



Cisco TelePresence Secure Communications and Signaling

Created: 9/17/09, OL-20744-01

Contents

Authentication and Encryption Framework Overview	2
Digital Certificates	3
Certificate Authority Proxy Function (CAPF)	4
CTL Provider, CTL Client, and the Certificate Trust List	4
Security Profiles	4
Device Security Modes	5
Configuration File Integrity and Encryption	5
Enabling Security for TelePresence Deployments	5
Security Profile Example	7
Cisco TelePresence Security Protocol Details	10
Transport Layer Security (TLS)	11
Secure RTP (SRTP) and Secure RTCP (SRTCP)	12
SRTP Key Exchange	15
Key Exchange via SDP messages (SDES)	15
Key Negotiation/Exchange via DTLS-SRTP	16
Key Exchange via Encrypted Key Transport (EKT)	18
TelePresence Multipoint Call Operation	19
Network Design Implications	21
Bandwidth Considerations	21
Video Media Streams	21



Americas Headquarters:
Cisco Systems, Inc., 170 West Tasman Drive, San Jose, CA 95134-1706 USA

Audio Media Streams	21
RTCP Streams	22
Overall Estimated Bandwidth Impact	22
Jitter and Delay	23
Encryption of Point-to-Point, Multipoint, and Interoperability Meetings	23
Network Visibility	26
Call Signaling Visibility	26
Media Visibility	26
Secure RTP versus IPSec Encryption	27
Summary	30
Glossary	30
Reference Documents	30

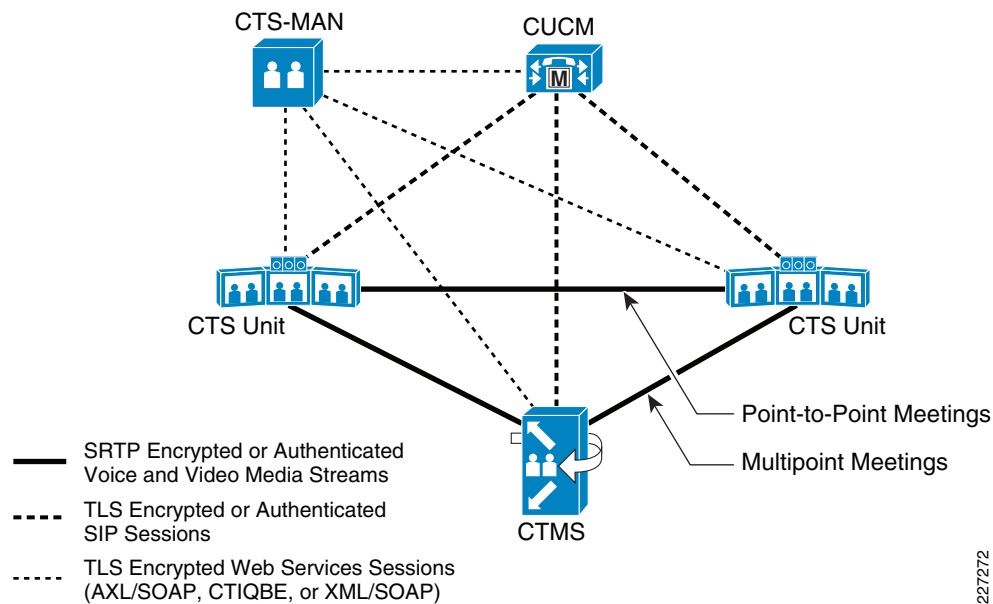
The Cisco TelePresence solution provides a robust framework for providing secure communications and signaling, encompassing much more than just media encryption. This framework was first deployed in the Cisco Unified Communications Manager (CUCM) Release 4.0 to support encryption of voice calls, and has since proven to be a reliable and comprehensive security architecture for supporting corporate-wide IP telephony deployments. Cisco TelePresence leverages this same Unified Communications security framework both for intra- and inter-company communications. This document focuses on the deployment of authentication and encryption for intra-company communications; although the protocols and functionality discussed equally apply to inter-company communications. It assumes the reader is not familiar with the authentication and encryption framework for Cisco IP Telephony, and therefore provides a high-level overview of the framework and the process for enabling authentication and encryption; specifically focused on TelePresence devices. In addition, the network design implications of enabling TelePresence secure communications and signaling are discussed.

Authentication and Encryption Framework Overview

As of CTS Version 1.5, Cisco TelePresence deployments support the security of all currently supported signaling and media streams through the following capabilities:

- Data authentication and confidentiality of the RTP voice and video media flows, using the Secure Real-time Transport Protocol (SRTP), for both point-to-point and multipoint TelePresence meetings.
- Data authentication and confidentiality of the SIP signaling between the CUCM and Cisco TelePresence System (CTS) endpoints, and between the CUCM and the Cisco TelePresence Multipoint Switch (CTMS); using Transport Layer Security (TLS).
- Data authentication and confidentiality of the web services signaling between the Cisco TelePresence System Manager (CTS-MAN), CTMS, CUCM, and CTS endpoints using TLS.

A summary of these capabilities is shown in [Figure 1](#).

Figure 1 Cisco TelePresence Secure Media Flows

The following sections briefly describe the various components that together provide the framework for providing secure communications for Cisco TelePresence deployments.

Digital Certificates

A digital certificate is a mechanism for providing authentication of a particular entity (device, application, etc.), using public-key cryptography. In public-key cryptography, entities generate public/private key pairs that have a unique feature—messages encrypted with the private key can only be decrypted with the public key, and vice versa. The private key is kept secure within the device and never exposed. The public key can be bound to one or more attributes relating to the identity of the entity to form a digital certificate. The binding is accomplished by a Certificate Authority which digitally signs the certificate. The digital signature itself is a hash of the message, encrypted via the private key of the Certificate Authority. The digital signature of the Certificate Authority can be verified by the recipient via the public key of the Certificate Authority. The format of a digital certificate is standardized through the ITU-T and ISO/IEC X.509 v3 standard.

Before any secure communications can occur between two TelePresence endpoints or between a TelePresence endpoints and the CUCM, a framework of authentication and trust must first be established. Each CTS unit (CTS-3200, CTS-3000, CTS-1300, CTS-1000, or CTS-500) comes from the factory with an X.509 v3 digital certificate installed within it, signed by the Cisco manufacturing certificate authority. This is known as a Manufacturer Installed Certificate (MIC). This certificate can provide the credentials used by the CTS units to perform a first-time authentication and enrollment into the security framework provided by the CUCM.



Note

The MIC is also used for establishing Datagram Transport Layer Security (DTLS) session between TelePresence endpoints. DTLS is used to negotiate Secure Real-time Transport Protocol (SRTP) master keys. DTLS is discussed further in following sections.

The CUCM also has installed certificates that are used to authenticate it to the TelePresence endpoints. Additional Locally Significant Certificates (LSCs) are installed within TelePresence endpoints by the Cisco Certificate Authority Proxy Function (CAPF) of the CUCM. This is briefly discussed in the following sections.

Certificate Authority Proxy Function (CAPF)

The Cisco Certificate Authority Proxy Function (CAPF) is a software service, installed as part of the CUCM. CAPF issues Locally Significant Certificates (LSCs) for TelePresence endpoints. CAPF can create certificates under its own authority, or it can be used as a proxy to request certificates from an external Certificate Authority (CA); and then proxy those certificates to the TelePresence endpoints and the CUCM servers. These certificates are then used to establish secure, authenticated connections for protocols such as SIP signaling over TLS.

CTL Provider, CTL Client, and the Certificate Trust List

The CTL Provider is another software service, installed as part of CUCM, which works together with the CTL Client to generate a Certificate Trust List (CTL). The CTL Client is a software plugin that can be downloaded from the CUCM server and run on a separate PC. The Certificate Trust List itself is a pre-defined list of trusted certificates stored on the CUCM server, which is downloaded as a file to the TelePresence endpoints when they boot up. The CTL indicates the list of CUCM servers that the TelePresence endpoints can trust when they initiate SIP sessions over TLS for call signaling. In order to provide authentication for the CTL itself, a minimum of two separate Cisco USB hardware security keys (tokens) need to be purchased from Cisco. These security keys are inserted into the PC running the CTL Client plugin during the CTL generation process.

Security Profiles

Security profiles are sets of security attributes that can be configured once, and then applied to multiple TelePresence endpoints; versus having to configure the same set of security attributes individually to each device. There are two types of security profiles used within Cisco TelePresence deployments.

- **Phone Security Profiles** must be defined for both the TelePresence primary codecs and their associated IP 7975G Phones in order to enable security on the CTS units.
- **SIP Trunk Security Profiles** must be defined in order to enable secure connectivity to CTMS devices, since they are configured as SIP Trunks to CUCM.

Security profiles include the desired device security mode for the TelePresence endpoint and the method by which the TelePresence endpoint should initially authenticate itself to the CAPF service within CUCM. There can be multiple instances of each type of security profile (i.e., multiple phone security profiles or multiple SIP trunk security profiles). For example, if some CTS units are dedicated for business-to-business (B2B) use, they may have a security profile different from CTS units dedicated for internal company use.

Device Security Modes

TelePresence devices use the security profiles defined within CUCM in order to enable one of the following three device security modes:

- *Non-Secure*—In this mode, neither message authentication nor encryption is enabled for the SIP signaling, the Real-time Transport Protocol (RTP) voice and video media streams, and the Real-time Transport Control Protocol (RTCP) control streams.
- *Authenticated*—In this mode, message authentication is enabled for the SIP signaling between the TelePresence device and the CUCM via TLS. However, encryption is not enabled. Further, in authenticated mode, neither message authentication nor encryption is enabled for the RTP voice and video media streams and the RTCP control streams.
- *Encrypted*—In encrypted mode, message authentication and encryption are enabled for the SIP signaling via TLS. Message authentication and encryption are also enabled for the RTP voice and video media streams via Secure RTP (SRTP).



Note

Only message authentication is enabled for most RTCP packets in encrypted mode, since intermediary devices such as the CTMS may require visibility into the RTCP flows. However, note that some RTCP packets may contain encrypted payloads as well.

