Telepresence Interoperability Protocol ("TIP") License

Published April 2010

This is a legal agreement ("Agreement") between the user of the Specification or the Profiles (either an individual or an entity) ("You"), and Cisco Systems, Inc. ("Cisco") (collectively, the "Parties"). If You want a license from Cisco to implement the TIP Specification or the Profiles (as defined below), You must agree to these terms. This is an offer to be accepted only on the terms set forth in this Agreement. You accept these terms by accessing, downloading, using, or reviewing the TIP Specification. Once accepted, Cisco 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 Agreement. If You attempt to alter or modify any of these terms, then the offer is no longer valid and is revoked.

The TIP Specification provides a protocol for interoperability between videoconferencing products, including streaming of audio, video, and data to and from videoconferencing products. We offer 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 Implementation Profiles ("Profiles") offer additional information regarding Cisco's implementation of the TIP Specification, and are intended to facilitate interoperability with Cisco TelePresence products. The TIP Specifications and the Profiles are referred to collectively as the "TIP Documents".

No part of the TIP Documents may be printed out, reproduced, distributed, resold, or transmitted for any other purpose, including, without limitation, commercial purposes, such as selling copies of this documentation or providing paid-for support services.

Definitions

"Compliant Implementation" means only the portions of Your product that (i) fully implements all required portions of the TIP Specification or any aspect of the Profiles that is required for interoperability with a Cisco TelePresence productand (ii) are fully compliant with the TIP Specifications.

"Essential Claim" means a claim of a patent, whether or not issued as of the date of this Agreement, that is necessarily infringed by implementing the required portions of the TIP Specification or any aspect of the Profiles that is required for interoperability with a Cisco TelePresence product. A claim is necessarily infringed only when there is no technically or commercially reasonable way to avoid infringement of that claim when following the requirements of the Specification or Profiles to make a Compliant Implementation. Notwithstanding the foregoing, Essential Claim does not include any claim that, if licensed, would require a payment of royalties by the licensor to a person or entity that is not affiliated with the licensor.

“TIP Profiles” means the profiles found at http://www.cisco.com/go/tip . Only the TIP Profiles published by Cisco will be considered the TIP Profiles for purposes of this Agreement. Any profiles not published by Cisco or one that incorporates any of the TIP Profiles in part, in whole or by reference shall not be considered a TIP Profile for purposes of this Agreement.

“TIP Specification” means the specification found at http://www.cisco.com/go/tip. Only the TIP Specification published by Cisco will be considered the TIP Specification for purposes of this Agreement. Any specification not published by Cisco 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 Agreement.

Patent License

Upon your acceptance of these terms, Cisco grants You a non-exclusive, royalty-free, nontransferable, non-sublicensable, personal, worldwide license under all of Cisco’s Essential Claims, to make, have made, use, sell, offer to sell, import and distribute Compliant Implementations. For the avoidance of doubt, nothing in the previous sentence grants You a license under (a) any claim of any patent licensable by Cisco that is not an Essential Claim; or (b) any Essential Claim that is infringed by any manufacture, use, sale, offer for sale, importation, or other disposition other than the manufacture, use, sale, offer for sale, importation, or other disposition of a Compliant Implementation.

Prohibited Uses

The rights and licenses granted by Cisco in the TIP Specification, including those granted in the Patent License, are conditioned upon Your agreement to use the TIP Specification for only the purpose of interoperating between videoconferencing products, and for no other purpose.
No right or license to any Cisco intellectual property is granted under this agreement for any other use.

Defensive Suspension

If You assert, threaten to assert, or participate in the assertion of a lawsuit, proceeding, claim or similar action directed against any other person or entity, including Cisco, claiming that the manufacture, use, sale, offer for sale, importation or other disposition of a Compliant Implementation of any of the TIP Documents infringes any of Your Essential Claims, then Cisco may terminate all license grants and any other rights provided under this Agreement to You. For the avoidance of doubt, nothing in the previous sentence shall cause the termination of the license or other rights provided under this Agreement to You upon the assertion by You of (a) any claim of any patent licensable by You that is not an Essential Claim; or (b) any Essential Claim that You allege is infringed by any manufacture, use, sale, offer for sale, importation or other disposition other than the manufacture, use, sale, offer for sale, importation, or other disposition of a Compliant Implementation.

Reservation of Rights

All rights not expressly granted herein are reserved.

Updates

Cisco reserves the right to issue updates to any of the TIP Documents. Updates will be provided on Cisco’s website, at http://www.cisco.com/go/tip.

Feedback and Improvements

Cisco welcomes, and will consider, feedback to and suggested improvements to, the TIP Documents. Feedback and suggested improvements may be sent by e-mail to ask-tip@cisco.com. If You would prefer to contact Cisco before sending your feedback or suggested improvements, please send an e-mail to that effect to ask-tip@cisco.com.

Please note that by providing such feedback or suggested improvements to Cisco, You acknowledge and agree that Cisco may use such feedback and/or suggested improvements in connection with the TIP Specification without obligation of any kind to You. Your submission of feedback constitutes Your agreement that Cisco may incorporate any feedback into the TIP Specification (including any Update), and that You are licensing Cisco, its affiliates, and all other implementers of the TIP Documents on the same terms on which Cisco is making the TIP Documents available to You and other actual and prospective implementers.

Third-Party Information

The TIP Specification may contain links to third-party websites and standards and specifications (for example, H.264 and the IETF RTP specification) that are not under the control of Cisco, and Cisco is not responsible for the content on any linked site or referenced standard or specification. If You access a third-party website mentioned in the Specification, then You do so at Your own risk. Cisco provides these links and references only as a convenience, and the inclusion of the link does not imply that Cisco endorses or accepts any responsibility for the content on those third-party sites or of the referenced standards and specifications. No right, license or interest is granted in any third party technology referenced in this guide. You acknowledge that the owners of intellectual property rights asserted to be necessary to implement any referenced standard may or may not make those intellectual property rights available to implementers.

