AN ARCHITECTURE FOR OBTAINING VOIP SESSION ENCRYPTION KEYS IN A CALEA COMPLIANT NETWORK

A Dissertation by

Stephen D. Guhl

M. S. Electrical Engineering, Wichita State University, 1993

Submitted to the Department of Electrical and Computer Engineering
And the faculty of the Graduate School of
Wichita State University
in partial fulfillment of
the requirements for the degree of
Doctor of Philosophy

May 2008

AN ARCHITECTURE FOR OBTAINING VOIP SESSION ENCRYPTION KEYS IN A CALEA COMPLIANT NETWORK

The following faculty members have examined the final copy of this dissertation for forma and content, and recommend that it be accepted in partial fulfillment of the requirement for the degree of Doctor of Philosophy.

Ravindra Pendse, Committee Chair

We have read this dissertation
and recommend its acceptance:

M. Edwin Sawan, Committee Member

Kamesh Namuduri, Committee Member

Krishna Krishnan, Committee Member

Kurt Soschinske, Committee Member

Accepted for the College of Engineering

Zulma Toro-Ramos, Dean

Accepted for the Graduate School

Susan Kovar, Dean

DEDICATION

I would like to dedicate this dissertation to my wife and family whose loyal support and encouragement were invaluable

ACKNOWLEDGEMENTS

I would like to acknowledge Dr. Ravi Pendse's efforts and encouragement. Without such, I would not have succeeded.

ABSTRACT

A number of laws have been passed in recent decades governing the wiretapping and interception of conversations on the PSTN and recently the internet. The Communications Assistance for Law Enforcement Act (CALEA) passed in 1994 requires that digitally switched telephone networks be designed and built with wiretap capabilities and that service providers assist Law Enforcement Agencies (LEA) in obtaining the desired surveillance.

The FCC has ruled that the CALEA also applies to Voice over IP (VoIP) conversations. This has generated considerable contention amongst the internet community regarding the extent to which wiretapping should be embedded into the applicable internet protocols. A number of industry opinions have been expressed that providing wiretap capability will reduce the security of the internet and compromise the basic internet design paradigm. That paradigm expresses the niew of placing the complexity in the applications implemented by end users and simplicity in the routing of information between them.

The IETF has provided general guidelines regarding architecture for the interception and availability of the information to the LEAs but has declined to produce a full fledged standard. Internet component manufacturers have driven the effort to meet the requirements of CALEA. However in an effort to meet the understood surveillance requirements none of the existing architectures adequately address the issue of end user encryption both in the signaling and content messages. Law enforcement agencies must rely on access to available decryption tools which may delay the timely response to threatening situations in the current war on terror. A key management system which allays the concerns with internet security and complexity is addressed in this investigation.

TABLE OF CONTENTS

Chapter Page

TABLE OF CONTENTS (continued)

LIST OF TABLES

Table Page

LIST OF FIGURES

Figure Page

LIST OF ABBREVIATIONS

AAA	Authentication Authorization and Accounting
ACF	Admission Confirm
AES	Advanced Encryption Algorithm
AOR	Address of Record
ARQ	Acknowledgement Request
B2BUA	Back-to-Back User Agent
BTS	Broadband Telephony Softswitch
CALEA	Communications Assistance for Law Enforcement Act
CCIAP	Call Content Intercept Access Point
CDT	Center for Democracy and Technology
CI	Critical Infrastructure
CMTS	Cable Modem Termination System
CTR	Counter
FCC	Federal Communications Commission
FISA	Foreign Intelligence Surveillance Act
HMAC	Hashed Message Authentication Code
IAP	Intercept Access Point
IETF	Internet Expert Task Force
ILD	Intermediate-Level Device
IPsec	Internet Protocol Security
IRI	Intercept Related Information
ISAKMP	IPsec Key Management Module

LIST OF ABBREVIATIONS (continued)

ISP	Internet Service Provider
IV	Initialization Vector
LEA	Law Enforcement Agency(s)
LI	Lawful Interception
MCU	Multipoint Control Unit
MD	Mediation Device
MD5	Message Digest 5
MIKEY	Multimedia Internet Keying
MKI	Master Key Index
MTA	Media Terminal Adapter
P2P	Peer-to-Peer
PSTN	Public Telephone Switch Network
Qop	Quality of Protection
QoS	Quality of Service
RAS	Registration, Admission, and Status
RFC	Request for Comments
ROC	Rollover Counter
RTP	Real-Time Transport Protocol
SBC	Session Border Controller
SD	Storage Device
SDP	Session Description Protocol
SEQ	Sequence Number