Configuration File Integrity and Encryption

The configuration files for CTS units, as well as their associated IP Phones, are stored within CUCM. These files are downloaded to the endpoints each time they boot up. Configuration files are also automatically downloaded to TelePresence devices any time a change in configuration is made within CUCM that would affect the endpoint's configuration; and the device is subsequently reset. The configuration files can be optionally encrypted through the security profile, so that only the intended TelePresence device (CTS codec or IP Phone) can read it. Furthermore, it can be digitally signed by the CUCM so that the TelePresence endpoint can trust that the file has not been altered.



Note

The CTS-MAN and CTMS do not store their configuration files on the CUCM. Therefore, the option for encryption of configuration files does not apply for these devices.

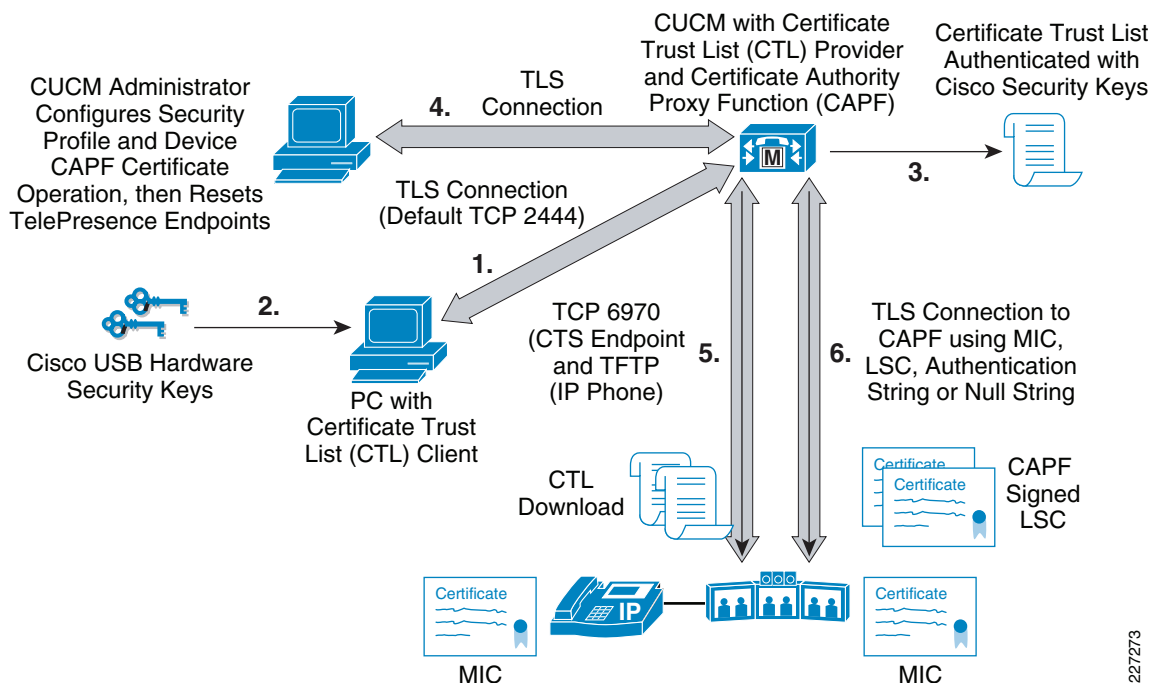
Enabling Security for TelePresence Deployments

This section provides a high-level overview of how the security components discussed in the previous sections work together to enable encryption and authentication for CTS endpoints. This section is not intended to be a detailed guide to enabling TelePresence security. For detailed procedures for enabling security on the CTMS, CTS Manager, and CTS endpoints, refer to the *Cisco TelePresence Security Solutions* document at the following:

http://www.cisco.com/en/US/docs/telepresence/security_solutions/CTSS.pdf

Figure 2 shows a high-level overview of the process for enabling authentication and encryption for CTS units. Note that the process has been somewhat simplified for clarity. To assist the reader in understanding Figure 2, the steps involved have been numbered and explained.

Figure 2 Enabling Authentication and Encryption for CTS Endpoints



The process begins when the network administrator executes the following:

- Downloads the CTL Client Plugin from the **Plugins** screen of the **Cisco Unified Communications Administration** window of the CUCM server
- Installs and launches the application on a separate PC.
- Uses the CTL Client to establish a TLS connection to the CTL Provider service running within CUCM (Step 1 in Figure 2).



Note

The TLS connection uses TCP port 2444 by default; however, the port is configurable within CUCM.

Together the CTL Client and CTL Provider are used to generate the Certificate Trust List (CTL) for the CUCM cluster. During the process, the network administrator is prompted to insert the Cisco USB hardware security keys that are used to authenticate the CTL file, which is then stored on the CUCM server (Steps 2 and 3 in Figure 2).



Note

A minimum of two security tokens are required in case one is lost.

Once the Certificate Trust List is completed, the network administrator must configure security profiles for both the TelePresence primary codec and its associated IP 7975G Phone; apply the security profiles to the devices; and reset the devices (Step 4 in Figure 2). The [“Security Profile Example”](#) section on page 7 shows an example of a security profile.

227273

When the TelePresence devices are reset, they download the CTL file and a partial configuration file from CUCM. TelePresence primary codecs use TCP port 6970, while IP 7975G Phones use TFTP, for the downloads (Step 5 in [Figure 2](#)). The partial configuration file indicates to the device that it needs to download an LSC. The devices then establish a TLS connection to CAPF, via TCP port 3804, using the authentication method specified within the configuration, as shown in the security profile example. They generate public/private key pairs, and the CAPF service binds the public keys with the device information, signs it, and returns the digital certificates to the devices. This results in the Locally Significant Certificates (LSCs) being installed into the CTS endpoints (Step 6 in [Figure 2](#)). Note that this occurs for both the TelePresence primary codecs and their associated IP 7975G Phones. After this, the endpoints download the CTL file again, and the full and encrypted configuration file from the CUCM. Note that this is not shown in [Figure 2](#) above. At this point, the CTS endpoint is ready to establish secure SIP sessions to the CUCM.

**Note**

During TelePresence meetings, audio media is not generated by the associated IP 7975G Phone itself. However, SIP signaling does exist between the IP 7975G Phone and the CUCM cluster. If a security profile is not defined and applied to the associated IP 7975G Phone, SIP signaling will not be protected. Further, any file downloads to the associated IP 7975G Phone will not be encrypted.

Security Profile Example

[Figure 3](#) shows an example security profile for a CTS-1000 codec. A similar profile must be defined for the IP 7975G IP Phone associated with the CTS-1000.

Figure 3 Example Security Profile for a CTS-1000 Codec

The screenshot displays the Cisco Unified CM Administration web interface. The top navigation bar includes links for System, Call Routing, Media Resources, Voice Mail, Device, Application, User Management, Bulk Administration, and Help. The main content area is titled "Phone Security Profile Configuration" and includes a "Save" button and a "Status: Ready" indicator. The "Phone Security Profile Information" section contains fields for Product Type (Cisco TelePresence 1000), Device Protocol (SIP), Name (Cisco TelePresence 1000 SIP By Auth String), Description (Cisco TelePresence 1000 SIP By Auth String), Nonce Validity Time (600), Device Security Mode (Encrypted), and Transport Type (TLS). Checkboxes for "Enable Digest Authentication", "TFTP Encrypted Config" (checked), and "Exclude Digest Credentials in Configuration File" are also present. The "Phone Security Profile CAPF Information" section shows Authentication Mode (By Authentication String) and Key Size (1024). The "Parameters used in Phone" section includes SIP Phone Port (5060). A "Save" button is at the bottom, and a note indicates that asterisks (*) denote required items.

As shown in Figure 3, the network administrator must choose the **Device Security Mode** in order to enable one of the three modes: non secure, authenticated, or encrypted. The **TFTP Encrypted Config** checkbox causes CUCM to encrypt the configuration file that is sent to the primary codec. The **Phone Security Profile CAPF Information** section is used to instruct the primary codec how to establish the initial authentication to the CAPF service within CUCM. The following authentication methods are supported between the CAPF service itself and devices:

- *By Null String*—This method provides little security, since any device can authenticate with the CAPF service. This is not recommended in production TelePresence deployments.
- *By Authentication String*—This method allows the network administrator to specify a one-time authentication string which is used to authenticate the TelePresence device to the CAPF service.
- *By Existing Certificate (Precedence to MIC)*—This method utilizes the manufacturer installed certificate on the device, in preference to any locally installed certificate on the device.
- *By Existing Certificate (Precedence to LSC)*—This method utilizes a locally installed certificate on the device, in preference to any manufacturer installed certificate on the device.

Once the security profiles have been defined for the TelePresence endpoints, they must be applied to their device configurations within CUCM. Figure 4 shows an example of part of the device configuration of a CTS-1000 codec, indicating where the device security profile is applied.

Figure 4 Example Device Security Profile Application within CTS Device Configuration

Protocol Specific Information	
Packet Capture Mode*	None
Packet Capture Duration	0
Presence Group*	Standard Presence group
SIP Dial Rules	< None >
MTP Preferred Originating Codec*	711ulaw
Device Security Profile*	Cisco TelePresence 1000 SIP By Auth String
Rerouting Calling Search Space	< None >
SUBSCRIBE Calling Search Space	< None >
SIP Profile*	Standard SIP Profile
Digest User	< None >
☐ Media Termination Point Required ☐ Unattended Port	

227275

CAPF must also install the locally significant certificate (LSC) needed for TLS authentication. [Figure 5](#) shows an example of the part of the device configuration of a CTS endpoint which triggers the installation of the LSC.

Figure 5 Example CAPF Section of CTS Device Configuration

Certification Authority Proxy Function (CAPF) Information	
Certificate Operation*	Install/Upgrade
Authentication Mode*	By Authentication String
Authentication String	7368139635
Generate String	
Key Size (Bits)*	1024
Operation Completes By	2009 8 16 12 (YYYY:MM:DD:HH)
Certificate Operation Status: None	
Note: Security Profile Contains Addition CAPF Settings.	

227276

The authentication mode selected within the configuration must match the authentication mode selected within the security profile applied to the CTS endpoint. Note that if the *By Authentication String* method has been chosen, the same string must be manually configured within the CTS endpoint via the GUI interface, in order for CAPF authentication to be successful. [Figure 6](#) shows an example of the **Unified CM Settings** screen within the TelePresence primary codec, where the authentication string is defined.

