACK
The endpoint starts receiving a 30 fps stream transmitted using the 30 fps auxiliary position
The endpoint receives a MediaOpts signaling a new frame rate of 1 fps
The endpoint ACKs the MediaOpts request
The endpoint receives a new REQTOSEND specifying the 1/5 fps auxiliary position
The endpoint transmits a REQTOSEND ACK
The endpoint starts receiving of 5 fps stream transmitted using the 1/5 fps auxiliary position

11 Audio mixing and synchronization

In a multi-point session, when audio streams are sent to a TIP triple screen endpoint, the multi-point device maintains a mapping between a source segment/position and a receive segment/position. In other words if a triple endpoint has received the audio from a single endpoint on its left segment, the triple endpoint will always receive the same single endpoint audio on its left segment.

As a result, a triple screen endpoint could receive three audio streams from a multi-point device all to be played on the same segment/position (e.g., left speaker). The receiver positions in the CSRCs of these audio streams will be different while the dynamic audio output position will be the same indicating that these are simultaneous streams to be mixed rather than alternative streams to be switched. The output positions in the CSRCs of these audio streams will be the same, indicating that these streams need to be mixed and played out together. A TIP endpoint MUST mix, prior to playing out, audio streams that have the same output position and different receiver positions.

In a multi-point session, or when communicating with a triple screen TIP endpoint, a single screen TIP endpoint will receive multiple audio streams. These audio streams will each have a different receiver position in the CSRC. The single screen endpoint must mix and play out these different audio streams.

Note that for all TIP endpoints some, or all, of the audio streams to be mixed could be coming from the same remote endpoint. Such audio streams can be identified as they share the same TIP-CSRC sampling clock ID. A TIP endpoint SHOULD use the audio SRs to synchronize audio between streams that share the same CSRC sampling clock ID.

12 Media

1. TIP endpoints MUST handle dynamic switches of SSRCs [4] and their associated CSRCs on received video and audio streams during a session. The CSRC receive position identifies that this new SSRC is to replace the previous SSRC that was associated with the CSRC with the same receive position.
2. TIP endpoints MUST change their audio and video streams' SSRCs and CSRCs after a hold/resume transaction.
3. Video SSRCs and CSRCs values MUST change when switching presentation streams from 5 fps to 1fps and vice versa
4. TIP endpoints MUST support remote RTP senders and receivers corresponding to the same position yet using different SSRCs. For example a TIP endpoint might receive an RTCP RR or an RTCP feedback packet from a center position receiver with a source SSRC different from the source SSRC of the center position's RTP packets.
5. TIP endpoints MUST support the flow control TIP packet, both for transmitting and receiving video streams. The receive flow control will be used by a multi-point device to inform the endpoint that no media will be transmitted by the server using the receive position specified in the flow control packet target CSRC. The transmit flow control will be used by the multi-point device to inform the endpoint that no media should be transmitted using the transmit position specified in the flow control packet target CSRC.
6. TIP Endpoints MUST always transmit media using all source positions negotiated in TIP, unless a transmit flow control packet to turn off that specific position was received.
7. TIP endpoints MUST transmit RTCP Sender Reports [4].
8. TIP endpoints MUST silently discard any unexpected non-RTP packets, such as DTLS

packets received on non-secure RTP sessions or STUN Keepalive packets received on video and audio RTP sessions.