LIST OF ABBREVIATIONS (continued)

SHA	Secure Hash Algorithm
SIP	Session Initiation Protocol
S/MIME	Secure/Multipurpose Internet Mail Extensions
SP	Service Provider
SRTP	Secure Real-Time Transport Protocol
SSRC(I)	Synchronization Source (Identifier)
TCP	Transport Control Protocol
TLD	Top-Level Device
TLS	Transport Layer Security
TEK	Traffic Encryption Key
UA	User Agent
UDP	User Datagram Protocol
URI	Uniform Resource Identifier
VoIP	Voice over IP
WIEP	Wiretap Information Exchange Protocol
XML	Extensible Markup Language
XOR	Exclusive ORing

CHAPTER 1

Introduction

Although wiretapping has existed since the 1920's the rapid advancement of digital technologies into the PSTN during the 1990's and the multiplicity of service providers forced Congress into action passing the Communications Assistance for Law Enforcement Act (CALEA) to preserve law enforcement's investigative capabilities. In 2005 the FCC ruled that CALEA also applies to VoIP conversations carried over the internet. Efforts have been made by various component providers to meet the requirements of CALEA as understood by the internet industry.

Two categories of intercepted information are defined as related to Lawful interception. Intercept related information (IRI), this information includes the source and destination phone numbers and IP addresses. This information would be carried in the signaling protocol such as H.323 or Session Initiation Protocol (SIP) and its related protocols. The second category is the actual content of the conversation which would be carried in a Real Time Protocol (RTP) packet. Each of these categories of information can be encrypted using different protocols. The current guidelines referenced in IETF RFCs require only that if the service provider supplied the encryption keys or has access to them they may be forwarded to the Law Enforcement Agency (LEA). With no standard approach to key access the LEA is left to what ever means it has available to decrypt the information. With timely decryption becoming an important part of threat detection and prevention a standard approach to key management needs to be available which will assist the LEA in its function while not degrading internet security and protecting individual privacy.

1.1 Legal Requirements of Lawful Interception

There are a number of specified requirements for Lawful Interception (LI). Some of the requirements for Lawful Interception are listed as follows:

1. Lawful Intercept (LI) should be undetectable by the intercept subject.
2. Unauthorized personnel should be limited from performing or knowing about lawful authorized intercepts.
3. Intercepted encrypted information that is encrypted by the service provider and the provider has access to the keys should be decrypted before delivery to the Law Enforcement Agency (LEA) or the keys passed to the LEA.
4. If the information is encrypted by the subject and the service provider has access to the keys, the service provider may deliver the keys to the LEA.
5. Multiple simultaneous intercepts done on a single subject at the request of different LEAs must be transparent or undetectable by the other requesting LEAs.
6. No unauthorized information should be delivered to a LEA.

All of these requirements place burdens upon the authentication, message integrity, and privacy of the LI architecture. Item (1) implies that if the subject chooses to encrypt that the encryption scheme not be detectable between a subject that is under surveillance verses one that is not. Items (3) and (4) suggest that it would be very helpful to law enforcement agencies for the service provider to provide encryption to their clients who want to use that feature. If the clients use the encryption feature then they must have confidence that no unauthorized personnel or entities (obviously excluding law enforcement agencies) have access to the encryption scheme employed. Lastly (2), (5), and (6) suggest an integrated architecture that that could meet all three requirements so that authentication, message integrity, and privacy is assured amongst all interested parties.

1.2 Lawful Intercept Architecture

The generic Lawful Intercept architecture has been generally defined by equipment manufacturers and consists of 4 major components/functions.

1.2.1 Lawful Intercept (LI) Administration Function:

This function is the provisioning interface for the intercept as required by a court order or warrant delivered by the Law Enforcement Agency. Typically this is a single interface to the Mediation Device (MD), which provisions the other components in the network. The provisioning interface is controlled to limit accessibility to authorized personnel.

1.2.2 Intercept Access Point (IAP):

The IAP device intercepts the information. It may be an existing device or a special device designed for that purpose. There are two types of IAP content and Intercept Related Information (IRI). A content IAP is used to intercept the IP traffic of interest. The IRI IAP provides the intercept related information.

1.2.3 Law Enforcement Agency (LEA):

This is the legal agency that requests the intercept and to which the information is delivered.