Figure 6 Configuration of CAPF Authentication String within the CTS Endpoint

The screenshot displays the Cisco TelePresence System Administration web interface. The top navigation bar includes the Cisco logo, the title "Cisco TelePresence System Administration", and user options: "admin", "Logout", "Help", and "About". The left sidebar contains a menu with categories like "Device Information", "Configuration", "System Status", and "Troubleshooting". The main content area is titled "Configuration > Cisco Unified Communications Manager Settings". It features a section for "Use Configuration TFTP Servers" with radio buttons for "Automatic" and "Specify" (selected). Below this are five TFTP server configuration rows. Further down, there are five rows for "Cisco Unified Communications Manager" status, all marked as "Not Available". A red rectangle highlights the "CAPF Authentication String" field, which contains the value "7368139635". Below this is a "Certificate Trust List (CTL)" field with a "Delete CTL" button. At the bottom right of the configuration area are "Apply" and "Reset" buttons. A vertical text "227277" is visible on the right edge of the interface.

Note that by using the MIC for initial authentication to the CAPF service, the network administrator can avoid having to “touch” every CTS unit; potentially saving administrative time when enabling secure signaling and communications for TelePresence deployments. However, configuring a unique one-time CAPF authentication string for each CTS unit provides a higher level of security, since both sides implicitly authenticate each other based on the shared knowledge of the authentication string.

Cisco TelePresence Security Protocol Details

The following sections provide the reader background information regarding the protocols used to provide secure communications and signaling within TelePresence deployments. These protocols include the following:

- *Transport Layer Security (TLS)*—Used to secure both the SIP call signaling as well as the web services signaling (AXL/SOAP, CTIQBE, and XML/SOAP messages) transported between the various TelePresence components.
- *Secure Real-time Transport Protocol (SRTP) and Secure Real-time Transport Control Protocol (SRTCP)*—Used to secure the actual RTP media flows and their corresponding RTCP control flows.
- *Protocols such as SIP/SDP Security Descriptors (SDS), Datagram Transport Layer Security (DTLS), and SRTP Encrypted Key Transport (SRTP/EKT)*—Used for SRTP master key negotiation and/or exchange.

Transport Layer Security (TLS)

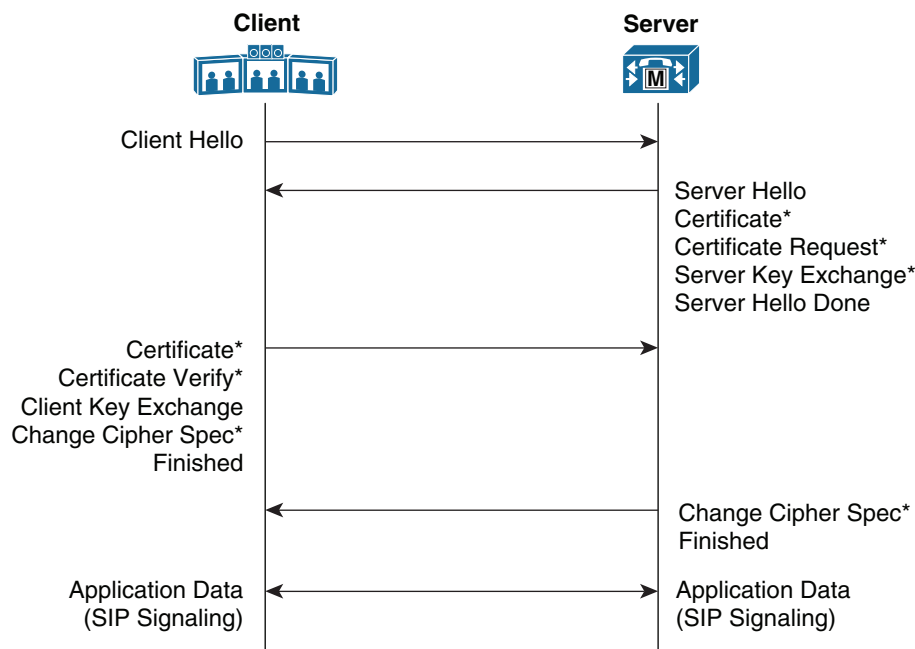
Transport Layer Security (TLS) is a protocol designed to provide authentication, data integrity, and confidentiality for communications between two applications. TLS is based on SSL Version 3.0, although the two protocols are not compatible. The latest version, TLS 1.2 is defined in IETF RFC-5246. TLS operates in a client/server mode with one side acting as the “server” and the other side acting as the “client”. TLS requires a reliable transport layer protocol such as TCP to operate over. The protocol has two primary layers—the TLS Handshake Protocol layer, and the TLS Record Protocol layer.

The TLS Handshake Protocol layer provides the following functionality:

- Allows the client and server to authenticate each other using public key cryptography (i.e. digital certificates). Authentication can be one-sided (the client authenticates the server) or mutual (both the client and server authenticate each other) as is the case with TelePresence.
- Allows the client and server to reliably negotiate a compression algorithm, message authentication algorithm, encryption algorithm, and the necessary cryptographic keys before any application data is sent.

Figure 7 shows a high level overview of the message flows in a TLS handshake from the perspective of the TelePresence endpoint being the client and the CUCM being the server. Note that both the TelePresence primary codec and its associated IP 7975G Phone would both perform a TLS handshake (although this is not shown in Figure 7 for simplicity), since both establish SIP sessions to the CUCM.

Figure 7 *Message Flows for TLS Handshake*



* Optional Messages

227278

The TLS Handshake protocol layer is designed to operate in a lock-step manner, meaning that messages received in incorrect order will cause the TLS handshake to fail. Since TLS is fairly flexible in terms of one-sided authentication versus mutual authentication and whether encryption and message authentication are used or not, many of the messages within the TLS handshake are optional. Bulk encryption of the application data is typically done via encryption algorithms such as the Advanced Encryption Standard (AES), using symmetric keying. Message authentication is typically done via hash

algorithms such as HMAC-SHA1. The negotiation of keying material is done securely within the TLS Handshake Protocol layer through the Client and Server Key Exchange messages. Note that the multiple TLS handshake messages (i.e. Certificate, Certificate Request, and Server Hello Done) appear within a single TCP packet in order to minimize the number of packets sent during the handshake. The TLS Record Protocol layer runs directly on top of a reliable transport layer protocol, such as TCP. It takes application data and fragments it into manageable blocks. The TLS Record Protocol then optionally compresses the block, applies a Message Authentication Code (MAC), and encrypts it; depending upon what was negotiated during the TLS Handshake Protocol for the particular TLS session.

Cisco TelePresence devices utilize TLS to secure the SIP signaling between the TelePresence primary codec and the CUCM server, as well as between the associated IP 7975G Phone and the CUCM server. Mutual authentication of both sides is performed during the TLS handshake, utilizing the LSCs downloaded to the TelePresence primary codec and associated IP 7975G phone. As of CTS Version 1.5, TLS protected SIP signaling uses the AES algorithm with a 128-bit key for encryption and HMAC-SHA1 for authentication. Note that TLS requires the devices to use SIP signaling over TCP, rather than UDP. Therefore, TelePresence endpoints and CTMS devices must use SIP signaling over TCP in order to use TLS for securing such signaling.

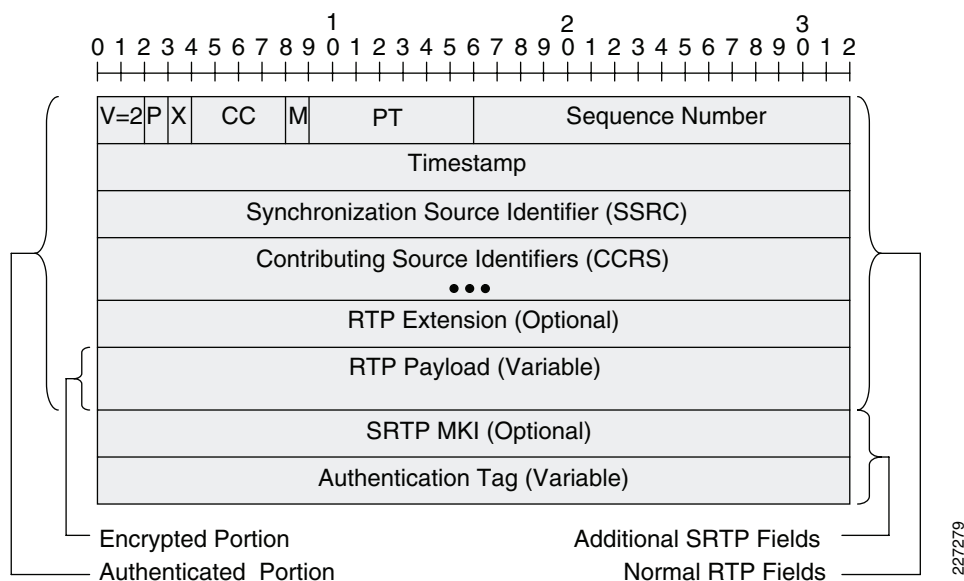
TLS is also used to secure the web services signaling that occurs between the CTMS, CTS-MAN, CTS units, and the CUCM, as shown in [Figure 1](#) above. TLS is also used to secure the connection between the CTS unit and the CAPF server when downloading the LSC. Note however, that the encryption and authentication algorithms used for each of these connections may differ from that used to secure SIP signaling.

Secure RTP (SRTP) and Secure RTCP (SRTCP)

Secure RTP (SRTP) and secure RTCP (SRTCP) are both defined in IETF RFC-3711. RFC-3711 details the methods of providing confidentiality and data integrity for both RTP voice and video media, as well as their corresponding RTCP streams; through the use of encryption and message authentication headers.

The format of an SRTP packet is shown in [Figure 8](#).

Figure 8 *SRTP Packet Format*



The following are the fields that appear within a normal RTP packet:

- *Version (V)*—2-bit field indicating the protocol version. The current version is 2.
- *Padding (P)*—1-bit field indicating padding at the end of the RTP packet.
- *Extension Header (X)*—1-bit field indicating the presence of an optional extension header.
- *Marker (M)*—1-bit marker bit, used to identify events such as frame boundaries.
- *Payload Type (PT)*—7-bit field which identifies the format of the RTP payload.
- *Sequence Number*—16-bit field which increments by one for each RTP packet sent. This field can be used by the receiver to identify lost packets.
- *Timestamp*—32-bit field which reflects the sampling instant of the first octet of the RTP packet.
- *Synchronization Source Identifier (SSRC)*—32-bit field which uniquely identifies the source of a stream of RTP packets.
- *Contributing Source Identifiers (CSRCs)*—Variable length field which contains a list of sources of streams of RTP packets that have contributed to a combined stream produced by an RTP mixer.
- *RTP Extension (Optional)*—Variable length field which contains a 16-bit profile specific identifier and a 16-bit length identifier, followed by variable length extension data.
- *RTP Payload*—Variable length field which holds the real-time application data (i.e. voice, video, etc).