9. TIP endpoints MUST silently discard any RTP packets with an unknown payload type.

12.1 Audio

1. All non-legacy audio streams MUST be transmitted using AAC-LD (see 13 for codec requirements) and using RFC 3640 with AAC-hbr mode [15]
2. TIP endpoints MUST transmit 2 AAC audio frames (10 msec per frame) msec in AAC audio packets
3. TIP endpoints MUST send audio activity metric byte for all AAC audio streams as specified in the TIP specs [11] when Audio Activity Metric support is negotiated in TIP.
4. TIP endpoints MUST set the Audio Activity Metric associated with presentation audio to a value of 99
5. TIP endpoints MUST NOT send the audio activity metric byte for G.711 or G.722 legacy streams
6. TIP endpoints MUST be able to deal with periods where no audio packets are received on one or more of its active audio positions without disrupting the call. The recommendation is to use video streams to detect network problems and act on them. Video streams will receive a receive flow control off packet when no video packets are available for transmission on a specific position.
7. TIP endpoints MUST be able to deal with receiving the same audio stream with discontinuous sequence numbers. Note that for video sequence numbers will only be discontinuous during an SSRC switch whereas in the case of audio this might not always be true.
8. TIP endpoints MUST support mixing of audio streams that are sent using the same output position and different receive positions
9. TIP endpoints SHOULD use received audio Sender Reports to synchronize audio between different SSRCs that share the same CSRC sampling clock ID.
10. TIP endpoints MUST generate accurate and frequent audio RTCP Sender Reports for audio sources to allow remote peers to perform cross-media synchronization
11. TIP endpoints MUST send all main audio streams using the same sampling clock ID
12. When dynamic output channels are not negotiated, TIP endpoints MUST include zero in the 4 TIP-CSRC bits corresponding to output positions

12.2 Video

1. TIP endpoints SHOULD include zero in the 4 video TIP-CSRC bits, corresponding to the dynamic output positions. TIP endpoints MUST accept non-zero TIP-CSRC bits.
2. All HD and auxiliary H264 video steams MUST be packetized using packetization mode "1" [7]. The following Network Abstraction Layer (NAL) types are supported:
    - 1 (non-IDR)
    - 5 (IDR)
    - 6 (SEI, only subtype 6 [recovery point] is supported for GDRs)
    - 7 (SPS)
    - 8 (PPS)
3. All 'legacy' CIF H264 video streams MUST be packetized using packetization mode 0 [7].
4. All video streams except for auxiliary MUST be transmitted using a 30 fps or 29.97 fps.
5. Auxiliary streams MUST be transmitted using the frame rate signaled in the TIP session.
6. TIP endpoints' video frames sampling instances, and their corresponding RTP timestamps, MUST always correspond to regular intervals that is based on the video clock rate and the video frame rate. For example a 30 fps TIP video stream using a 90KHz clock can only generate video frames every 3000 clock ticks. A 5 fps TIP video stream using a 90 KHz clock can only generate video frames every 18000 clock ticks.
7. TIP endpoints MUST send the video refresh flag for all video streams when support for video refresh flag is negotiated in TIP.
8. TIP endpoints MUST support in-band transmission of the SPS and PPS as the first two NALs in each I-frame. TIP endpoints MUST support in-band transmission of the SPS, PPS, and SEI recovery point NAL as the first three NALS in each GDR.
9. TIP endpoints MUST support transmission and reception of RTCP video feedback as specified in the TIP SPECS. [11]
10. TIP endpoints MUST transmit a repair frame as soon as a negative video feedback report is processed.

13 Restricted and Unrestricted Media Interoperability

13.1 Video multiplex MediaOpts field for bits 12 or 13

Section 4.2.5.11 of the TIP v7 protocol specification [11] defines the "Restricted Media" attributes as well as the device behavior expected for the combinations expressed in the bit "12" and "13" positions in the MEDIAOPTS for video streams. The tables shown below expand on that section of the specification for interoperability with Cisco TIP devices.

A value of "0" in the bit "12" and "13" positions indicates that the media must adhere to the list of media restrictions in section 4.2.5.11. A value of "1" in the bit "12" and "13" positions indicates there is no restriction on the media (it is "unrestricted" and does not require adherence to all those restrictions, except for Cisco TIP device constraints described below).

When a Cisco TIP device indicates it can receive "unrestricted" video media in TIPv7, it still has a couple of constraints per section 14.2-14.5. To summarize, each transmitted NAL MUST be less than 3320 bytes in size and no CABAC compression or 8x8 transforms can be used.

The combination that will provide the greatest compatibility with Cisco TIP devices as shown in the left hand side of the first table below is when a third-party TIP device implementer can support RX=1 and TX=0 for a particular resolution. This indicates that an implementer's transmitter will adhere to the list of TIP media restrictions for video it sends, but can receive unrestricted video media.

Additional combinations are shown in the second and third tables on the next page to illustrate where incompatibilities may exist with Cisco TIP device.