1.2.4 Mediation Device (MD)

The MD requests intercepts from IAPs through interfaces. The service provider controls the mediation device which formats the information and replicates the information (if required) for each Law Enforcement Agency. The LEA's access is limited to the Administration Function and the Mediation Device.

The operation is implied by the description of the function. The intercept process is initiated when the service provider receives a court order or warrant from a law enforcement agency with a specific subject identified and the information to be obtained (Intercept related information only or both IRI and content messages) and the start time and duration of the surveillance. A number of manufactures of internet equipment, such as Cisco, have implemented architectures that reflect that described above. The current architectures make no effort to decrypt encrypted information unless the encryption scheme is readily available to the service provider.

1.3 Session Initiation Protocol and Applicable Security Protocols

There are two major signaling protocols in use over the internet to set up and tear down VoIP connections, H.323 and Session Initiation Protocol (SIP). The International Telecommunication Union (ITU) H.323 is the older recommendation originally approved by the ITU in 1996. H.323 is an umbrella recommendation covering multimedia communication over packet networks. There are number of major elements in a H.323 network. Terminals, gateways, and multipoint control units (MCUs) are considered network end-devices. An end-device or end point originates and terminates media streams which could be audio, video, data, or any combination of the three. A gateway interfaces the H.323 network to another protocol network such as the PSTN. An MCU provides conferencing services for the end terminals. A gatekeeper is a server that controls a zone. If a gatekeeper is present, all end points within that zone register with the gatekeeper which controls authorization of the network services, placing and accepting a call. A gatekeeper also provides gateway location, address translation, bandwidth management, and feature implementation services to the terminals. H.323 signaling protocol uses TCP as the transport layer. H.323 has dominated in the videoconferencing application field. However, SIP is fast becoming the dominate internet telephony signaling protocol.

Session Initiation Protocol was developed by the IETF and originally released in 1999 and updated in 2003. Session Initiation Protocol strengths lie in its original design compatibility with internet protocols. SIP is text-based protocol similar to HTTP and SMTP. SIP while a signaling protocol also has both presence and instant message capability. The SIP signaling process is implemented with the exchange of request messages and numbered responses. SIP implementations use User Datagram Protocol (UDP) port 5060 to which proxy servers listen for incoming requests. The destination addresses in a SIP request are a uniform resource identifier (URI). Two types of URIs are used by SIP, an Address of Record (AOR) or a device URI. The AOR URI is associated with a user or a service. A Contact or device URI is related to a particular device or endpoint. SIP has built in reliability mechanisms allowing it to use User Datagram Protocol (UDP) as its transport protocol.

SIP end points are known as user agents (UAs). User Agents typically open a session with a proxy server which is one of a number of servers utilized by SIP systems. The proxy server receives and forwards requests.

Session Description Protocol (SDP) is used to describe media sessions for which SIP is establishing the connection. The information typically contained in a SDP message are; type of session, media codec, IP address and port number for receiving data. SDP messages have no built in security but rely on standard internet security protocol applications to transport SIP information when required. This is one of the distinguishing features of SIP from H.323. SIP relies on the standard internet security protocols such as Transport Layer Security (TLS) or Secure/Multipurpose Internet Mail Extensions (S/MIME). TLS cannot provide end to end protection if proxy servers are present. A separate TLS connection is needed between each server exposing the SDP information. S/MIME can provide complete end to end protection. This then presents a problem for a Lawful Intercept of Intercept Related Information in that available security schemes such as S/MIME limit the points at which IRI can be obtained. The details of TLS and S/MIME and their impact on LI will be discussed in chapter two.

In addition to security measures that can be applied to the signaling information in SIP security schemes can also be applied to the media or information content of the conversations. A VoIP media stream uses Real-Time Transport Protocol which utilizes UDP as the transport. Secure RTP (SRTP) adds authentication, confidentiality, and integrity protection to RTP sessions. SRTP uses the Advanced Encryption Algorithm (AES) to encrypt the data stream. Message authentication can be implemented by using a HMAC SHA-1 hash. HMAC (Hashed Message Authentication Code) is a message digest hash using a shared secret encryption key. SHA-1 (Secure Hash Algorithm) generates a 160 bit message digest. HMAC SHA-1 provides authentication and integrity of the message. SRTP keying and configuration information can be carried in the SDP. Multimedia Internet Keying (MIKEY) is a key exchange protocol for SRTP that can be initiated in SDP. A number of key exchange methods are supported by MIKEY including , preshared key, public key and Diffie-Hellman key generation. MIKEY provides its own encryption and integrity protection. The ability to encrypt the media stream again raises challenges to the ability of Law Enforcement Agencies to decrypt media streams in a timely manner. The impact of media stream encryption on Lawful Intercept will also be discussed in detail in chapter two.