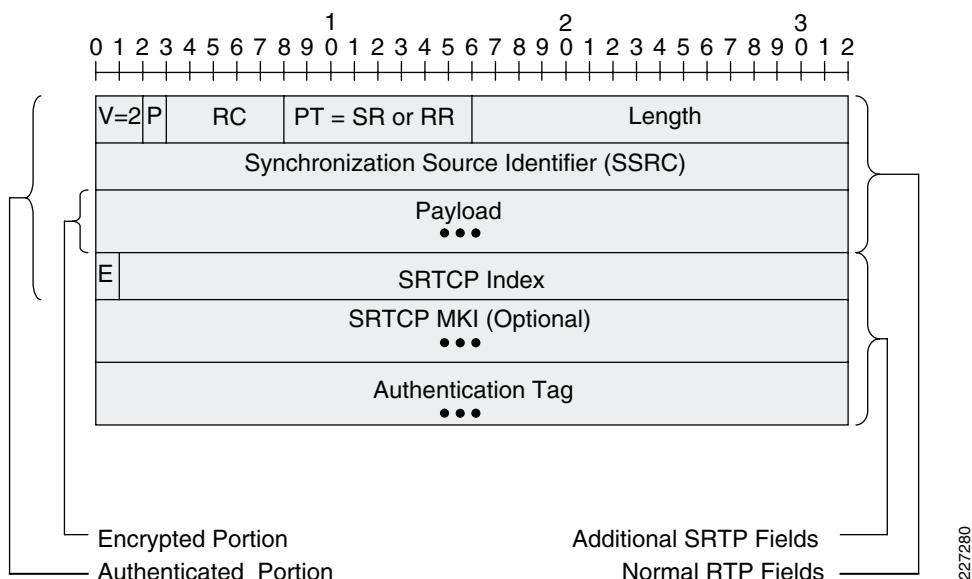
SRTP adds the following two additional fields to the packet:

- *Master Key Identifier (MKI)*—Optional field of configurable length, used to indicate the master key, from which the individual session keys were derived for encryption and/or authentication, within a given cryptographic context.
- *Authentication Tag*—Recommended field of configurable length, used to hold the message authentication data for the RTP header and payload for the particular packet.

With SRTP, encryption applies only to the payload of the RTP packet. Message authentication, however, is applied to both the RTP header as well as the RTP payload. Since message authentication applies to the RTP sequence number within the header, SRTP indirectly provides protection against replay attacks.

The default cipher for encryption defined within RFC-3711 is again AES with a 128-bit key. The AES encryption cipher yields no expansion of the RTP payload itself. The default message authentication code defined within RFC-3711 is HMAC-SHA1. RFC-3711 defines a default authentication tag length of 80 bits (although HMAC-SHA1 produces a 160-bit digest), but smaller tags may be used. Note that the MKI field is currently not used with Cisco TelePresence.

The format of an SRTP packet is shown in [Figure 9](#).

Figure 9 Typical SRTCP Packet Format

The following are the fields that appear within a normal RTP packet:

- *Version (V)*—2-bit field indicating the protocol version. The current version is 2.
- *Padding (P)*—1-bit field indicating padding at the end of the RTP packet.
- *Reception Report Count (RC)*—5-bit field indicating the number of reception report blocks contained within the packet.
- *Payload Type (PT)*—8-bit field which identifies the format of the RTP payload. Compound RTP packets begin with either a Sender Report (SR) or Receiver Report (RR).
- *Length*—16-bit field indicating the length of the RTP packet including the header and any padding.
- *Synchronization Source Identifier (SSRC)*—32-bit field which uniquely identifies the source of a stream of RTP packets.
- *Payload*—Since compound RTP packets typically consist of multiple blocks of data beginning with a SR or RR, the payload can be of variable length containing additional payload types.

SRTCP adds up to the following four additional fields to an SRTCP packet:

- *E-Flag (E)*—Required 1-bit field indicating whether the SRTCP packet is encrypted (1) or unencrypted (0).
- *SRTCP Index*—Required 31-bit field which is incremented for each SRTCP packet sent.
- *SRTCP MKI*—Optional field of configurable length, used to indicate the master key, from which the individual session keys were derived for encryption and/or authentication, within a given cryptographic context.
- *Authentication Tag*—Required field of configurable length, used to hold the message authentication data for the RTP header and payload for the particular packet.

As with SRTCP packets, encryption applies only to the payload of the SRTCP packet, when utilized. Message authentication, however, is applied to both the RTP header as well as the RTP payload.

SRTP Key Exchange

Symmetric keying is used for bulk encryption algorithms such as AES, which in turn are used to encrypt SRTP and SRTCP payloads. The actual process for exchanging keying material for SRTP and is outside the scope of RFC-3711. Prior to CTS Version 1.5, the two methods used by Cisco TelePresence included the use of SIP Session Description Protocol (SDP) messages (sometimes referred to as Security Descriptions or SDES) to exchange keying material within the call signaling during call establishment; and the use of Datagram TLS (DTLS) within the RTP media flows (referred to as DTLS-SRTP) to exchange keying material after call establishment.



Note

TelePresence endpoints use keying material derived from DTLS-SRTP, over keying material derived from SDES if both methods succeed.

These two methods were sufficient for point-to-point TelePresence calls. With CTS version 1.5 and higher, Cisco TelePresence uses Encrypted Key Transport (EKT) for key exchange. EKT is an extension to SRTP, which is necessary to support multipoint Cisco TelePresence calls. Note that EKT still makes use of either DTLS-SRTP or SDES in order to establish the initial SRTP master keys. Each of the three methods is discussed in the sections below.

Key Exchange via SDP messages (SDES)

This method of exchanging keying material is defined within IETF RFC-4568. The RFC defines a method for exchanging the cryptographic suite and keying material information for SRTP flows through SIP Session Description Protocol (SDP) messages during call establishment.



Note

RFC-4568 does not specify how the keys are generated and/or negotiated; but only how the keys are exchanged.

SRTP key exchange is accomplished through the “crypto” attribute (`a=crypto`) under the particular media type (`m=video` or `m=audio`) within the SDP. The format of the “crypto” attribute is shown below.

```
a=crypto:<tag> <crypto-suite> <key-params> [<session-params>]
```

The *tag* field is a decimal number used to uniquely distinguish multiple crypto attributes from each other. The *crypto-suite* field contains the cipher and message authentication algorithms for the particular crypto attribute. The *key-params* field itself has the format shown below.

```
key-params = <key-method> ":" <key-info>
```

The *key-method* field indicates the method by which keying material is exchanged. For example, *inline* indicates the keying information is contained within the SDP. For inline keying, the *key-info* can contain further information in the following format.

```
<key||salt> ["|" lifetime] ["|" MKI ":" length]
```

The *key||salt* field is a concatenation of the master key and master salt.



Note

A *salt* is a series of random bits either used in the key derivation process or appended to the key of a cryptographic algorithm; often used to make dictionary attacks less effective.

The *lifetime* field indicates the lifetime in terms of packets which can use the master key. The *MKI* field indicates whether or not the Master Key Identifier will be included within SRTP packets. Finally, the *length* field indicates the length of the MKI field within SRTP packets.

An example of the use of the crypto attribute within an SDP is shown in the following example:

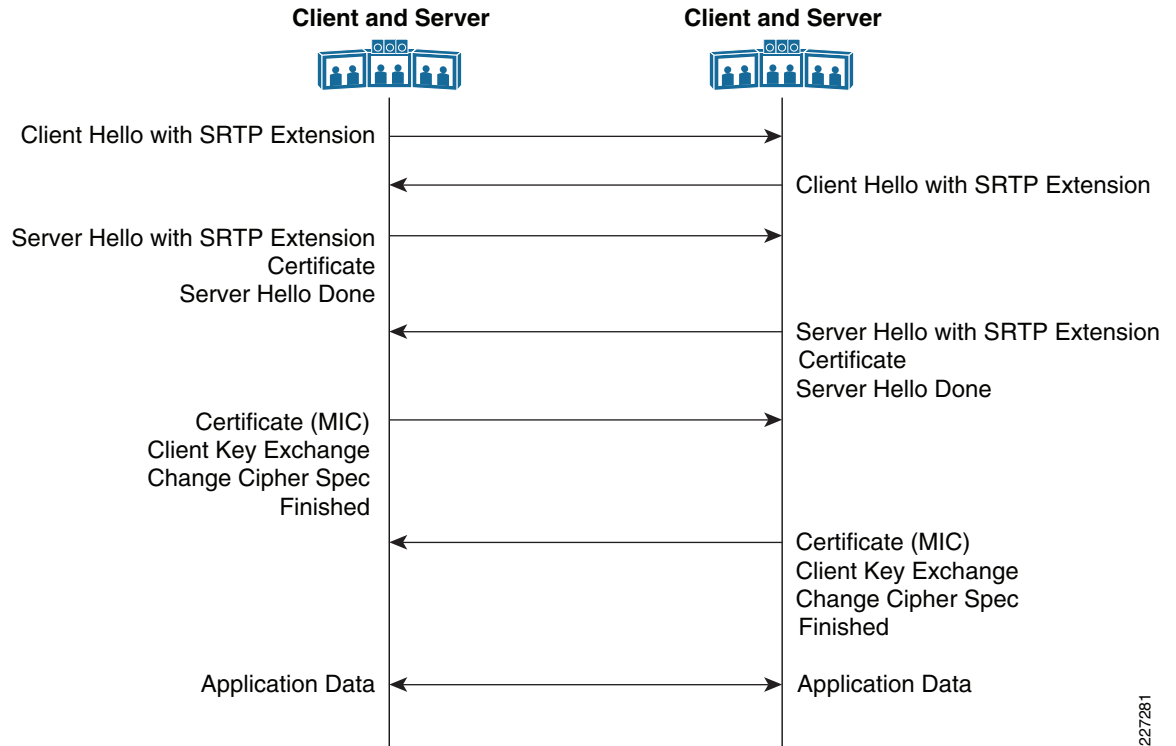
```
m=video 16386 RTP/SAVP 112
a=crypto:1 AES_CM_128_HMAC_SHA1_80
inline:d0RmdmcmVCspeEc3QGZiNWpVLFJhQX1cfHawJSoj|2^20|1:32
```

The reader is encouraged to review IETF RFC-4568 regarding the use of the crypto attribute for this purpose. Since keying material is sent within SIP signaling, encryption of the SIP signaling itself via TLS is necessary to secure the key exchange. The reader should note that Cisco IP Phones only support this method of SRTP key exchange. When an IP Phone is added-on to an existing TelePresence meeting, the encryption keys utilized for the SRTP media flows between the TelePresence primary codec and the add-on IP Phone are exchanged via SDP messages within the SIP signaling.