Cisco TIP devices per software release	RX	TX	Third-party TIP indicates RX=1 and TX=0
CTS 1.7.x all HD resolutions CTS 1.8.x or later for 1080p CTMS 1.7.x or later	0	0	MUST only send restricted media per TIPv7 spec
TelePresence Server 2.2 CTS 1.8.x or later for 720p only CTMS 1.7.x or later	1	0	SHOULD send restricted video media [5]
CTMS 1.8.x or later (in some cases only)	1	1	SHOULD send restricted video media [5]

Additional combinations are shown where incompatibilities may exist with Cisco TIP devices in some cases or with certain software releases.

Cisco TIP devices per software release	RX	TX	Third-party TIP indicates RX=0 and TX=0
CTS 1.7.x all HD resolutions CTS 1.8.x or later for 1080p CTMS 1.7.x or later	0	0	MUST only send restricted media per TIPv7 spec
TelePresence Server 2.2 CTS 1.8.x or later for 720p only CTMS 1.7.x or later	1	0	SHOULD send restricted video media [5]
CTMS 1.8.x or later (in some cases only)	1	1	Incompatible

Cisco TIP devices per software release	RX	TX	Third-party TIP indicates RX=1 and TX=1
CTS 1.7.x all HD resolutions CTS 1.8.x or later for 1080p CTMS any release	0	0	Incompatible
TelePresence Server 2.2 CTS 1.8.x or later 720p only CTMS 1.7.x or later	1	0	Can send unrestricted (eg, baseline H264) media
CTMS 1.8 (in some cases)	1	1	Can send unrestricted (eg, baseline H264) media

[5] When connected in multipoint calls, such as with the CTMS, it is possible that during any given session that the state may have to be changed to RX=0. The third-party TIP device SHOULD only send restricted media given that possibility, even though the TIP peer indicated it could receive unrestricted media. This provides the greatest compatibility and predictability with Cisco TIP devices across the scenarios.

14 TIP v8 “Unrestricted Media Constraints” Interoperability

14.1 Video multiplex MediaOpts field for bit 15

Section 4.2.5.12 of the TIP v8 protocol specification [11] defines the "Unrestricted Media Constraints" attribute as well as the device behavior expected for the combinations expressed in the bit "15" position in the MEDIAOPTS for video streams in TIP.

Section 6.3 of the TIP v8 protocol specification adds further information about maximizing compatibility with other devices. Only Cisco CTS 1.10 release will set the receiver to require Unrestricted Media Constraints (RX=1) for the 720p resolution and the shared auxiliary (a.k.a. presentation) positions for up to an XGA resolution. Like the releases before it, the CTS 1.10 release requires receipt of “Restricted Media” for the 1080p resolution.

Like what is recommended in Section 6.3 of the TIP v8 protocol specification [11], a device that wishes to be compatible with the greatest installed base of Cisco TIP devices will advertise TX=1 and RX=0 for Unrestricted Media Constraints bit and adhere to the small number of constraints described in Section 4.2.5.12 of the TIP v8 protocol specification [11].

15 Codec Specs

15.1 Cisco TIP HD Video Encoder Requirements (Restricted Media)

1. MUST generate H264 Main Profile or High Profile (HiP), if negotiated, compatible bit stream
2. MUST support CALVC
3. MAY support CABAC as a negotiated option
4. MUST support 1280x720 or 1920x1080 resolution
5. MUST support fixed frame rate: 30 fps or 29.97 fps
6. MUST generate one macro block-row per slice

   -1280x720: 45 slices per frame, each slice must be 80 macro blocks in length

   -1920x1072: 67 slices per frame, each slice must be 120 macro blocks in length
7. MUST support deblocking_filter_control_present = 1 & disable_deblocking_filter_idc = 1
8. MUST only use inter-prediction blocks of size 16x16
9. MUST only use one reference picture and that reference picture must be the immediate previous frame or a long term reference picture
10. Each NAL MUST be less than 3320 bytes in size