1.4 Goals of the Research

The purpose of this dissertation is to examine the issues generated by the application of CALEA to IP telephony and the current LI architectures.

- Examine the impact of encryption on the SIP signaling messages and media streams
- Propose an architecture to obtain session encryption keys used in VoIP conversations that will not unduly reduce the security of the internet and its users but will provide law enforcement with a more timely ability to obtain and decrypt signaling and media information.
- Model this proposed architecture and to determine its applicability to a CALEA compliant network.

CHAPTER 2

Literature Review

This chapter reviews the literature relevant to the historical background of VoIP surveillance and the security mechanisms applicable to VoIP specifically SIP signaling and media security. Following this, information will be presented regarding lawful intercept architectures. This will provide the background for discussions of a new approach for surveillance protocols and key management meeting the necessary requirements of lawful intercept and efficient decryption by law enforcement agencies.

2.1 Historical Background of Wiretapping

Wiretapping has been a legally evolving concept, initially applied in the 1920's during the time of prohibition [1]. Since that point there has been a legal tight-rope walk between the invasion of privacy and warranted legal searches for information. The 1934 US Federal Communications Act prohibited the "interception and divulgence" of wired communications. The US Supreme Court in a series of rulings over the next 30 years continued to narrow the circumstances under which law enforcement could electronically "bug" individual communications. The result was an underground wiretapping effort which produced significant abuses by the legal system.

In 1967 President Johnson signed the Omnibus Crime Control and Safe Streets Act of 1968. Title III of this act listed only 26 crimes that would warrant wiretap investigations. A judge was required to rule on whether the warrant request met specific conditions before the process could proceed. Additional abuses by the Intelligence community prompted Congress to pass the Foreign Intelligence Surveillance Act (FISA), in 1978, which authorized procedures and safeguards for national security wiretaps.

Congress passed the Communications Assistance for Law Enforcement Act (CALEA) in 1994. It mandated that digitally switched telephone networks must be designed and built wiretap enabled. The US Department of Justice would determine the appropriate technology standards.

As a response to the terrorist attacks on September 11, 2001 Congress passed the Uniting and Strengthening America by Providing Appropriate Tools Required to Intercept and Obstruct Terrorism Act (USA PATRIOT Act) which was signed into law on Oct 26, 2001. This legislation expanded the authority of law enforcement to fight terrorism in the United States and abroad. Title II of the act permits surveillance to use pen register (records call-identifying information for calls originated by the subject), and or trap and trace devices (records call-identifying information for calls received by the subject) to obtain information. Sections 215 and 210 in Title II also allow the collection of communications records pertaining to session times, electronic communication durations and any identifying addresses as well as the equipment used by the surveillance subject [23].

Thus, since 1995 CALEA has required telecommunications carriers to deploy equipment and services capable of supporting law enforcement surveillance activities. The question is who determines what capabilities the law requires? CALEA was enacted because of the new technologies that had emerged in the 90's, the migration of analog to digital communications, presenting enormous technical challenges to law enforcement wire tapping efforts. In addition competitive new entrants into the service provider arena provided many more alternatives to local access for individuals and institutions.

With these issues in mind Congress passed CALEA with the following goals:

- Preserve a narrowly focused capability for law enforcement agencies to carry out properly authorize intercepts

- Protect privacy in the face of increasingly powerful and personally revealing technologies

- Avoid impeding the development of new communications services and technologies.

The government passed CALEA to preserve law enforcement's investigative capabilities. These are defined in Section 103 of CALEA:

- Isolate expeditiously the content of targeted communications transmitted with in the carrier's service area

- Isolate expeditiously information identifying the targeted communications' originating and destination numbers, but not targets' physical locations

- Provide intercepted communications and call-identifying information to law enforcement in a format transmittable over lines or facilities leased by law enforcement to a separate location

- Carry out intercepts unobtrusively, so electronic surveillance targets aren't aware of the interception, and in a way that doesn't compromise other communications' privacy and security.

Prior to CALEA the telecommunications industry did not consider the needs of law enforcement in their design criteria. Designing a "back door" for anything but problem determination and prevention would be designing a security flaw into the product. In addition manufacturers and service providers would be concerned about additional costs in a very competitive environment [2].