Key Negotiation/Exchange via DTLS-SRTP

The Datagram Transport Layer Security (DTLS) protocol is designed to provide authentication, data integrity, and confidentiality for communications between two applications, over a datagram transport protocol such as UDP. The protocol is defined in IETF RFC-4347. DTLS is based on TLS, and is designed to provide equivalent security guarantees. However, in order to account for the underlying unreliable transport protocol, mechanisms such as sequence numbers and retransmission capability have been added to the DTLS handshake. DTLS-SRTP is an extension to DTLS which provides for the negotiation of SRTP keying material within DTLS. It is currently an IETF Internet-Draft (<http://www.ietf.org/internet-drafts/draft-ietf-avt-dtls-srtp-07.txt>).

Figure 10 shows a high-level overview of the message flows in a DTLS-SRTP handshake from the perspective of two TelePresence endpoints, both functioning as client and server to each other.

Figure 10 **Message Flows for DTLS-SRTP Handshake**

227281

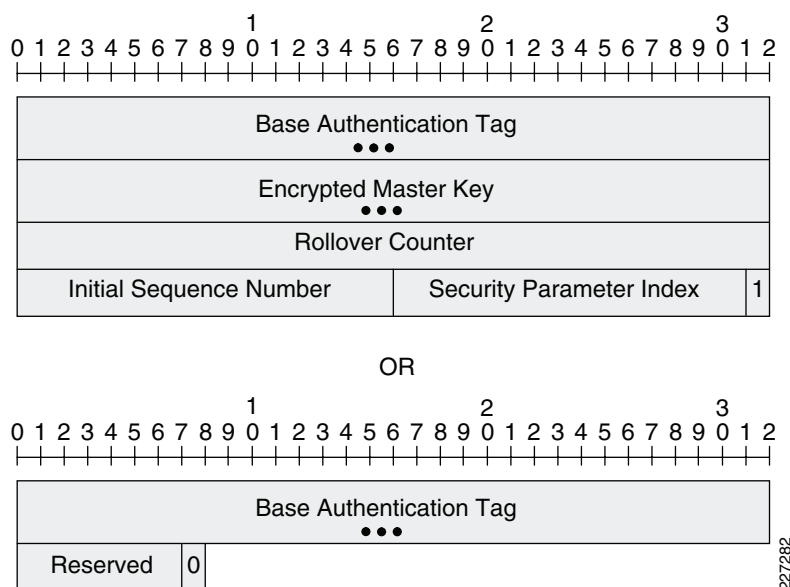
One difference that should be noted from the TLS handshake shown in [Figure 7](#) is that the DTLS handshake occurs directly between the TelePresence endpoints. Therefore, the DTLS-SRTP session is established from TelePresence primary codec to TelePresence primary codec, or between TelePresence primary codec to the CTMS; within the RTP media streams between the two devices. The SRTP keying material, algorithms, and other parameters is securely negotiated directly between the devices. There is no “transitive trust” issue where the SIP back-to-back user agent (B2BUA) within the CUCM participates within the SRTP key exchange (as occurs with SDES); and therefore has knowledge of the SRTP keying material. Although this may not be a significant issue in a TelePresence deployment within a single corporate entity, it is important from a B2B TelePresence perspective; where additional components such as Session Border Controllers (SBCs) are outside the administrative control of the corporate entity.

Note that only the TelePresence primary codecs perform DTLS-SRTP handshakes. Since all voice and video media originate from the primary codec during a TelePresence meeting, SRTP keying material needs to be negotiated only between primary codecs, and not between their associated IP 7975G phones. Since there are both voice and video media streams, two DTLS-SRTP handshakes occur—one over each media stream—and keying material is negotiated for encryption and authentication of both the voice and the video media. This is not shown in [Figure 10](#) above for simplicity. As of CTS Version 1.5, AES with a 128-bit encryption key, is currently negotiated within DTLS-SRTP for encryption of RTP payloads. Likewise, as of CTS Version 1.5, HMAC-SHA1 with an authentication tag size of 80 bits (10 bytes) is negotiated within DTLS-SRTP for packet authentication.

Key Exchange via Encrypted Key Transport (EKT)

EKT is an extension to SRTP which provides for the transport of SRTP master keys securely within SRTP and SRTCP packets. It is currently an IETF Internet-Draft (<http://www.ietf.org/internet-drafts/draft-mcgrew-SRTP-ekt-04.txt>). EKT accomplishes this by sub-dividing the SRTP and/or SRTCP Authentication Tag (shown in Figure 8 and Figure 9 above) into one of the two formats shown in Figure 11.

Figure 11 Encrypted Key Transport (EKT) Packet Formats



In the first (long) format, the final bit of the Authentication Tag is set to 1. This indicates the presence of the following fields:

- *Base Authentication Tag*—Configurable length field used to hold the message authentication data for the RTP or RTCP header and payload for the particular packet.
- *Encrypted Master Key*—Variable length field which contains the encrypted SRTP master key corresponding to the SSRC within the SRTP or SRTCP packet.
- *Rollover Counter*—32-bit field used to hold the value of the SRTP rollover counter associated with the SSRC contained within the SRTP or SRTCP packet. Since the RTP sequence number is only a 16-bit long field, the rollover counter is necessary for codecs to maintain synchronization and minimize re-keying in long term media streams.
- *Initial Sequence Number*—Indicates the RTP sequence number of the first RTP packet which will be protected by the SRTP master key contained within the Encrypted Master Key field of this SRTP or SRTCP packet.
- *Security Parameter Index*—16-bit field similar to the IPSec SPI. It is used to identify a particular security context (Key Encrypting Key (KEK) used to produce the Encrypted Master Key, KEK cipher, SRTP cipher, SRTP master salt, etc.) corresponding to a particular SRTP flow.

In the second (short) format, the final bit of the Authentication Tag is set to 0. This indicates the presence of the following fields:

- *Base Authentication Tag*—Configurable length field used to hold the message authentication data for the RTP or RTCP header and payload for the particular packet.

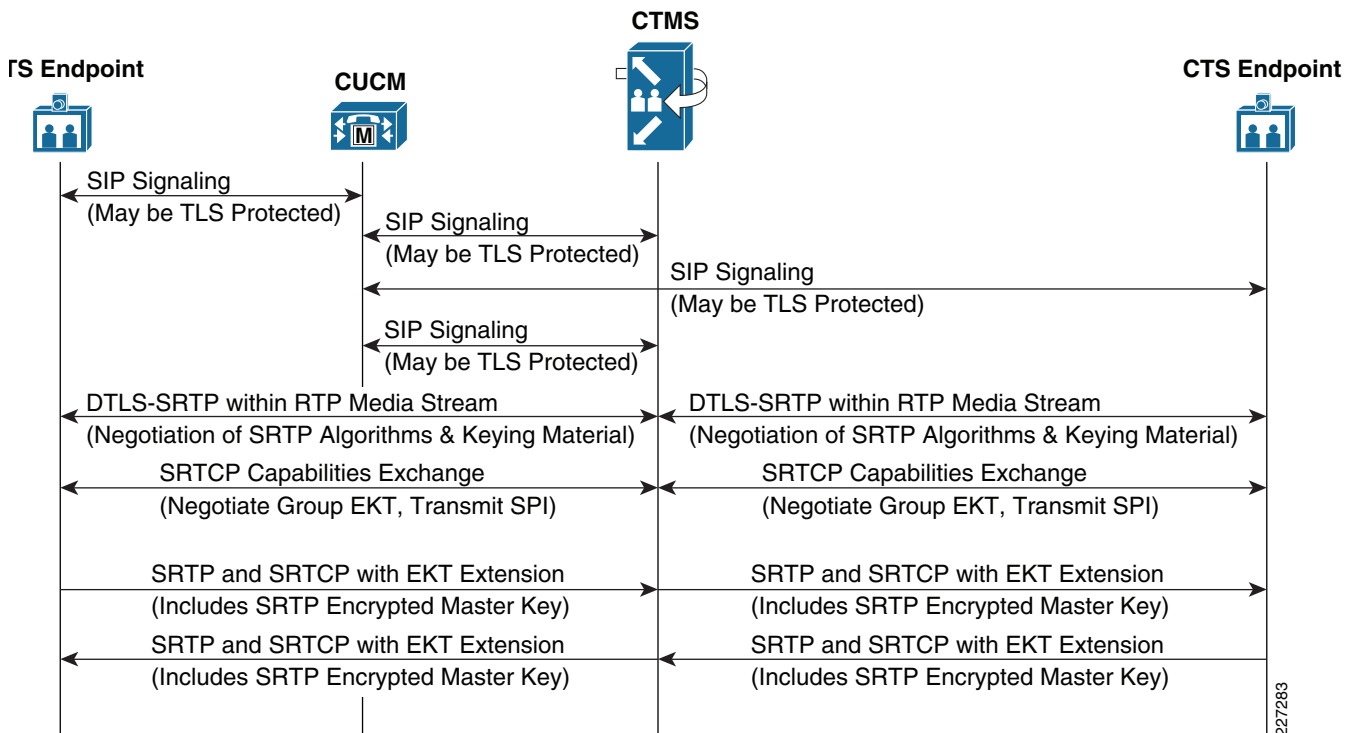
- *Reserved*—7-bit field reserved for future use.