11. POC (Picture Order Count) MUST increment by 2 every frame
12. MUST NOT USE SPS IDs in the range of 10 to 20
13. Long term reference pictures (LTRPs) MAY be used for error concealment. LTRPs requires max_num_ref_frames>=3. LTRPs use H.264 standard picture buffer management. MUST support long_term_frame_idx < 2, MUST use memory_management_control_operation = 6. LTRPs are optional and can be disabled.
14. Gradual decoder refresh pictures (GDRs) MAY be used at the start of a sequence instead of an instantaneous decoder refresh picture (IDR). A recovery point SEI precedes the GDR. The SEI always has the broken_link_flag set to 0. GDRs are optional and can be disabled.
15. When high profile is negotiated, MUST support seq_scaling_matrix_present_flag=0 and pic_scaling_matrix_present_flag=0, MUST NOT use monochrome pictures, MUST NOT use quantization matrices.

15.2 Cisco TIP HD Video Encoder Requirements (Unrestricted Media)

1. MUST support H264 Baseline Profile compatible bit stream with the following constraints
2. MUST support CALVC
3. MUST NOT negotiate or encode using CABAC or 8x8 Transforms
4. MUST support 1280x720 resolution
5. MUST support fixed frame rate: 30 fps or 29.97 fps
6. MUST only use inter-prediction blocks of size 16x16
7. Each NAL MUST be less than 3320 bytes in size
8. MUST NOT USE SPS IDs in the range of 10 to 20
9. Long term reference pictures (LTRPs) MAY be used for error concealment. LTRPs requires max_num_ref_frames>=3. LTRPs use H.264 standard picture buffer management. MUST support long_term_frame_idx < 2, MUST use memory_management_control_operation = 6. LTRPs are optional and can be disabled in the TIP negotiation.
10. Long term reference pictures (LTRPs) MUST NOT be used if also transmitting disposable p-frames, thus disable LTRP in the TIP negotiation if disposable p-frames might be used.
11. Gradual decoder refresh pictures (GDRs) MAY be used at the start of a sequence instead of an instantaneous decoder refresh picture (IDR). A recovery point SEI precedes the GDR. The SEI always has the broken_link_flag set to 0. GDRs are optional and can be disabled in the TIP negotiation.

15.3 Cisco TIP CIF Video Encoder Requirements

1. MUST support H264 Baseline Profile compatible bit stream
2. MUST support 352x288 resolution
3. May support variable frame rate up to 30 fps
4. Uses one reference picture from previous frame and that reference picture must be the immediate previous frame or a long term reference picture
5. Each NAL MUST be less than 3320 bytes in size

15.4 Cisco TIP Presentation Video Encoder Requirements (1 & 5 fps)

1. MUST support H264 Baseline Profile compatible bit stream with the following constraints;
2. MUST support CALVC
3. MUST NOT negotiate or encode using CABAC or 8x8 Transforms

4. MUST support 1024x768 resolution
5. MUST support Fixed frame rate: 1 or 5 fps
6. MUST only use inter-prediction blocks of size 16x16
7. Each NAL MUST be less than 3320 bytes in size

15.5 Cisco TIP Presentation Video Encoder (30fps) Requirements

1. MUST support H264 Baseline Profile compatible bit stream with the following constraints
2. MUST support CALVC
12. MUST NOT negotiate or encode using CABAC or 8x8 Transforms
3.
4. MUST support 1024x768 resolution
5. MUST support fixed frame rate: 30 fps or 29.97 fps
6. MUST only use inter-prediction blocks of size 16x16
7. Each NAL MUST be less than 3320 bytes in size
8. MUST NOT USE SPS IDs in the range of 10 to 20
9. Long term reference pictures (LTRPs) MAY be used for error concealment. LTRPs requires max_num_ref_frames>=3. LTRPs use H.264 standard picture buffer management. MUST support long_term_frame_idx < 2, MUST use memory_management_control_operation = 6. LTRPs are optional and can be disabled.
10. Gradual decoder refresh pictures (GDRs) MAY be used at the start of a sequence instead of an instantaneous decoder refresh picture (IDR). A recovery point SEI precedes the GDR. The SEI always has the broken_link_flag set to 0. GDRs are optional and can be disabled in the TIP negotiation.