Although the law excluded "information services", in August 2005 the Federal Communications Commission (FCC) ruled that broadband voice over IP (VoIP) must comply with CALEA. This resulted in a multipronged legal challenge to the Federal Communications Commission's (FCC) decision mounted by a number of organizations. The Center for Democracy and Technology (CDT) and many other organizations presented arguments to counter the FCC's assertion that certain VoIP services were to be considered as telecommunication services subject to CALEA regulations. The CDT and others viewed the decision as a harbinger of greater intrusions into privacy rights possibly including peer-to-peer (P2P) voice technology hampering innovation at the application layer [3]. As of June 2006 the courts have continued to uphold the FCC's application of CALEA to carriers providing VoIP services.

Applying CALEA to VoIP will extend insecurities into the internet possibly inhibiting the technological and economic growth which the internet has spawned. The internet was fundamentally designed to place the intelligence at the endpoints with a relatively simple network in between. Due to the layered structure of the internet this will require pushing surveillance capabilities deeper into the protocol stack to achieve 100 percent success for legal VoIP wiretaps [1].

This inherent conflict between the end user's desire for confidential secure information exchange and the requirements of CALEA applicable VoIP Lawful intercept are extenuated by the determination of the U. S. government that information and telecommunications is a Critical infrastructure (CI). A CI sector is defined as "systems and assets," whether physical or virtual, so vital to the United States that the incapacity or destructions of such systems or assets would have a debilitating impact on security, economic security, national public health or safety, or any combination of those matters [24].

The recent exposure of hacker cell-phone "bugging" of a significant number of high ranking government and military officials in Greece has emphasized the need for proper administrative management of the "wiretapping" capabilities of such communication systems [25]. Although the "Athens Affair" dealt directly with cell-phone exposure the increasing migration of telephony to VoIP has made the security of the media stream through the use of Secure RTP a valid and necessary protocol to provide confidentiality and integrity in VoIP sessions. This requirement necessitates the use of key management to protect the exchange of encryption keys as a parameter in Secure RTP. Multimedia Internet Keying (MIKEY) is one recommended protocol for key management [24]. Encryption keys may also be passed in the message body of Session Description Protocol, the message body encrypted with Secure/Multipurpose Internet Mail Extensions (S/MIME), as part of a Session Initiation Protocol exchange in establishing a VoIP session.

2.2 Research and Standards

In 2000 an Internet Engineering Task Force (IETF) Network Working Group determined that adding wiretapping capabilities to Internet protocols would make the Internet less secure providing a feature that could be used both for and against U.S. institutions and its citizens. The addition of wiretapping to the internet protocols poses risks to the U.S. economy through loss of corporate information and national security through large-scale intelligence gathering. The risks outweigh the benefits of applying CALEA to VoIP likely decreasing rather than increasing our security [1].

The IETF's RFC 2804 is strictly an informational document relating the IETF's policy on wiretapping [4]. In this document the IETF decided not to create an IETF standards document encompassing wiretapping. Some of the reasons for this position are stated briefly below:

- IETF as an international standards body is the wrong forum.

- The IETF sets standards for communications across the internet that encompass a wide range of security requirements. The users of the internet are best served by the security properties of the connections being well know and as free from loopholes as possible.

- The IETF does not see an engineering solution that allows wiretapping when end systems take adequate measures to protect their communications.

- Adding a requirement for wiretapping will make protocol designs more complex, jeopardizing security in unintended ways.

- There should be wide availability of strong cryptographic technology for users of the internet to protect against illegal intrusion.

- The technologies designed for wiretapping should be openly described so that they adhere to their design constraints as well as publication of the technology's weaknesses.

The IETF's decision not to publish a surveillance standard has placed the internet service providers and equipment manufacturers in the position of proposing standards to meet the CALEA requirements. Also the support by the IETF of strong end user cryptographic technology reduces law enforcement to dependence on the available decryption capabilities.

2.2.1 Research

Some efforts to develop standards have come from university projects [5]. One such project describes the surveillance of phone calls in an H.323 network. In this paper four different surveillance techniques were described based on the ITU-T H.323 standard: Wiretap on Gateways, Wiretap Routing on Gatekeeper, Fixed Route Wiretap, and Promiscuous Wiretap.

Gateways are the interface between H.323-based IP telephony networks and the PSTN. They perform translation functions for the signaling messages and voice streams between the H.323 protocol and the PSTN and visa versa. Since the gateway sits astride the H.323 network and the PSTN it represents a fixed point of call routing and interception. The gateway has access to the entire signaling and voice content of the call passing between the two networks.