No Warranty

THE TIP SPECIFICATION AND THE LICENSES GRANTED UNDER THIS AGREEMENT ARE PROVIDED “AS IS,” AND CISCO MAKES NO REPRESENTATION, WARRANTY, CONDITION OR OTHER TERM, EXPRESS OR IMPLIED (BY COMMON LAW, STATUTE OR OTHERWISE), AND DISCLAIMS ANY AND ALL REPRESENTATIONS, WARRANTIES, CONDITIONS OR OTHER TERMS WHICH MAY BE IMPLIED (BY COMMON LAW, STATUTE OR OTHERWISE) WITH RESPECT TO THE TIP DOCUMENTS AND/OR THE LICENSES TO EACH OF THE TIP DOCUMENTS GRANTED UNDER THIS AGREEMENT, IN WHOLE OR IN PART, ALL OF WHICH ARE EXCLUDED, INCLUDING, WITHOUT LIMITATION, ANY IMPLIED TERM OF FITNESS FOR A PARTICULAR PURPOSE, MERCHANTABILITY, TITLE OR NONINFRINGEMENT. CISCO AND ITS AFFILIATES SHALL NOT 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 ANY USE OR IMPLEMENTATION OF THE TIP SPECIFICATION, IN WHOLE OR IN PART, INCLUDING, BUT NOT LIMITED TO, ANY USE, MANUFACTURE, OFFER TO SELL, SALE, IMPORTATION, OR DISTRIBUTION OF ANY COMPLIANT IMPLEMENTATION.

Limitation Of Liability

CISCO AND ITS AFFILIATES SHALL NOT BE LIABLE UNDER OR IN CONNECTION WITH THIS AGREEMENT (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 AGREEMENT WHATSOEVER,

IN EACH CASE HOWEVER CAUSED OR ARISING, PROVIDED THAT NOTHING IN THIS AGREEMENT SHALL EXCLUDE OR LIMIT CISCO'S AND ITS AFFILIATES' LIABILITY:

(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 CISCO WOULD NOT HAVE ENTERED INTO THIS AGREEMENT ON THE TERMS HEREOF WITHOUT THE LIMITATIONS OF LIABILITY SET FORTH HEREIN.

Governing Law
This Agreement shall be construed in accordance with, and all disputes hereunder shall be governed by, the laws of England and Wales. The Courts of England and Wales shall have exclusive jurisdiction over any disputes relating to the Agreement.

No Requirement to Implement
Nothing in this Agreement will be construed as requiring You to use or implement any of the TIP Documents, or limit You from developing any products that compete with Cisco products provided that You do not violate any of Cisco's intellectual property rights with such development.

Telepresence Interoperability Protocol (TIP)

Version 6.0

This document can only be distributed under the Cisco TIP License Agreement.

Modification History

Revision	Date	Originator	Comments
1.0	01/19/2010	Cisco Systems, Inc.	Initial document
1.1	04/05/2010	Cisco Systems, Inc.	Clarified the reference to implementation profile documents

Table of Contents

1 Introduction

Generally, TIP systems are high-end, high-definition video conferencing devices built to operate seamlessly with Cisco's TelePresence Systems. A TIP compliant device is capable of handling multiple audio and video streams. The devices negotiate the number of media streams they will exchange, and media sources are switched when necessary to always present the viewer with the active conference participants.

The devices can be endpoints communicating in a point-to-point call or they may communicate with a TIP-compliant multipoint control unit (MCU) to enable more than two endpoints to communicate simultaneously.

The TIP protocol provides a variety of ways to indicate or negotiate endpoint or MCU configuration, preferences or restrictions. A set of accompanying documents, called "Cisco TIP Implementation Profile" guides, define what protocol and media options are used to achieve tested interoperability with Cisco and these may be updated from time to time as product capabilities evolve.

2 Background Information

The TIP protocol is designed around the IETF standards for Voice over IP and Video Conferencing. A TIP-compliant device would use SIP [1] for connection management and signaling, and RTP/RTCP [4] for media transmission. In the call scenarios discussed in this document, the media transmission uses only IP unicast. No IP multicast capabilities are used.

TIP devices will offer both audio and video streams during the SDP offer/answer exchange [2] that is part of the SIP INVITE or re-INVITE process. TIP devices do not express their multiple stream capability in SDP [3].

All supported media formats are carried in their respective IETF RTP payloads, specifically:

- Audio
  - AAC-LD
    - Bitrate: 64 kbps/channel
    - RTP Payload: IETF RFC 3640, AAC-hbr mode
    - Default Dynamic Payload Number: 96
  - G.711 (u-law)
    - RTP Payload: IETF RFC 3351
    - Static Payload Number: 0
  - G.722
    - RTP Payload: IETF RFC 3351
    - Static Payload Number: 9

DTMF
RTP Payload: IETF RFC 2833
Default Dynamic Payload Number: 101

Video
H.264 Baseline Profile
Image sizes: 1080p, 720p, 1024x768, 352x288
Bitrates: 4 Mbps to 300 kbps
RTP Payload: IETF RFC 3984, packetization mode 1 and mode 0
Default Dynamic Payload Number: 112

The following discussion assumes detailed knowledge of the RTP protocol [4] and [5]. As a convenience, the RTP and RTCP packet diagrams are replicated from [4] below:

RTP Header

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|V=2|P|X|  CC   |M|     PT      |       sequence number         |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                           timestamp                           |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|           synchronization source (SSRC) identifier            |
+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+
|            contributing source (CSRC) identifiers             |
|                             ....                              |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

RTCP Header

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|V=2|P| subtype |       PT      |             length            |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                              SSRC                             |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

The following types of RTCP packets may be used by TIP devices:

Description	Abbreviation	Type
Sender Report	SR	200
Receiver Report	RR	201
Session Description	SDES	202
Goodbye	BYE	203
Application	APP	204
Feedback	FB	205
Extended Report	XR	207

3 Media Channel Establishment

This document begins its discussion at the point where call signaling has been established via SIP [1] and the SDP offer/answer exchange [2] has determined the potential network addresses to be used for the media channel between two entities in the media path. This section details how the media channel and the rules for its operation are established.

3.1 NAT/Firewall Traversal

The first challenge to the establishment of the media channel is the traversal of any NAT/Firewalls that may exist between the two media processing entities (endpoints or MCUs).

To help with the NAT/Firewall traversal TIP mandates the use of symmetric RTP/RTCP ports [6]. A TIP-complian device MUST transmit all packets in a symmetric fashion, i.e., that the IP address and UDP port of a transmitted packet exactly match the IP address and UDP port for reception of the corresponding media stream. However, when receiving packets from the remote peer, implementations SHOULD accept packets with a UDP source port that differs from the destination UDP port that is being used when transmitting.

3.2 TIP Negotiation

This section describes how an implementation detects that its remote peer is capable of understanding TIP. If the remote peer does not implement TIP, an implementation SHOULD operate in a strictly standards compliant mode where only a single audio or video stream is sent in the media channel, and no TIP specific control extensions are used.