With EKT, the SRTP master key is distributed directly between CTS endpoints (CTS-3200s, CTS-3000s, CTS-1300s, CTS-1000s, and CTS-500s) in a multipoint call by encrypting it with a Key Encrypting Key (KEK), and sending it within SRTP and SRTCP packets which contain the long format of the EKT extension shown in Figure 11 above. This is done for each of the voice and video media streams. The following section provides a high-level overview of the key exchange process within a multipoint TelePresence call using DTLS-SRTP and EKT.

TelePresence Multipoint Call Operation

Figure 12 shows the process that occurs for establishing a secure multipoint TelePresence meeting. Figure 12 shows two CTS-1000 systems in a multipoint meeting via a CTMS. As with previous figures, only one set signaling and media has been shown for simplicity.

Figure 12 Multipoint TelePresence Key Exchange



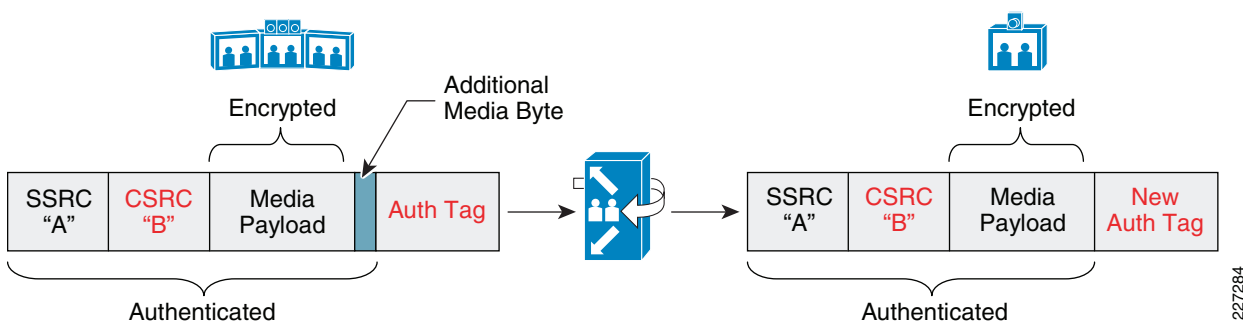
SIP signaling occurs first between the CTS endpoints and the CUCM, as well as between the CTMS and the CUCM. SIP establishes the RTP media streams and the RTCP control streams between the CTS endpoints and the CTMS. Immediately after the RTP media streams are set up, a DTLS-SRTP session is established over each media stream between the CTS endpoints and the CTMS. This is used to negotiate the SRTP encryption and authentication algorithms, as well as keying material. Following this, the CTS endpoints and the CTMS exchange SRTCP packets (using the SRTP keys established within DTLS) in order to discover their capabilities. During this step the CTS endpoints discover that they are communicating with a CTMS, and that EKT is required for SRTP master key exchange between the CTS endpoints. EKT is necessary for SRTP master key exchange because the CTMS does not actually de-encrypt the SRTP media packets and re-encrypt them. Therefore, each CTS endpoint needs the SRTP master keys used to encrypt the media flows from each of the other CTS endpoints in order to decrypt

the media. Also during this step, a group EKT parameter set is established. The group EKT parameter set includes the Key Encrypting Key (KEK), which will be used to encrypt the SRTP master keys sent within SRTCP and SRTP packets which include the long format of the EKT extension shown in [Figure 11](#) above.

SRTP encrypted master keys are sent by the individual CTS endpoints to the CTMS via SRTCP Source Description (SDS) packets which contain the long format of the EKT extension. The CTMS will forward these packets to other CTS endpoints within the multipoint call. Additionally, for video streams, the first SRTP packet of each video frame contains the EKT extension with the SRTP encrypted master key. All other packets of the frame contain the shorter, second format of the EKT extension shown in [Figure 11](#) above. For voice, every third RTP audio packet contains the EKT extension with the SRTP encrypted master key. The other audio RTP packets contain the shorter EKT extension. By forwarding these voice and audio SRTP EKT packets to the other CTS endpoints, the CTMS further guarantees that each endpoint acquires the SRTP encryption key for each transmitter.

[Figure 13](#) shows a high-level example of the video media switching performed by the CTMS during a secure multipoint TelePresence meeting. The SRTP packet has been intentionally simplified for this example.

Figure 13 Media Switching in a Secure Multipoint TelePresence Meeting



With CTS Version 1.5, the TelePresence endpoints use the Contributing Source Identifier (CSRC) field to indicate the source camera or microphone position of the media stream, as well as the destination display or speaker of the media stream, as opposed to the Synchronization Source Identifier (SSRC) which was previously used. The CTMS may modify the CSRC value in cases such as when the video from CTS-3000 *right* camera needs to be displayed on the *right* display of a CTS-3000 and on the *center* display of a CTS-1000 within the same multipoint meeting. However, the SSRC value of the media packets is not modified, as they are switched through the CTMS. CTS endpoints also append an additional unencrypted byte to the payload of audio and video media packets before sending the packets to the CTMS. For audio packets, the additional byte is used to determine the level of audio energy within the packet. This is used by the CTMS to determine which CTS endpoints should be transmitting video streams. For video packets, the additional byte is used to determine if the video packet is the beginning of a new reference frame (IDR). For both the audio and video packets, the CTMS strips off the additional byte before forwarding it to the CTS endpoints. Therefore, the CTMS does need to recompute the authentication tag as it switches the media, as shown in [Figure 13](#) above. However, the CTMS does not decrypt and re-encrypt the actual media streams. This helps to ensure high performance and low delay of the CTMS even with encryption enabled on the multipoint call, as well as preserving the existing hardware investment of the CTMS.

Network Design Implications

This section discusses some of the network design implications for enabling encryption within a Cisco TelePresence deployment. These include the following:

- Possible additional bandwidth utilization considerations
- Possible additional jitter and/or delay incurred when enabling encryption
- Support for encryption across multipoint meetings, as well as meetings which include Interoperability
- Potential issues with regard to visibility and manageability of the TelePresence deployment

Bandwidth Considerations

The current Cisco TelePresence bandwidth requirements can be found in the Quality of Service Design for TelePresence section of the *Cisco TelePresence Network Systems 2.0 Design Guide* at the following URL:

<http://www.cisco.com/en/US/docs/solutions/Enterprise/Video/TP-Book.html>

Since the encryption of the RTP payload yields no payload expansion and the MKI field is not used with Cisco TelePresence, the only expansion to RTP packets when using SRTP with EKT is the addition of the modified Authentication Tag as specified by EKT. The effects on the audio and video media streams are discussed in the following subsections.

Video Media Streams

For video packets, the first packet of each video frame contains the long form of the EKT extension. This adds approximately 34 bytes of overhead (10 byte Authentication Tag + 16 byte Encrypted Master Key + 4 byte Rollover Counter + 2 byte Initial Sequence Number + 2 byte Security Parameter index and last bit set to 1). All other video packets contain the short form of the EKT extension. This adds approximately 11 bytes of overhead (10 byte Authentication Tag + 1 additional byte as shown in [Figure 11](#) above).

Since the average unencrypted Cisco TelePresence video RTP packet is approximately 1,100 bytes in size, the addition of the long form of the EKT extension represents about a 3 percent increase in overhead. The addition of the short form of the EKT extension represents about a 1 percent increase in overhead. Given an average number of approximately 12 packets per video frame (often observed in normal Cisco TelePresence operation), the estimated impact of implementing secure communications on Cisco TelePresence video is slightly over 1 percent increase in video traffic levels over non-secure communications.

Audio Media Streams

For audio packets, every third packet contains the long form of the EKT extension. Again, this adds approximately 34 bytes of overhead (10 byte Authentication Tag + 16 byte Encrypted Master Key + 4 byte Rollover Counter + 2 byte Initial Sequence Number + 2 byte Security Parameter index and last bit set to 1). The other two audio packets contain the short form of the EKT extension. This adds approximately 11 bytes of overhead (10 byte Authentication Tag + 1 additional reserved byte with the last bit set to 0).

Since the average Cisco unencrypted Cisco TelePresence audio RTP packet is approximately 225 bytes in size, the addition of the long form of the EKT extension represents about a 15 percent increase in overhead. The addition of the short form of the EKT extension represents about a 5 percent increase in overhead. Given the ratio of long to short EKT extension headers within audio packets observed the estimated impact of implementing secure communications on Cisco TelePresence audio is slightly over 8 percent increase in audio traffic levels over non-secure communications.

RTCP Streams

Based on observations, very few SRTCP packets contain the long form of the EKT extension. However, SRTCP packets with the long EKT extension contain approximately 38 bytes of overhead (4 bytes for the encryption bit and SRTCP Index + 10 byte Authentication Tag + 16 byte Encrypted Master Key + 4 byte Rollover Counter + 2 byte Initial Sequence Number + 2 byte Security Parameter index and last bit set to 1) over normal RTCP packets. The majority of SRTCP packets contain the short form of the EKT extension which contains approximately 15 bytes of overhead (4 bytes for the encryption bit and SRTCP Index + 10 byte Authentication Tag + 1 additional Reserved byte with the last bit set to 0) over normal RTCP packets.

Based on observations, the average size of RTCP packets within Cisco TelePresence is around 100 bytes. Therefore, the estimated impact of implementing secure communications on Cisco TelePresence RTCP is approximately a 15 percent increase in RTCP traffic levels over non-secure communications.

Overall Estimated Bandwidth Impact

The design engineer should note that the amount of video traffic greatly exceeds the amount of voice traffic in terms of bandwidth usage during a TelePresence meeting. Each TelePresence video stream can go up to 4.4 Mbps (4 Mbps + 10 percent temporary burst = 4.4 Mbps) without network layer header overhead. Each audio stream only uses 64 Kbps without network layer overhead. Likewise, the amount of RTCP traffic is effectively negligible compared to the video component of a Cisco TelePresence meeting.

The use of SIP over TLS does result in some increase in packet size and number of packets (TLS negotiation) sent for SIP signaling between the TelePresence endpoints and the CUCM server. However, the amount of SIP signaling messages sent between devices is also considered negligible from a bandwidth perspective compared to the amount of video traffic. Likewise, the use of TLS for web services messaging does result in some increase in packet size and number of packets (TLS negotiation) sent for web services signaling between the CTMS, CTS endpoints, CTS-MAN, and CUCM. Again, the amount of web services signaling messages sent between the devices is considered negligible from a bandwidth perspective compared to the amount of video traffic.