In order to intercept a specific call the gateway must maintain a list of endpoints designated for surveillance.

The Wiretap Routing on Gatekeeper approach makes use of the Gatekeeper component of an H.323 network The Gatekeeper component is responsible for call setup procedures and maintains a list of registered users with their aliases and IP addresses. The H.225 Registration, Admission, and Status (RAS) messages are exchanged between the endpoint and the gatekeeper. Call signaling messages and call control messages are either exchanged directly between endpoints or routed from endpoints to another over the gatekeeper. Voice media streams are routed directly between endpoints. If one of the parties has been designated for surveillance the gatekeeper sends a message to the wiretap device communicating that a call between endpoints A and B is being established and needs to be intercepted. The message sent to the wiretap device includes the IP addresses of both endpoints. The gatekeeper replies to the ARQ message from the caller with an Admission Confirm (ACF) message that contains the IP address of the wiretap device instead of the callee. The rest of the call setup progresses with the wiretap device inserted in between end points A and B. The entire call content may be stored in the wiretap device.

The Fixed Route Wiretap method is similar to the Wiretap Routing on the Gatekeeper method. Since it is possible to detect an intercepted call by comparing user and Wiretap device IP addresses all calls are routed through the Wiretap device by the Gatekeeper. The Gatekeeper then informs the Wiretap device of which calls need to be intercepted and recorded. The Wiretap device then becomes a new component of the H.323 network.

. In the Promiscuous Wiretap method a device is added which monitors the entire network traffic. The device would extract media streams and signaling information designated for surveillance The Promiscuous Wiretap Device is connected to the local area network switch. The switch must be configured to forward all traffic to the monitoring port to which the wiretap device is connected. The Promiscuous Wiretap device is then able to extract the signaling and media stream of the entire conversation.

This same group of researchers also proposed a distributed architecture for lawful interception. Elements of which have been incorporated into existing industry standards [6]. The proposed architecture used a proprietary protocol, Wiretap Information Exchange Protocol (WIEP) to define communication procedures and distribute the necessary information.

The hierarchical architecture had a Top-Level Device (TLD) which is the central management component of the system. The TLD contained the defined policies for a specific LI. The policies were distributed to the Intermediate-Level Device ILD. The ILD served a mediating function passing along the LI policies to other ILD units. The ILD could also intercept VoIP calls and record some categories of Wiretap data. The Bottom Level device had no descendents and served to intercept the VoIP calls. The components exchanged information which fell into two categories: Wiretap Information and Wiretap Data. Wiretap Information consisted of a set of VoIP endpoints that were defined by the LI as well as additional intercept policies. All intercepting components of the system used this information to identify the designated calls to be intercepted. Wiretap data was acquired during the interception of the VoIP calls. Wiretap data included information such as recorded call start and end times, participating IP address and the recorded content of the call. The last component of the system was the Storage Device (SD). The SD was not a part of the hierarchical control structure but provided a means of remote storage of Wiretap Data. The system was capable of functioning in a multiprotocol environment such as H.323 and SIP.

WIEP defined the communication procedures in the distributed system. Wiretap Information and Wiretap Data were the two categories of information exchanged. The message format used in WIEP was XML for both categories. Four methods were used in endpoint identification as part of the Wiretap Information: define a host ip address, define a subnet address or range of address, and finally an alias as a high-level identification element. Also included in the Wiretap Information list were specifications for interception policies which control the interception process for the target call. The attributes which defined these policies were whether the LI system will intercept just call signaling, call content, or both as defined by CALEA.

Also included in the Wiretap information were time and date pairs for the start and end times of the interception as well as storage method, local or remote. It was important to protect the recorded information therefore, WIEP used authentication, authorization and secure transmission to transmit the data to be stored and limit its access.

Research is on going in the area of new CALEA compliant protocols which interact with VoIP gateways to permit the continuation or termination of a VoIP session based on the availability of the session keys to decrypt the VoIP content subject to a lawful intercept [26]. However, due to the various key protection and usage scenarios available in SIP the challenge remains in obtaining these session keys in the least obtrusive and detectable manner and making them available to a Law Enforcement Agency when a User Agent is subject to a lawful intercept.

As previously discussed the proposed CALEA architecture [5] does not provide for the decryption of signaling or content data in a Lawful intercept. However as mentioned in the previous paragraphs authentication, authorization, and secure transmission are important requirements to meet the general security constraints of CALEA. It is the integration of the end users desire for encryption and privacy as well as the CALEA protection requirements in accessing and transmitting the intercepted data with the associated encryption keys which provide the impetus for this research.