TIP uses the RTCP extension mechanism, the APP packet, which allows for private uses to signal all of its control information. TIP uses the RTCP private application name “xcts” ,eXtended Control for Telepresence Systems, for all of its control packets.

TIP defines an RTCP APP MUXCTRL packet which informs the remote peer of media multiplexing capabilities. This includes the number of media streams that a device is willing to transmit and willing to receive for a given media channel of a particular type (audio or video). It also allows the receiver to specify which media positions it is willing to transmit and receive. The reception of an RTCP APP MUXCTRL packet or an RTCP APP MUXCTRL ACK packet specifying the TIP private application name, “xcts”, is used as acknowledgement that TIP is understood by the remote peer.

TIP requires that the local system’s MUXCTRL be acknowledged by the remote peer and that the MUXCTRL of the remote peer be received before communication under the rules of TIP can begin. The RTCP APP packets and processing rules are detailed in section 4.2

The TIP negotiation begins by sending the RTCP APP MUXCTRL on the RTCP UDP channel(s) indicated by the media channel establishment.

For a short interval (recommendation is 15 seconds), the MUXCTRL packet should be periodically retransmitted, until an acknowledgement is received or the interval elapses without a response.

The reception of either a MUXCTRL or MUXCTRL ACK packet on the RTCP UDP channel SHOULD be interpreted as support of TIP.

An implementation SHOULD acknowledge all MUXCTRL packets it receives. The lack of any response MUST be interpreted as a lack of support for TIP, and hence the standard RTCP traffic should be sent on the RTCP UDP channel as per the standard mode of operation for [4], unless the peer has signaled otherwise in SDP [9].

4 TIP Multiplex

4.1 Positional Multiplex

The key function of TIP is to enable multiple media streams, belonging to the same media type, to be transported on the same UDP channel within the boundaries of standards compliance. To accomplish this there needs to be a multiplex point for the multiple streams based on a label, here after referred to as the “position”. This section describes the mechanisms for this multiplex.

All SSRC fields MUST be random 32 bit values, that are unique across the RTP session as per the RTP and SRTP standards [4]. Each unique entity that is a RTP or RTCP transmitter should have a unique SSRC value.

The TIP positional multiplex value MUST be added to all RTP packets as the first CSRC value. (Note that multiple CSRC values are still allowed, but the first one must contain the TIP positional values.) The RTP CC field should be set appropriately to indicate the addition of the CSRC value. This document will henceforth refer to this value as the MUX-CSRC.

The following packet diagram illustrates these rules.

TIP RTP Packet

```
      0                   1                   2                   3
     0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
    +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+<+
    |V=2|P|X| CC=1  |M|     PT      |       sequence number         | |
    +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+ |
    |                           timestamp                           | |
    +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+ |
    |           synchronization source (SSRC) identifier            | |
    +=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+ |
    |    contributing “positional” source (MUX-CSRC) identifier     | |
  +>+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+ |
  | |                           payload  ...                        | |
  | |                               +-------------------------------+ |
  | |                               | RTP padding   | RTP pad count | |
  +>+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+<+
```

TIP multiplexes multiple RTP media streams of the same type (audio or video) on one UDP channel.

For the video stream multiplex, typicall all the streams use the same encoding format, H.264. Hence the RTP payload number is the same for all streams. Each video stream identifies its particular media configuration options via use of inband H.264 parameter sets (SPS/PPS).

For the audio stream multiplex, the situation is more complex. Depending on device capabilities and the call topology multiple encoding formats such as AAC-LD, and/or G.711 and G722 audio streams may be used. Also any of the audio streams may include DTMF sent via RFC 2833 encoding. Hence multiple RTP payload numbers may be seen within the audio multiplex.

The RTP media streams sent on the same UDP channel are multiplexed by using different RTP MUX-CSRC values for each stream. Note that each RTP stream has its own independent RTP sequence number and RTP timestamp space, hence a receiver should be prepared for a change in the value of the SSRC to be accompanied by a change of sequence number and timestamp. Note that multiplexing different ,yet realted, RTP streams of the same media is allowed in RTP as long as certain requirements are met [4]. Such multiplexing is actually the norm in multicast sessions.

Additionally, the bits of the MUX-CSRC field are subdivided and assigned specific semantics:

```
0                   1                   2                   3
0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|           Sampling Clock ID           | OutPos|XmitPos|RcvPos |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Bits (Big-Endian)	Mask (Big-Endian)	Description
31-12	0xFFFFF000	Sampling Clock ID
11-8	0x00000F00	Output Position
7-4	0x000000F0	Transmitter Position
3-0	0x0000000F	Receiver Position

The Sampling Clock ID is a random number chosen by the transmitting endpoint to designate the media sampling clock used for this stream. If media streams share the same clock source then they should use the same value of the Sampling Clock ID. This allows TIP receivers to detect when two media streams should be strictly synchronized. This is consistent with [4]. In situations where there is no literal sampling clock, e.g. for control packets, a unique random number for the transmitter should be selected and used for this subfield.

The stream positions denote a spatial relationship or functional use of the stream. Audio and video streams with the same positions are intended to accompany each other -- i.e., there is a physical correspondence between the left audio position and the left video position.

The transmitter position denotes a both an RTP transmitter and an input channel as these are tightly bound.

The receiver position denotes an RTP receiver and a default output channel.

The output position is an option that can be used to direct a receiver to output media to a non-default channel. If the option is not used, the value of the output position SHOULD be zero. The Audio Dynamic Output option described later in this document uses this output position.

The definition of the video stream position values are:

Value	Label	Description

Value	Label	Description
0	Control	Multiplex control
1	Center	Physical center
2	Left	Physical left
3	Right	Physical right
4	Aux	Auxiliary e.g. presentation, doc camera
5	<Not Used>	Reserved
6	<Not Used>	Reserved
7-8	<Not Used>	NA
9	Legacy Center	Legacy interoperability version of center
10	Legacy Left	Legacy interoperability version of left
11	Legacy Right	Legacy interoperability version of right
12	<Not Used>	NA
13-15	<Not Used>	NA

Table 1: Video stream position values

The definition of the audio stream position values are:

Value	Label	Description
0	Control	Multiplex control
1	Center	Physical center
2	Left	Physical left
3	Right	Physical right
4	Aux	Auxiliary audio
5	<Not Used>	NA

Value	Label	Description
6	<Not Used>	NA
7-8	<Not Used>	NA
9	<Not Used>	NA
10	<Not Used>	NA
11	<Not Used>	NA
12	Legacy Mix	Legacy interoperability mix of all audio positions
13-15	<Not Used>	NA

Table 2: Audio stream position values

The transmitting source is allowed to switch which media streams (i.e. MUX-CSRC's) are sent at any time based on a number of local mechanisms. However the total number of streams sent must not exceed the value negotiated with the receiver. For example, a transmitter may be capable and allowed to send any one of it's three camera streams to a receiver, but it may only send one at any given time. (Note that for video streams, the transmitting source should make provisions to ensure that a newly switched video stream can be decoded by the receiver, e.g. generation of an IDR frame for H.264.)

Note with the RTCP APP extensions, there are situations of packets that are generated and/or consumed by a control entity that does not have a definitive position. In these situations the reserved value of "Control" (0) SHOULD be used for the position in the MUX-CSRC.

4.2 TIP RTCP APP Extensions

A number of control extensions are used in TIP. This section describes the TIP specific extensions carried in RTCP APP packets.

Note that each of the TIP specific RTCP APP messages MUST be sent within compound RTCP packets as per [4]. The way to accomplish this is to prepend an "empty" receiver report (RR) and session description (SDES with CNAME) before the APP packet.

RTCP Packet with Empty RR, SDES CNAME, and TIP APP extension

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