Therefore, the impact of enabling secure communications on the bandwidth utilization estimates for a TelePresence meeting is estimated to be slightly over a 1 percent increase in overall traffic. However, since the rough estimate of approximately 20 percent network protocol overhead for Cisco TelePresence meetings is a “rough estimate” in the first place, it is considered to hold for both secure and non-secure TelePresence deployments still. Therefore, existing bandwidth design guidance still holds for both secure and non-secure Cisco TelePresence deployments.

Jitter and Delay

Based on observations of TelePresence endpoints themselves via the web-based administrative interface, enabling encryption on the TelePresence endpoints seems to have negligible impact on jitter and delay for both point-to-point and multipoint TelePresence meetings. This was expected since TelePresence endpoints utilize dedicated hardware codecs which provide sufficient processing capability for the encryption and decryption process without affecting performance. Likewise, since the CTMS does not decrypt and re-encrypt the media itself, it has sufficient processing capability for calculating new authentication tags and switching the media packets, without significantly affecting performance.

Encryption of Point-to-Point, Multipoint, and Interoperability Meetings

Encryption of point-to-point TelePresence meetings has been supported as of CTS Version 1.2. As of CTMS Version 1.5, encryption of multipoint TelePresence meetings is also supported. Encryption of TelePresence meetings that require interoperability is currently not supported. The video/audio media streams required for TelePresence interoperability are cascaded from the CTMS to the CUVIC 3500 Series MCU. Since the CUVIC 3500 Series MCU does not support DTLS-SRTP or EKT; encryption is not supported. Therefore, TelePresence meetings that require interoperability can only be configured as non-secured meetings.

When a point-to-point TelePresence call is attempted between a CTS device that supports encryption and a CTS device which does not support encryption—such as a TelePresence endpoint with no security profile applied—the TelePresence meeting will automatically fall back to an unsecured (not authenticated or encrypted) meeting. The end user is made aware of the security status of the meeting through one of the following icons which appear briefly at the beginning of a TelePresence meeting. See [Figure 14](#).

Figure 14 *TelePresence Meeting Security Indicators*



With multipoint TelePresence meetings, the default meeting behavior can be determined via a meeting security policy configured by the administrator through the **Security Settings** page of the CTMS, as shown in [Figure 15](#).

Figure 15 CTMS Security Settings Page

The screenshot shows the Cisco TelePresence Multipoint Switch Administration interface. The left sidebar contains a navigation menu with options like System Information, System Configuration, System Settings, Import/Export Files, Unified CM, CTM Settings, Access Management, Software Upgrade, Security Settings, Interface Failover, Service Settings, Meetings Management, Default Settings, Static Meetings, Adhoc Meeting, Scheduled Meetings, and Active Meetings. The main content area is titled 'System Configuration > Security Settings'. It features a 'Meeting Security Policy' section with three radio buttons: 'Non-Secured', 'Secured', and 'Best-Effort' (which is selected). Below this is a 'Web Services Security' section with a 'Secured' radio button. There are 'Apply' and 'Reset' buttons. Below these sections is a 'Digital Security Certificates' section with a 'Category' dropdown set to 'All' and a 'Unit' dropdown set to 'All'. A 'Filter' button is present. Below the filters is a table showing 6 records of certificates. The table has columns for 'Unit', 'Category', and 'Certificate Name'. The records are: tomcat (OWN, tomcat.pem), CTM-trust (TRUST, CUCM2.pem), CTM-trust (TRUST, CUCM0.pem), CAPF-LSC (OWN, CTMS_Cert_Chain.pem), CAPF-LSC (OWN, CTMS.pem), and CAPF-trust (TRUST, CAPF.pem). At the bottom of the table are buttons for 'Upload', 'Download LSC', 'View', and 'Delete All'.

Unit	Category	Certificate Name
☐ tomcat	OWN	tomcat.pem
☐ CTM-trust	TRUST	CUCM2.pem
☐ CTM-trust	TRUST	CUCM0.pem
☐ CAPF-LSC	OWN	CTMS_Cert_Chain.pem
☐ CAPF-LSC	OWN	CTMS.pem
☐ CAPF-trust	TRUST	CAPF.pem

The meeting security policy can be one of the following three choices:

- *Non-Secured*—With this setting, the meeting utilizes no security, regardless of whether all CTS devices support TelePresence security (configured with a security profile). All CTS endpoints are allowed to join the TelePresence meeting with no authentication or encryption enabled.
- *Secured*—With this setting, all meeting participants are required to support TelePresence security in order to join the TelePresence meeting. CTS endpoints that are not configured to support security (no security profile configured) are not allowed to join the multipoint meeting.
- *Best-Effort*—With this setting TelePresence security will be enabled if all CTS endpoints which join the multipoint TelePresence meeting are configured to support security (configured with a security profile). However, if an endpoint which does not support TelePresence security joins the multipoint meeting, all endpoints will fall back to an unsecured (no authentication or encryption) state.

Note also that the status of the web services signaling to and from the CTMS (referred to as Web Services Security) can also be verified from this screen. The default meeting security policy for an individual meeting can also be overridden by the administrator or meeting scheduler when scheduling the meeting. [Figure 16](#) shows an example of this for a static meeting.

Figure 16 Static Meeting Configuration Specifying Security Policy

Meetings Management > Static Meetings

Access Number:	9193926001
Meeting ID:	9193926001
Meeting Description:	Best Effort *
Switching Policy:	☒ Room ☐ Speaker
Maximum Rooms:	16 *
Video Announce:	☒ Yes ☐ No
Hosted Meeting:	☐ Yes ☒ No
Host Room Number:	*
Quality:	Highest Detail, Best Motion: 1080p
Interop:	☐ Yes ☒ No
Meeting Security Policy:	☐ Non-Secured ☐ Secured ☒ Best-Effort

Save Close

The status of the Web Services Security between the CTS-MAN and the CTMS can also be verified by the administrator through the **Multipoint Conference Unit** screen of the CTS-MAN. An example is shown in Figure 17.

Figure 17 CTS-MAN Multipoint Conference Unit Screen

Cisco TelePresence Manager

admin | Logout | Preferences | Help | About

Host: tp-c1-ctm...

Support > Multipoint Conference Unit

Summary Capability

MCU Devices

Status: All MCU: Filter

Showing 1 - 2 of 2 records

Status	Hostname	Type	Version	Switching	Conference Termination	Interop	Web Services Security
OK	tp-ctms-1.tp.com	CTMS	1.5.1.0 (2)	✓	✓	✓	🔒
OK	tp-c1-vc3500-1.tp.com	CUVC		✗	✗	✗	🔓

First < Previous Next > Last Rows Per Page: 10

System Status

Today's Meetings:

With Error:	0
In Progress:	0
Scheduled:	0
Other Errors:	3

Finally, as with point-to-point calls, the end user is made aware of the security status of the multipoint meeting through one of the icons shown in Figure 14, which appear briefly at the beginning of the TelePresence meeting.

Network Visibility

When deploying encryption within Cisco TelePresence deployments, the network administrator should be aware of potential issues regarding visibility into the various protocols which many network devices require in order to operate. These visibility issues can be separated into two categories—call signaling visibility and media visibility. Each are discussed below.

Call Signaling Visibility

Cisco TelePresence is sometimes deployed in an environment where firewalls exist between TelePresence endpoints. Firewalls often use application layer inspection of the SIP signaling in order to open the necessary dynamic RTP port ranges required for the audio and video media streams, as well as enforce security policy. When TLS encryption is enabled for SIP signaling, such firewalls will not be able to inspect the SIP/SDP messages in order to dynamically open the necessary dynamic ports. In these scenarios, the network administrator may be left with one of two alternatives—statically open the range of IP addresses and UDP ports that correspond to the RTP audio and video media streams that must cross the firewall; or implement a firewall that uses TLS proxy functionality for Cisco TelePresence endpoints. With TLS proxy functionality, the TLS session is established between the TelePresence endpoint and the firewall. A second TLS session is established between the firewall and the CUCM server. The firewall is then able to decrypt the TLS packets, inspect the SIP/SDP messages to determine the necessary dynamic UDP ports to open for audio and video media. The signaling is then re-encrypted and sent to its respective endpoint. A third alternative is to implement IPSec encryption on network devices between the two TelePresence endpoints, and statically allow IPSec traffic to pass through the firewall.

**Note**

Testing of the TLS proxy functionality of the ASA 5500 Series Firewall with Cisco TelePresence is currently being looked into as of the time this document was written, and is currently not a supported feature. This document will be updated with results when available.

Media Visibility

Since SRTP only encrypts the payload of the RTP packet, and not the RTP header; much of the visibility into the RTP audio and video media streams remains with the deployment of encryption for Cisco TelePresence. As can be seen in Figure 8 above, the RTP Payload Type (PT) field, as well as the SSRC and any CSRCs are still visible to network level devices. Therefore any deep-packet inspection (DPI) implemented within the network, such as the hardware-accelerated NBAR capability of the Cisco Catalyst 6500 Sup-32 PISA, can still identify the RTP payload types for TelePresence traffic and apply appropriate policy to the traffic if so configured. Likewise any static access control lists which utilize RTP port numbers or NBAR capabilities to characterize traffic based on RTP payload types, will be able to recognize encrypted TelePresence audio and video media. Note however that protocol analyzers deployed within the network will not be able to decode into the Network Abstraction Layer (NAL) of the TelePresence H.264 video streams when encryption is deployed. However, this is not considered to be a normal requirement of network administration.

Secure RTP versus IPsec Encryption

IPsec encryption of Cisco TelePresence traffic is briefly discussed in Chapter 6 – Branch QoS Design for TelePresence of the *Cisco TelePresence Network Systems 2.0 Design Guide* found at the following URL:

<http://www.cisco.com/en/US/docs/solutions/Enterprise/Video/TP-Book.html>

Table 1 provides a high level comparison of the use of SRTP / TLS versus IPsec encryption for Cisco TelePresence meetings.