2.2.2 Industry Standards

The IETF has taken the position that it is not a forum to generate a standard for Lawful intercept architecture. Internet industry component manufacturers have step into this vacuum. Cisco has developed its own internal standards which selected component offerings will support. Cisco has published two documents describing these architectures.

PacketCable™ is an initiative led by CableLabs of Colorado [7]. The organization defines standards for the Cable TV industry. The purpose is to provide interoperable interface specifications to deliver realtime multimedia services over two-way cable networks. The PacketCable networks use IP to enable a range of multimedia services, such as IP telephony. This capability has brought them under the Lawful Intercept (LI) requirements imposed by CALEA. To support these requirements for Cisco supplied components, Cisco has developed the **PacketCable Lawful Intercept Architecture for BTS** (Broadband Telephony Softswitch) **Architecture for BTS Versions 4.4 and 4.5** documents [8].

PacketCable Lawful Intercept Architecture for BTS (Broadband Telephony **Architecture for BTS Versions 4.4 and 4.5** addresses voice traffic only. The PacketCable specifications are considered a SafeHarbor for compliance with CALEA.

The key requirements of any LI architecture are listed below:

- LI must be undetectable by the intercept subject. Therefore, the tapping must take place within the domain of trust of the SP or ISP (on a cable modem termination system [CMTS] or trunking gateway) and performed along the normal path of data.

- Multiple LEAs intercepting the same subject must not be aware of each other. This means a one way flow of information from the mediation device which is duplicating the information to each LEA collection point.

- Unauthorized personnel must be prohibited from gaining knowledge of the LI capability.

- The information identifying intercepts, phone numbers, IP addresses, etc., must be correlated to the content of the intercepts.

- The reliability of delivery of intercepted information to the LEAs must be the same as the original delivery of information to the customers.

The following paragraphs describe the PacketCable Lawful Intercept Architecture. Figure 1 is a diagram of the generic network. The topology of networks supporting LI consist of five major components; LI Administration Function, Mediation Device, the Intercept-Related Information Access Point (IRI IAP), the Call Content Intercept access point (CC IAP), and the collection function. The first four are part of the service provider network. The collection function is the LEA responsibility.

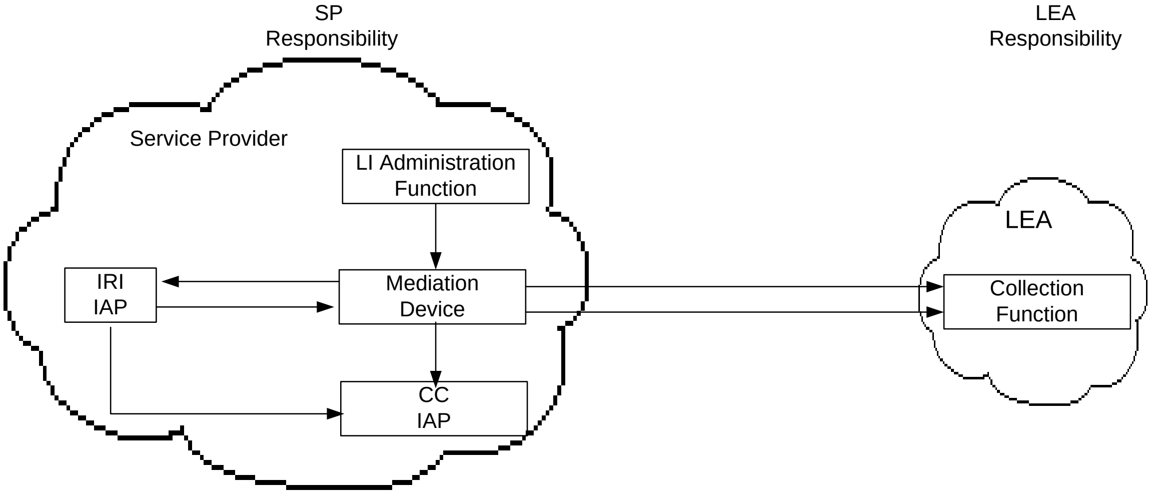

Figure 1. Generic IP Network That Supports LI of Voice and Data Traffic

The following paragraphs describe the functional components of a PacketCable LI Network. Figure 2 shows a functional description of the LI Network.