```
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|V=2|P|   RC=0    |     PT=RR=201   |          length=1           |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                               SSRC                              |
+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+
|V=2|P|   SC=1    |   PT=SDES=202   |          length             |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                               SSRC                              |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|    CNAME=1      |     length      |   user and domain name    ...
+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+
|V=2|P| subtype |    PT=APP=204   |          length             |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                               SSRC                              |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                            name=xcts                            |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
~                               ...                               ~
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Note that by default the TIP APP extensions are hop-by-hop, and do not need to be forwarded by MCU's. However some of the extensions do need to be forwarded. These extensions either specify a rule by which forwarding should be performed, or they include a "target" field that specifies the MUX-CSRC value of the intended recipient (from the perspective of the RTCP packet transmitter).

Note that currently TIP does not attempt to abide by the RTCP bandwidth limitations discussed in [4] with respect to the RTCP APP packets. The packets are small in size, limited in number, and used only over unicast connections, and hence will minimally perturb the RTCP bandwidth allocation.

4.2.1 Multiplex Control

This packet is used to negotiate and establish the parameters of the TIP multiplex. It MUST be sent after the media channel is established but before any media is transmitted. It MAY be sent later in the call to update some of the multiplex parameters (see below for details).

RTCP APP MUXCTRL packet format

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|V=2|P|subtype=1|    PT=APP=204   |          length             |
```

```
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                              SSRC                             |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                           name=xcts                           |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| MV=6  | profile |  options    |numXmitStreams | numRcvStreams |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|             NTP timestamp, most significant word              |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|             NTP timestamp, least significant word             |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|  Conference ID, most significant word                         |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|  Conference ID, least significant word                        |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|         xmitPositions         |         rcvPositions          |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

RTCP APP Packet field definitions, should follow [4] and:

`subtype`: 5 bits

MUST be set to '1' to designate the MUXCTRL packet type.

`SSRC`: 32 bits

MUX SSRC field which is a arndom number

`name`: 4 bytes

MUST be set to the value 'xcts'.

MUXCTRL Packet field definitions:

`mux version (MV)`: 4 bits

The version of the TIP multiplex. Implementations of this document MUST specify the value '6'.

`profile`: 4 bits

TIP employs RTP profiles beyond the standard AVP profile described in [5], specifically support for a special flavor of the Feedback Profile. This is used to carry packet ACK/NACK information in a manner aligned with [8]. The MUXCTRL packet can be used to negotiate use of extended RTP profiles if it has not already been accomplished by call signaling. The allowed values assign numeric values to the IANA defined profiles. The values are:

Value	Name	Specification

0	RTP/AVP	RFC 3351
1	Unused	NA
2	RTP/AVPF	TIP flavor of RFC 4585
3-15	Unused	NA

`options`: 8 bits

The following option bits are currently defined:

Bits	Value	Description
0	0x01	Sender is a multipoint focus (aka an MCU)
1-7	NA	Unused, should be zero.

`num xmit streams`: 8 bits

The number of simultaneous media streams that can be transmitted within this multiplex by the system sending the MUXCTRL information.

`num rcv streams`: 8 bits

The number of simultaneous media streams that can be received within this multiplex by the system sending the MUXCTRL information. A transmitter MUST never send more that the number of streams the receiver is willing to accept.

`ntp timestamp`: 64 bits

The NTP time associated with the creation of the MUXCTRL information. Note that retransmissions with the same NTP timestamp is permitted.

`conference identifier`: 64 bits

A conference identifier used to associate with a particular multipoint call. The value of zero is reserved to indicate that the field is unused and invalid, e.g. in a point to point call.

`xmit positions:` 16 bits

A bitmask of the available transmit positions. Transmit position i is available if the i'th bit is set to 1, otherwise it is not allowed. See Table 1 for video position values and Table 2 for audio position values.

`rcv positions:` 16 bits

A bitmask of the available receive positions. Receive position i is available if the i'th bit is set to 1, otherwise it is not allowed. See Table 1 for video position values and Table 2 for audio position values.

It is RECOMMENDED to send MUXCTRL packets every 250 ms, for the first 15 seconds of the call (for a maximum of 60 transmissions). The NTP timestamp should be the same value for all retransmissions of the MUXCTRL packet. Once a MUXCTRL ACK packet is received from the peer with a received NTP time that matches the transmitted MUXCTRL packet, no more transmissions should be made.

If later during the call, the local number of streams that can be accommodated changes, further MUXCTRL packets can be sent. The same retransmission strategy as above should be used in this situation. If the updated MUXCTRL packet is not acknowledged, the call can continue with the previously negotiated values. The value of the profile is not allowed to change during these renegotiations. The profile can only be negotiated once during the initial exchange.

These requests MUST be explicitly acknowledged by the receiver using the RTCP APP ACK scheme described in section 4.2.6.

4.2.2 Network Path Measurement

To enable accurate measurements of the network path between peers, TIP specifies the RTCP APP packet extension “ECHO”. This packet is similar in function to an ICMP ECHO, but has some significant advantages. By performing the echo function in-band, the same network path and most importantly the same QoS is applied to the measurement packets as the media packets. Also ICMP ECHO packets are frequently filtered by intermediate systems such as firewalls. The in-band RTCP messages are less likely to suffer this fate.