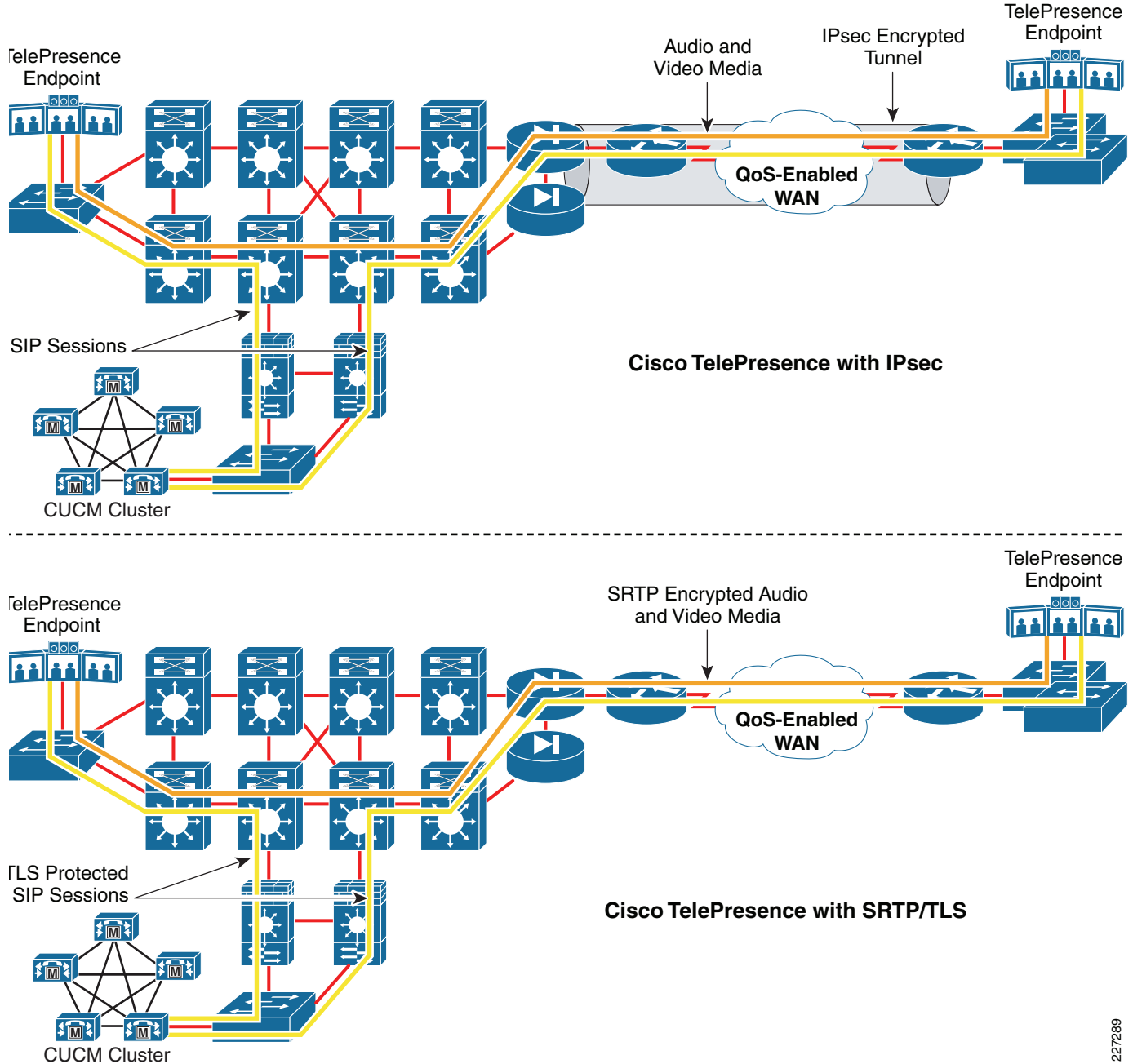
Table 1 **Comparison of SRTP/TLS versus IPsec**

	SRTP / TLS	IPsec
Media Encryption Considerations	Encryption and authentication of the RTP media is deployed end-to-end, between the primary codecs of the TelePresence endpoints.	Encryption and authentication is deployed between network devices to which the TelePresence endpoints are attached. Further, in typical deployments, IPsec encryption is only deployed across corporate WAN circuits between routers; with the corporate LAN connectivity considered somewhat secure although the TelePresence media and signaling may not be encrypted or authenticated.
Signaling Encryption Considerations	Encryption and authentication of the SIP signaling is deployed end-to-end, between the primary codec of the TelePresence endpoint and the CUCM server. Additionally, encryption and authentication of the web services signaling is deployed between the CTMS, CTS endpoints, CUCM server, and CTS-MAN.	Encryption and authentication is deployed between network devices to which the TelePresence endpoint and the CUCM are attached. Further, in typical deployments, IPsec encryption is only deployed across corporate WAN circuits between routers; with the corporate LAN connectivity considered somewhat secure, although the TelePresence SIP signaling is not encrypted or authenticated.
Media Bandwidth Considerations	The AES encryption algorithm yields no expansion of the RTP payload itself. The expansion due to SRTP / EKT is estimated to be about 8 percent on average for audio packets, and slightly over 1 percent for video packets. Overall, the additional bandwidth increase is estimated to be slightly over 1 percent based on the proportion of audio to video traffic within a Cisco TelePresence call.	IPsec adds up to an additional 56 bytes to all voice and video RTP packets; resulting in up to approximately 25 percent increase in voice packet size and approximately 5 percent increase in average video packet size. Overall, the additional bandwidth increase is estimated to be slightly over 5 percent based on the proportion of audio to video traffic within a Cisco TelePresence call. The use of GRE tunneling within IPsec adds additional overhead.

Table 1 **Comparison of SRTP/TLS versus IPSec**

Call Signaling Visibility	The network has visibility that the call signaling is SIP. Therefore, QoS and other policy can be applied based on the traffic type, as well as DSCP marking. Visibility into SIP/SDP messages in order to dynamically open firewall ports for RTP media requires a TLS proxy functionality to be deployed within the firewall. Otherwise, static ranges of ports for the TelePresence RTP media can be opened.	The network has no visibility into the traffic, other than it is IPSec encapsulated traffic. Note however, that this may be desirable from a security perspective over WAN links. QoS can be applied before the traffic is encapsulated within IPSec in order to apply policy based on DSCP marking. However, native IPSec tunnel mode as well as IPSec protected GRE tunnels limit visibility to the actual traffic. QoS pre-classify can be used to look at the DSCP settings prior to encryption and then setup a pointer to the packet. After encryption, the pointer is used to determine interface queuing. Unless the firewall terminates the IPSec connection, it must be statically configured to pass IPSec encapsulated traffic. In this situation, the firewall has no visibility into the traffic it is passing.
Media Visibility Considerations	The network has complete visibility into the RTP media up to the NAL layer for H.264 video traffic. Deep packet inspection engines such as NBAR/PISA are able to identify and apply policy to the traffic based on RTP payload type.	The network has no visibility into the traffic, other than it is IPSec encapsulated traffic. Again, this may be desirable from a security perspective over WAN links. Deep packet inspection engines such as NBAR/PISA will not be able to identify and apply policy based on RTP payload type.
Point-to-Point versus Multipoint Considerations	As of CTS Version 1.5 SRTP/ TLS is supported for both point-to-point and multipoint TelePresence meetings.	IPSec can be deployed on network devices in order to support both point-to-point and multipoint TelePresence meetings.

The primary benefits of the use of TelePresence encryption via SRTP/TLS is end-to-end security with minimal additional bandwidth overhead due to packet expansion. IPSec provides an alternative, which can secure both point-to-point as well as multipoint meetings. IPSec encryption is not supported on the TelePresence endpoints themselves. Instead, IPSec is deployed on networking equipment between the TelePresence endpoints. Typical deployments implement IPSec across WAN circuits only, leaving the TelePresence media and call signaling unencrypted or authenticated across the corporate LAN, which is often considered secure from a corporate standpoint. [Figure 18](#) visually highlights TelePresence over IPSec versus SRTP/TLS in this type of deployment.

Figure 18 *Example TelePresence Deployment over IPsec versus SRTP/TLS*

With the addition of encrypted multipoint calls and secure web services signaling between TelePresence devices; CTS Version 1.5 provides a complete enterprise solution for providing secure communications and signaling for TelePresence meetings. Therefore, it is recommended to use SRTP/TLS instead of IPsec where possible within Cisco TelePresence enterprise deployments in order to provide secure communications and signaling for both point-to-point and multipoint TelePresence meetings.

227289

Summary

This document presents the robust authentication and encryption framework used to secure the audio/video media, call signaling, and web services signaling of the Cisco TelePresence solution. The framework consists of the use of Transport Layer Security (TLS) in order to provide authentication and encryption of SIP signaling between the CUCM and CTS endpoints; as well as authentication and encryption of the XML signaling between the CUCM, CTMS, CTS Manager and CTS endpoints. Audio and video media can be encrypted through the use of SRTP and SRTCP, which provides little packet expansion, resulting in minimal increase in bandwidth utilization when secure Telepresence meetings are required. As of CTS Version 1.5, the use of SRTP/TLS to secure TelePresence meetings applies to both point-to-point and multipoint meetings; and is the recommended approach to providing secure communications and signaling for Cisco TelePresence deployments.

Glossary

- *CTMS*—The Cisco TelePresence Multipoint Switch (CTMS) provides multipoint meeting services for Cisco TelePresence deployments.
- *CTS Unit*—Any Cisco TelePresence System device, including the CTS-3500, CTS-3000, CTS-1300, or CTS-500.
- *CTS-MAN*—The Cisco TelePresence System Manager (CTS-MAN) provides management and meeting scheduling services for Cisco TelePresence deployments.
- *CUCM*—The Cisco Unified Communications Manager (CUCM) provides SIP call signaling services for Cisco TelePresence deployments.
- *TelePresence Endpoint*—Generic term for any Cisco TelePresence device, including CTS Units the CTMS or the CTS-MAN.

Reference Documents

- *Cisco Unified Communications Manager Security Guide, Release 7.1(2)*
http://www.cisco.com/en/US/partner/docs/voice_ip_comm/cucm/security/7_1_2/secugd/sec712-cm.html
- *Cisco TelePresence Network Systems 2.0 Design Guide*
<http://www.cisco.com/en/US/docs/solutions/Enterprise/Video/TP-Book.html>
- *Cisco TelePresence Security Solutions*
http://www.cisco.com/en/US/docs/telepresence/security_solutions/CTSS.pdf
- *Cisco TelePresence Fundamentals*
Tim Szigeti, Kevin McMenamy, Roland Saville, and Alan Glowacki
Indianapolis, IN: Cisco Press, 2009.