15.6 Cisco TIP Audio Encoder Requirements

Codec: G.711 (u-law)
Frame Size: 80 samples
CNG (Comfort Noise Generation) is OFF
Packetization interval: 20ms

Codec: G.722
Frame Size: 160 samples
Sampling Rate: 16KHz
Bit Rate: 64 kbps
CNG is OFF
Packetization interval: 20ms

Codec: AAC-LD
Frame Size: 480 samples
Sampling Rate: 48KHz
Bit Rate: 64 kbps
Mono channel
Transmux = 0
PNS (Perceptual Noise Substitution): enabled
LTP (Long Term Prediction): disabled
Packetization interval: 20ms

15.7 Cisco TIP Audio Activity Metric Calculation

The audio activity metric is an integer value from 0-99 computed by comparing the current audio power level to a measured noise floor. If the current power is less than or equal to the floor then the metric value should be 0. If the current power is between 1dB and 20dB above the floor then the metric value should be between 1 and 89. If the current power is between 20dB and 40dB above the floor then the metric value should be between 90 and 98. If the current power is greater than 40dB above the floor then the metric value should be 99.

The calculation of the noise floor should be dynamic during the course of the call, although historical information may be used as the initial value. The goal of the noise floor is to eliminate artificially high audio activity metric values due to constant background noise.

16 SIP Invite – TIP Indicators in Contact Header

The initial offer SHOULD send the follow TIP indicators in the SIP Contact Header. Note that these indicators can only be sent or may be received in a CUCM 8.5.x or later deployment.

16.1 TIP indicator

Definition: tip-indicator = "x-cisco-tip"

Example:

```
contact: <>;x-cisco-tip
```

16.2 Number of TIP Screens indicator

Definition: Multi-screen = "x-cisco-multiple-screen=" integer

Examples:

```
contact: <>; x-cisco-multiple-screen=1 # one screen TIP endpoint

contact: <>; x-cisco-multiple-screen=3 # three screen TIP endpoint
```

17 Point-to-point, Native SIP Interoperability Considerations

This section provides additional information for TIP implementers to use to interoperate with Cisco TelePresence devices that can establish either TIP or native, SIP -only (TIP-less), point-to-point connections. These point-to-point TIP-less connections will only support a single screen experience, with a maximum of one audio stream, one video stream and one AUX/shared stream (eg, shared presentation content). Cisco immersive TelePresence endpoints as of release 1.8.x will support TIP or native, SIP –only point-to-point connections

17.1 Initial SIP Invite Considerations

The initial SDP offer MAY offer extra audio and/or video codecs, as well as an extra media lines for presentations/AUX, specifically if needing to be ready to interoperate in a native, SIP-only connection. If electing to include extra codecs in the offer, the initial SDP offer MUST list the audio and/or video codecs required by this Cisco TIP Implementation profile document first and include any extra codecs subsequent to those.

17.2 Determining TIP or SIP-only Negotiation

In the TIP protocol specification [11], section titled 'TIP Negotiation,' it also states, "For a short interval (recommendation is 15 sec), the MUXCTRL packet should be periodically retransmitted, until an acknowledgement is received or the interval elapses without a response."

Circumstances in which a TIP device MAY decide to abort sending MUXCTRL packets early, and operate in a native, SIP-only mode are:

1. If the remote device starts sending non-TIP formatted RTCP/SRTCP packets before any TIP MUXCTRL packets are received from it.
2. If the remote device starts sending RTP/SRTP media before TIP negotiation completes for that respective media.

When codecs negotiated in the initial SDP offer from the remote device are not consistent with the MUST requirements in this Cisco TIP Endpoint Profile document, a TIP device MAY decide not to start sending TIP MUXCTRL packets at all and operate in a native, SIP-only mode instead.

18 Summary of Changes to this document

Initial version of document; no changes.