TIP devices can use these RTCP APP ECHO packets to measure network path latency. Indirectly this also serves as aliveness test of the network and the remote peer. The current best practice is to transmit an ECHO packet once per second during a call on each UDP channel, and measure averages over 10 second periods.

RTCP APP ECHO packet format

```
0                   1                   2                   3
0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|V=2|P|subtype=4|   PT=APP=204  |             length            |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                              SSRC                             |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                           name=xcts                           |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|         Transmit NTP timestamp, most significant word         |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|         Transmit NTP timestamp, least significant word        |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|         Receive NTP timestamp, most significant word          |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|         Receive NTP timestamp, least significant word         |
```

```
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

RTCP APP Packet field definitions should follow [4] and:

`subtype`: 5 bits

MUST be set to ‘4’ to designate the ECHO packet type.

`SSRC`: 32 bits

MUX SSRC field which is a random number

`name`: 4 bytes

MUST be set to the value ‘xcts’.

ECHO Packet field definitions:

`transmit ntp timestamp`: 64 bits

The NTP time of the transmission of the ECHO request. An ECHO response should contain the same transmission time as the request which generated the response.

`receive ntp timestamp`: 64 bits

The NTP time of the reception of the ECHO request. If this field is zero, then the packet is an ECHO request. If this field is non-zero, then the packet is an ECHO response.

Note that consistent with the IETF RTP model, there is no guarantee that the transmitter and receiver share a common NTP clock. Hence the receive timestamp may not fall between the transmit timestamp and the NTP time of the ECHO requestor’s reception of the ECHO response. The ECHO requestor can measure round trip time, and estimate one-way latency if symmetric network paths are assumed. The requestor can also detect if a common NTP clock is shared with the ECHO responder based on the expected value of the receive timestamp and the value actually returned.

These requests MUST NOT be explicitly acknowledged by the receiver using the RTCP APP ACK scheme described in section 4.2.6. These requests use their own acknowledgement mechanism.

4.2.3 Media Flow Control

This flow control applies to the media codec and not the RTP transmitter or receiver. The RTP entity should remain active awaiting either a request to continue media processing or call termination via the SIP control plane.

There is one packet format defined for FLOWCTRL, but two subtypes are used to distinguish between messages destined for transmitters and those destined for receivers.

RTCP APP FLOWCTRL packet format

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|V=2|P| subtype |   PT=APP=204  |             length            |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                              SSRC                             |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                           name=xcts                           |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|         Transmit NTP timestamp, most significant word         |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|         Transmit NTP timestamp, least significant word        |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                       Flow Control State                      |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                              Target                           |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

RTCP APP Packet field definitions should follow [4] and:

`subtype`: 5 bits

MUST be set to '5' to designate the transmit flow control packet type or '6' to designate the receive flow control packet type.

`name`: 4 bytes

MUST be set to the value 'xcts'.

FLOWCTRL Packet field definitions:

`transmit ntp timestamp`: 64 bits

The NTP time of the transmission of the FLOWCTRL request. This can be used by the receiver to detect out of order or duplicate requests.

`flow control state`: 4 bytes

The desired state of flow control. A value of 0 indicates that media should be traversing the network, a value of 1 indicates that media should be stopped. All other values are reserved for future use.

`target`: 4 bytes

The MUX-CSRC of the target that is being requested to act on the flow control state. Typically this value would be provided from the RTP media packets being received from the transmitter.

These requests MUST be explicitly acknowledged by the receiver using the RTCP APP ACK scheme described in section 4.2.6.

4.2.4Video Refresh Request

In scenarios that employ switching of video streams there is a need to have a mechanism to request that video encoders generate a new I Frame (IDR Frame in H.264). This RTCP APP packet is the mechanism for that request.

Please note this mechanism is not intended as the mechanism for repair of packet loss. The feedback mechanism detailed below should be used for that purpose.

RTCP APP REFRESH packet format

```
0                   1                   2                   3
0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|V=2|P|subtype=8|   PT=APP=204  |          length               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                              SSRC                             |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                           name=xcts                           |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|         Transmit NTP timestamp, most significant word         |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|         Transmit NTP timestamp, least significant word        |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                              Target                           |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                               Flags                           |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

RTCP APP Packet field definitions should follow [4] and:

`subtype`: 5 bits

MUST be set to ‘8’ to designate the REFRESH packet type.

`name`: 4 bytes

MUST be set to the value ‘xcts’.

REFRESH Packet field definitions:

`transmit ntp timestamp`: 64 bits

The NTP time of the transmission of the REFRESH request. This can be used by the receiver to detect out of order or duplicate requests.

`target`: 4 bytes

The MUX-CSRC of the target encoder that is being requested to refresh. Typically this value would be provided from the RTP media packets being received from the transmitter.

`flags`: 4 bytes

Flags field to control the type of refresh being requested.

Bits	Value	Description
0	0x00000000	IDR
0	0x00000001	GDR
1-31	NA	Reserved

These requests MUST be explicitly acknowledged by the receiver using the RTCP APP ACK scheme described in section 5.2.10.

4.2.5 Media Options

TIP provides several media encoding options, and extensions of the standard RTP media payloads to allow TIP devices to fully utilize advanced capabilities. By default, these options are disabled. They can be enabled via use of the RTCP APP MEDIAOPTS packet. The details of each media option are described in subsections below.

RTCP APP MEDIAOPTS V2 packet format

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|V=2|P|subtype=7|   PT=APP=204  |             length            |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                              SSRC                             |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                           name=xcts                           |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|         Transmit NTP timestamp, most significant word         |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|         Transmit NTP timestamp, least significant word        |
```

```
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|           Version             |           Positions           |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                         Transmit options                      |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                          Receive options                      |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|   Option tag  |           Option value                        |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
~                              ...                              ~
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|   Option tag  |           Option value                        |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

RTCP APP Packet field definitions should follow [4] and:

`subtype`: 5 bits

MUST be set to '7' to designate the MEDIAOPTS packet type.

`name`: 4 bytes

MUST be set to the value 'xcts'.

MEDIAOPTS Packet field definitions:

`transmit ntp timestamp`: 64 bits

The NTP time of the transmission of the request. This can be used by the receiver to detect out of order or duplicate requests.

version: 2 bytes

The version of the MEDIAOPTS packet. Implementations compliant with this specification MUST set the value to '2'.

positions: 2 bytes

This field is a bitmask of positions within the multiplex for which the media options apply. This can be used to set a media option for just a subset of the active media streams. For this version of the specifications this field MUST be set to 0xFFFF.

`transmit options`: 4 bytes

The media options that the sender of the MEDIAOPTS packet is willing to transmit. See table below for media option value assignments.

`receive options`: 4 bytes

The media options that the sender of the MEDIAOPTS packet is willing to receive. See table below for media option value assignments.

Zero or more option tag/value pairs MAY exist in the MEDIAOPTS packet. The number of option tag/value pairs can be determined by the length of the RTCP length field of the MEDIAOPTS packet.

Each option tag/value is 32 bits, and is subdivided as follows:

`Option tag`: 1 byte

A unique identifier for the option. (See below for assignments).

`Option value`: 3 bytes

A 24 bit value whose interpretation is determined by the option tag. (See below for definitions).

These MEDIAOPTS requests MUST be explicitly acknowledged by the receiver using the RTCP APP ACK scheme described in section 4.2.6.

The media options field bits for an audio multiplex are assigned the following values:

Bits	Value	Description
0	0x00000001	Audio Activity Metric
1	0x00000002	Audio Dynamic Output Channels
2	0x00000004	Capable of G.722 for Legacy Audio
3	0x00000008	Using G.722 for Legacy Audio
4-31	N/A	RESERVED set to 0

The media options field bits for a video multiplex are assigned the following values:

Bits	Value	Description
0	0x00000001	Refresh Flag
1	0x00000002	Inband Parameter Sets (SPS/PPS)
2	0x00000004	Arithmetic Coding (CABAC)
3	0x00000008	Long-Term Reference Pictures (LTRP)
4	0x00000010	Reserved1 (set to 0)
5	0x00000020	Aux Video FPS Bit 0
6	0x00000040	Reserved2 (set to 0)
7	0x00000080	Reserved3 (set to 0)
8	0x00000100	Gradual Decoder Refresh (GDR)
9	0x00000200	Aux Video FPS Bit 1

10-31	N/A	RESERVED set to 0

An option is enabled if it has been offered by the local system and the complementary option has been offered by the remote peer. For example, if the local system has offered to transmit an option, and the remote peer has offered to receive the option, the option is enabled. Otherwise it is not. In pseudo-code:

```
bool transmitOptionEnabled = localTransmitOffered & remoteReceivedOffered;
bool receiveOptionEnabled = localReceivedOffered & remoteTransmitOffered;
```

Note that an implementation can independently offer to transmit or receive a media option. For instance, a TIP endpoint might offer to transmit the Audio Activity Metric, but not offer to receive this option.

As necessary, new MEDIAOPTS packets can be sent to renegotiate the media options in use. This can occur in a multipoint conference due to the entry or departure of an endpoint with more limited capabilities that the other members of the conference.

Note that any intermediate entity that has enabled media options and is forwarding the media packets to other entities must take responsibility for reversing these options if the downstream receiver has not enabled them – i. e., the burden of interoperability with standards compliant implementations falls on the option enabler.

4.2.5.1 Audio Dynamic Output Channels

Typically, audio RTP streams are statically mapped to an output channel (speaker) configuration. In the highly dynamic environment of a multi-channel, multipoint conference that static mapping is suboptimal. TIP supports the capability of multiple audio streams with dynamically changing output channel mappings . This enables TIP devices to mix multiple audio streams when more than one stream is mapped to a particular output channel.

To allow for highly dynamic switching of audio streams to different output channels, this media option enables the interpretation of the output position of the RTP MUX-CSRC field. The bitmask 0x00000F00 is used to select the bits of the MUX-CSRC that represent the desired output channel. The position values defined in Table 2 apply for this position.

4.2.5.2 Audio Activity Metric

TIP supports devices capable of generating and receiving a real-time voice activity confidence metric for audio input channels. This media option transmits that 1 byte metric at the end of the standard RTP audio payload for the codec in use. If more than one audio frame is sent per packet, the metric is the average value of the included audio frames.

Implementations that allow for transmission of this media option should ensure that the additional byte does not induce IP layer fragmentation due to MTU limitations.

4.2.5.3 G.722 Legacy Audio

TIP supports devices capable of sending or receiving an extra audio stream encoded with a different encoding format to that negotiated in SIP. This is achieved within the multiplex via use of the Legacy Mix position. By default, the encoding of this audio stream is G.711. This media option allows signaling of the capability to use G.722 as an alternate encoding. Support for this alternate encoding may be only for transmit, only for receive, or both depending on the capabilities of the system.

If a given transmitter/receiver pair between peers is capable of supporting the G.722 encoding for the legacy audio, then an additional media option bit is defined to signal the actual desire to use the capability.

Note that these two audio encodings use different, static RTP payload numbers (G.711 uses 0 and G.722 uses 9). Also no encoding specific parameters are needed. Hence these two media option bits are sufficient to accomplish the equivalent of an SDP offer/answer [2] negotiation between these two encodings. Receivers can also easily distinguish the encodings to enable graceful handling of mid-call media transitions between the two audio encodings.

4.2.5.4 Video Refresh Flag

For multipoint sessions, it is desirable for receivers to be able to easily distinguish video refresh points in the RTP packet stream. This can be done with video codec and RTP payload specific knowledge, but it requires fairly complicated code. To ease this task, TIP provides an option where such video refresh packets can be marked with an explicit refresh flag. This flag is carried in a single byte appended to the end of the standard RTP video payload for the codec in use. The extra byte is sent for every video packet, with the byte set as follows:

Byte	Description
0	Not a refresh packet
1	First packet of an IDR frame
2	First packet of a GDR frame
3	First packet of an LTRP 0 candidate frame
4	First packet of an LTRP 1 candidate frame

5	First packet of a repair frame associated with LTRP 0
6	First packet of a repair frame associated with LTRP 1
7	First packet of a repair frame not using LTRP

Implementations that allow for transmission of this media option should ensure that the additional byte does not induce IP layer fragmentation due to MTU limitations.

4.2.5.5 Video Parameter Sets

To support highly dynamic video switching among a set of compatible video streams, there is a need to have video configuration information travel in-band with the video data. In H.264 this video configuration information is contained in "parameter sets", specifically the sequence parameter set (SPS) and the picture parameter set (PPS). Since video switching can only usefully occur at video refresh points (IDR frames), the video parameter sets should accompany these frames. When enabled, this media option instructs the video transmitter to include the parameter sets with the refresh frames.