19 References

[1] IETF RFC 3261 "SIP: Session Initiation Protocol"
[2] IETF RFC 3264 "An Offer/Answer Model with the Session Description Protocol (SDP)"
[3] IETF RFC 2327 "SDP: Session Description Protocol"
[4] IETF RFC 3550 "RTP: A Transport Protocol for Real-Time Applications"
[5] IETF RFC 3551 "RTP Profile for Audio and Video Conferences with Minimal Control"
[6] IETF RFC 4961 "Symmetric RTP/RTP Control Protocol (RTCP)"
[7] IETF RFC 3984 "RTP Payload Format for H.264 Video"
[8] IETF RFC 4585 "Extended RTP Profile for Real-time Transport Control Protocol (RTCP)-Based Feedback (RTP/AVPF)"
[9] IETF RFC 5761 "Multiplexing RTP Data and Control Packets on a Single Port"
[10] "Cisco Unified Communications Manager SIP Line Messaging Guide (Standard)" can be found at http://developer.cisco.com/web/sip/docs
[11] Telepresence Interoperability Protocol (TIP), version 6, 7 or 8 and all the Cisco TIP Endpoint Implementation Profile revisions can downloaded at http://www.imtc.org/tip/
[12] IETF RFC 2833 "RTP Payload for DTMF Digits, Telephony Tones and Telephony Signals"
[13] IETF RFC 4730 "SIP Event Package for Key Press Stimulus (KPML)"
[14] IETF RFC 5168 "XML Schema for Media Control"
[15] IETF RFC 3640 "RTP Payload Format for Transport of MPEG-4 Elementary Streams"
[16] IETF RFC 3890 "A Transport Independent Bandwidth Modifier for the Session Description Protocol (SDP)"
[17] "Cisco Unified Communications Manager SIP Trunk Messaging Guide" can be found at http://developer.cisco.com/web/sip/docs
[18] To use the CUCM Extended SIP interface, one must first be members of the Cisco Developer Network (CDN). Learn more by visiting http://developer.cisco.com/web/cdc/home. If you would like to contact Cisco directly about becoming an SIP extension licensee, please send all inquiries to: ipcbu-busdev@external.cisco.com. Please note that no SIP extension license is required to interoperate with Cisco TelePresence systems implementing TIP.
[19] IETF RFC 5764 "DTLS Extension to Establish Keys for SRTP" and IETF RFC 5763 "Framework for Establishing a SRTP Security Context Using DTLS"
[20] [EKT] Internet-Draft "Encrypted Key Transport for Secure RTP" version 03 http://tools.ietf.org/html/draft-ietf-avt-srtp-ekt-03
[21] Cisco TelePresence Unified Communications Compatibility Information http://www.cisco.com/en/US/products/ps8332/products_device_support_tables_list.html
[22] "Security for Generic Secure Video SIP Endpoints" UCM document can be found at http://www.cisco.com/en/US/docs/voice_ip_comm/cucm/security/8_5_1/secsipvideo.pdf

Glossary

AAC-LD	MPEG-4 Advanced Audio Coding - Low Delay Audio
AES	Advanced Encryption Standard
AVP / SAVP	Audio Video Profile / Secure Audio Video Profile
CABAC	Context-adaptive binary arithmetic coding
CALVC	Context-adaptive variable length coding
CIF	A video format that supports both NTSC and PAL signals
CSRC	Contributing Source in RTP
DTLS	Datagram Transport Layer Security
EKT	Encrypted Key Transport
GDR	Gradual Decoder Refresh
IDR	Instantaneous Decoder Refresh
LTRP	Long-Term Reference Picture
MCU	Multipoint Control Unit
MUX	Multiplexer/Multiplexing
NAL	Network Abstraction Layer
NTP	Network Time Protocol
PPS	Packets per second or Picture Parameter Set
RSA	Rivest, Shamir and Adleman (an encryption protocol)
RTCP	Real-Time Control Protocol
RTP	Real-Time Protocol
SDP	Session Description Protocol
SEI	Supplemental Enhancement Information for H.264 Frames
SIP	Session Initiation Protocol
SPIMAP	Serial Peripheral Interface Map
SPS	Sequence Parameter Set
SRTCP	Secure Real-Time Control Protocol
SRTP	Secure Real-Time Protocol
SSRC	Synchronization Source in RTP
STUN	Simple Traversal of UDP through Network Address Translators (NATs)
TIAS	Transport Independent Application Specific descriptions in SIP
TLS	Transport Layer Security
UCM	(Cisco) Unified Communications Manager