This in-band option for parameter sets is allowed by [7] but only when a means of providing reliable transport exists. For TIP devices, this reliable transport is accomplished via the RTP feedback mechanism described later in this document.

4.2.5.6 Video Coding Mode

TIP allows devices which operate in H.264 baseline profile ,which is defined by the ITU to use context adaptive variable length coding (CAVLC), to use context adaptive binary arithmetic coding (CABAC), that yields increased performance, i.e., a lower bitrate. Although CABAC is part of the H.264 standard, it is not allowed under a strict (i.e., interoperable) definition of the baseline profile. When enabled this media option allows the video codec to use CABAC in conjunction with the baseline profile.

4.2.5.7 Video Long Term Reference Pictures

This media option enables use of this encoding feature if both peers are capable of using it.

4.2.5.8 Aux Video FPS

TIP supports three discrete framerate values for auxiliary video streams;1 frame per second, 5 frames per second, or 30 frames per second. A TIP device will advertise the highest the number of frames per second it can support, lower levels are supported implicitly. This advertisement is done using the two Aux Video FPS bit as shown in the following table:

Bit 0	Bit 1	Description
0	0	5 FPS
0	1	1 FPS
1	0	30 FPS
1	1	N/A

4.2.5.9 Gradual Decoder Refresh

This bit indicates that the sender supports sending (TX options) or receiving (RX options) gradual decoder refresh frames. When enabled on the transmitter, the video encoder will generate GDR frames in place of IDR frames.

4.2.5.10 Options Tags and Values

The media options tags are assigned as follows:

Tag	Description
0	Reserved
1	Transmitter Profile (see below for values)
2	N/A
3	N/A
4	N/A
5-255	Unused

The transmitter profile values are assigned as follows:

Value	Description
0	Reserved
1	Satellite deployment

Note: The purpose of the current options is to enable TIP devices to adapt their network error condition algorithms to values that are appropriate for varying deployment scenarios that can occur in different calls/conferences.

4.2.6 Acknowledgement

Most of the RTCP APP extensions described in this document benefit from having an explicit acknowledgement from the receiver. In this section, a generic ACK packet is defined for use with these extensions and any future extensions that may require it.

The most significant bit of the RTCP APP subtype is defined to represent an ACK bit. The least significant nibble (4 bits) of the subtype is used to carry the extension identifier of the packet being acknowledged. This results in the following allocation of the subtype values for TIP extensions.

Subtypes within the 'xcts' namespace

Subtype	Extension
0	Reserved
1	MUXCTRL
2	Deprecated
3	Deprecated
4	ECHO
5	TXFLOWCTRL
6	RXFLOWCTRL
7	MEDIAOPTS
8	REFRESH
9	Deprecated
10	Reserved
11	Reserved
12	Reserved
13	Reserved
14	Reserved
15	Unused
16	Reserved
17	MUXCTRL ACK
18	Reserved
19	Deprecated
20	Reserved for ECHO ACK
21	TXFLOWCTRL ACK
22	RXFLOWCTRL ACK
23	MEDIAOPTS ACK
24	REFRESH ACK
25	Reserved

26	Reserved
27	Reserved
28	Reserved
29	Reserved
30	Reserved
31	Unused

RTCP APP ACK packet format

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|V=2|P| subtype |   PT=APP=204  |             length            |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                              SSRC                             |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                           name=xcts                           |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|         Transmit NTP timestamp, most significant word         |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|         Transmit NTP timestamp, least significant word        |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

RTCP APP Packet field definitions, should follow [4] and:

`subtype`: 5 bits

MUST be set to the subtype field of the RTCP APP packet being acknowledged plus 16 (bitwise-or of 0x10)

`name`: 4 bytes

MUST be set to the value ‘xcts’.

ACK Packet field definitions:

`transmit ntp timestamp`: 64 bits

The NTP time contained in the RTCP APP extension packet being acknowledged.

Reception of RTCP APP packets with NTP timestamps that are older than the last ACK transmitted SHOULD NOT be acknowledged.

Reception of RTCP APP packets with NTP timestamps equal to the last ACK transmitted SHOULD be re-acknowledged (as the ACK may have been lost).

4.3 RTP Feedback

Depending on the capabilities of the encoder to adjust to packet loss, TIP devices may use an RTP profile such as [8] that allows for timely receiver feedback on specific packet reception or loss. Use of this profile is indicated in the RTCP APP MUXCTRL packet discussed in section 4.2.1 by setting the profile field value to ‘2’ signaling the RTP/AVPF profile.

When used, TIP devices will generate transport layer feedback messages (RTCP packet type RTPFB). The FMT parameter of these messages types is set to private value 30, and a custom feedback control information (FCI) block is used.

RTCP RTPFB packet format with custom FCI

```
0                   1                   2                   3
   0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
   |V=2|P|  FMT=30 |  PT=RTPFB=205 |          length               |
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
   |                  SSRC of packet sender                        |
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
   |                  MUX-CSRC of media source                     |
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
   |              PID              |  PPA, most signifcant 16 bits |
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
   |                              PPA                              |
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
   |                              PPA                              |
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
   |                PPA, least significant 32 bits                 |
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
   |           Reserved            | PPAm, most significant 16 bits|
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
   |                              PPAm                             |
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
   |                              PPAm                             |
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
   |                PPAm, least significant 32 bits                |
```

```
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

RTCP RTPFB fields:

See [8].

FCI fields:

`packet ID (PID)`: 16 bits

The RTP sequence number of the media SSRC that is being acknowledged as received.

`previous packet acknowledgements (PPA)`: 112 bits

A bitmask indicating whether the previous 112 RTP packets relative to the packet identified by the PID field have been received or not. The i'th least significant bit of the PPA refers to the packet with RTP sequence number (PID –i) modulo 2^16. A bit value of 0 indicates no packet reception, a bit value of 1 indicates reception.

The length of this bit field is designed to provide adequate redundancy against loss of the feedback packets themselves up to the expected maximum bitrate of a typical session.

`previous packet acknowledgements mask (PPAm)`: 112 bits

A bitmask indicating which of the 112 bits in the PPA field are valid. The i'th least significant bit of the PPA refers to the packet with RTP sequence number (PID –i) modulo 2^16. A bit value of 0 indicates the corresponding PPA bit is invalid, a bit value of 1 indicated it is valid. Invalid PPA bit values should be ignored by the receiver as they do not indicate an ACK or a NACK.

The PPAm field is optional, its presence should be detected by the length of the RTPFB packet.

The primary purpose of the PPAm field is to allow RTPFB packets to be sent out of order without implicitly indicating reception or no reception for the previous packets in the PPA bitfield. A PPAm bit set to zero indicates that the correspoinding PPA bit CAN NOT be used to establish whether the associated packet has been received or not.

TIP only supports the feedback profile for video streams. The TIP RTCP RTPFB packets should be generated by the receiver at the time that a video frame is consumed. For a typical implementation of feedback, this is when the packets corresponding to a video frame are dequeued from the receiver's jitter buffer. The number of RTPFB packets generated should have a one-to-one correspondence with the received video frame rate. Note however that the PPA field will correspond to both the current video frame, and a number of previous video frames. This is intentional, and provides for retransmission of the ACK/NAK data to protect against loss of the RTPFB packets themselves.

Note that each of the TIP specific RTCP RTPFB messages SHOULD be sent within compound RTCP packets as per [4] and [8]. The simplest way to accomplish this is to prepend an "empty" receiver report (RR) and session description (SDES with CNAME) before the APP packet.

RTCP Packet with empty RR, SDES CNAME and TIP RTPFB

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|V=2|P|  RC=0   |    PT=RR=201  |          length=1             |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                              SSRC                             |
+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+
|V=2|P|  SC=1   |  PT=SDES=202  |          length               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                              SSRC                             |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|    CNAME=1    |     length    |  user and domain name      ...
+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+
|V=2|P|  FMT=30 | PT=RTPFB=205  |          length               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                  SSRC of packet sender                        |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                  MUX-CSRC of media source                     |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|              PID              |  PPA, most signifcant 16 bits |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                             PPA                               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                             PPA                               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                PPA, least significant 32 bits                 |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

4.4 Positional Rules

4.4.1 Control Position

The control position (value 0) of the multiplex is reserved for use with RTCP packets that carry TIP APP extension packets that pertain to the entire multiplex and not a specific media stream within the multiplex, e.g., MUXCTRL.

RTP packets should never use the control position for either the transmitter or receiver positions. They may use it as an "undefined" value for the output position when that option is not used.

4.4.2 Center, Left, Right Positions

The center (value 1), left (value 2), and right (value 3) positions correspond to the physical segments of a typical “three screens” TIP device. A typical “single screen” TIP device will only have the center position available.

4.4.3 Auxiliary Position

The auxiliary positions are typically used for two purposes: local 3-way audio conferencing and audio and video presentation(aka auto-collaborate). These functions are not always used in a call, and hence the availability of these positions dynamically changes within the TIP multiplex. Note if a TIP device is performing both functions the auxiliary audio stream is a mix of the presentation audio and the conference sources.

With respect to audio, TIP endpoints will typically offer to receive an auxiliary position stream. They will transmit an auxiliary position stream, when one of the functions that require it becomes active. A new MUXCTRL packet will be sent to the peer, and assuming the peer offers to receive the auxiliary position, the media stream will be transmitted.

With respect to video, a TIP endpoint will either offer to transmit or receive the auxiliary position, but not both simultaneously. The TIP model is that there is only one active presentation in a call at any given time, and the recommended user experience policy is that the last unit to assert a new, active source should be given the floor. Hence, a TIP endpoint can send a number of mid-call MUXCTRL packets to update the availability of the auxiliary position video transmitter and receiver, given the user’s actions with respect to enabling and disabling the presentation video at that endpoint.

Only one presentation requestor (MUXCTRL with auxiliary position transmitter enabled), is sent a MUXCTRL from a TIP MCU, which makes the auxiliary position receiver at the MCU available. All other TIP endpoints receive a MUXCTRL, which indicates that the auxiliary receiver position is not available at the MCU peer. When the current presenting endpoint releases the floor (MUXCTRL with auxiliary position transmitter disabled), then the MCU will allow the next most recent presenter to regain control of the floor, by sending that endpoint a new MUXCTRL which indicates availability of the auxiliary position receiver at the MCU. The new endpoint can either transmit or indicate that it is no longer interested in transmitting auxiliary video (via MUXCTRL).

4.4.4 Legacy Positions

The legacy video positions (values 9-11) provide TIP devices the ability to transmit lower image resolution versions of the HD video streams. If enabled, they allow an MCU to simultaneously provide both high definition video streams to TelePresence sites and a standard definition streams to devices that do not support HD video.

Only if both peers offer the legacy positions will the streams be transmitted from the TIP endpoint to the MCU. The positional MUX-CSRC of the video legacy stream will have the legacy value for both the transmitter and receiver positions.

The HD video decoders should be capable of handling SD video.

An MCU can transmit a legacy video stream to the center, left, or right position video receivers. The inband SPS/PPS should be inserted immediately before the IDR frame that begins the first frame of the new image size video. Hence, the MCU is able to map the legacy video stream to any available display according to the user experience policies of the product.

The positional CSRC of the video stream transmitted from the MCU to the TIP endpoint should have the legacy value for the transmitter position, and the selected physical position (center, left , or right) for the receiver position.

4.5 RTCP Rules

TIP devices operating under TIP will generate RTCP sender and receiver reports (SRs and RRs) in alignment with [4]. The receipt of the sender reports is used by TIP receivers to achieve correct synchronization of the media streams (though a less precise fallback mechanism should be used if no SRs are received).

5 Call Flow Examples

5.1 Basic TIP MUX Setup

Note that the sequence of MUXCTRL and MEDIAOPTS exchanges is just an example; several other possibilities exist as each peer's exchange is independent. It is required that a successful MUXCTRL occur before any other TIP MUX extension, including MEDIAOPTS, is sent.

```
Peer A                    Peer B
  |                         |
  |         MUXCTRL         |
  |------------------------>|
  |       MUXCTRL ACK       |
```

```
|<-----------------------|
|        MUXCTRL         |
|<-----------------------|
|      MUXCTRL ACK       |
|----------------------->|
|                        |
|       MEDIAOPTS        |
|----------------------->|
|     MEDIAOPTS ACK      |
|<-----------------------|
|       MEDIAOPTS        |
|<-----------------------|
|     MEDIAOPTS ACK      |
|----------------------->|
|                        |
|        RTP/RTCP        |
|<======================>|
|                        |
```

6 References

The following documents provide the context for the issues discussed in this document:

[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 DRAFT draft-ietf-avt-rtp-and-rtcp-mux-07 “Multiplexing RTP Data and Control Packets on a Single Port”

End of Document