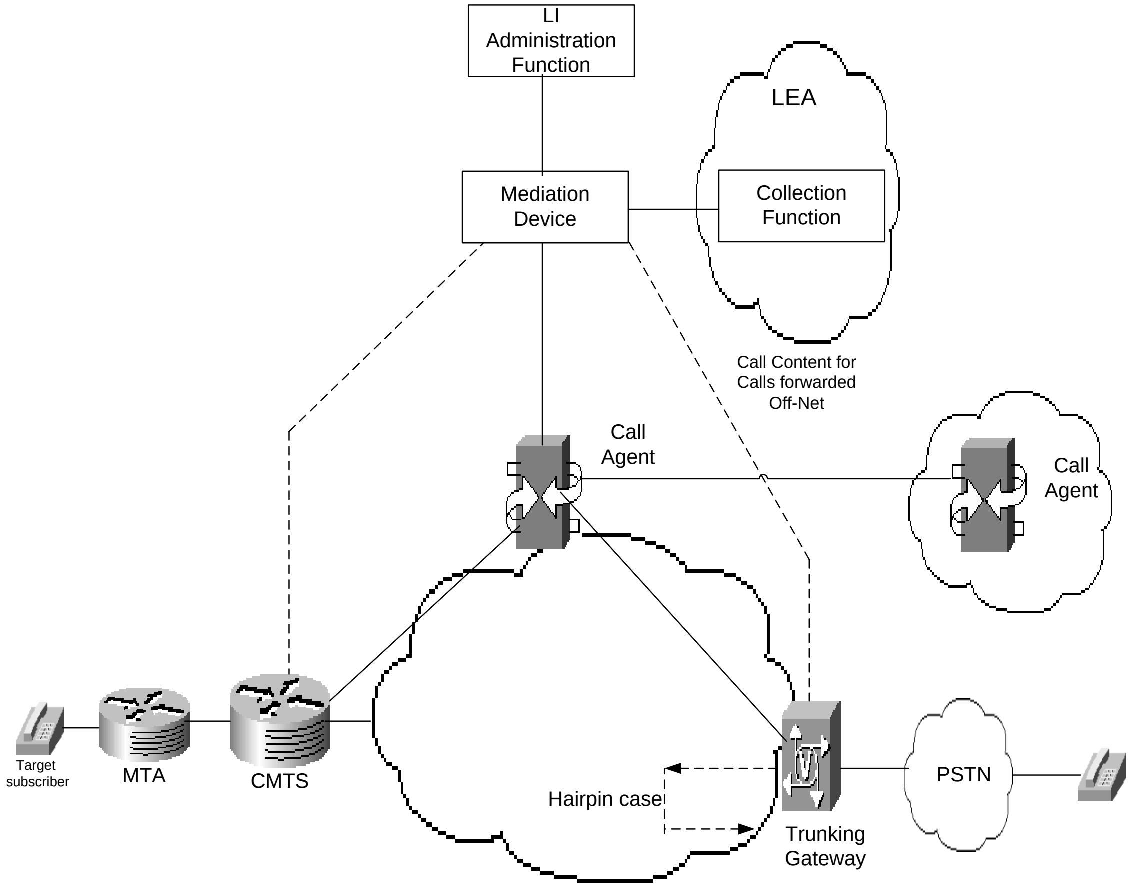

Figure 2. Functional Description of a PacketCable LI Network

The LI administration function is used by the Service provider to provision the intercepts to interface with the other network components. It administers intercept orders, tracking and maintaining intercept information. It also supervises the security and integrity of the LI process.

The mediation device (MD) is an SP or ISP maintained component. It sends configuration commands to the Intercept Access Points (IAPs) to enable intercepts. It receives intercept information, encapsulates it and delivers it to the LEAs.

The Intercept-Related Information Intercept Access Point (IRI IAP) is the component that provides identification information to the mediation device. This information includes the source and destination phone numbers and IP addresses. It also includes any post call-establishment messaging such as three-way calling or call forwarding. The architecture would determine the IRI IAP component. It could be either a call agent, the call modem termination system (CMTS), Session Initiation Protocol (SIP) proxy, or the gatekeeper. The call agent supplies call control-related information and the CMTS supplies QoS- related information.

The Call Content Intercept Access Point (CC IAP) intercepts call content information replicates it and forwards the replicated information to the mediation device. The CC IAP is located as close to the source as possible minimize the number of monitored calls to ensure the CC will be intercepted.

The collection function is a storage device maintained by the LEA. It receives intercept information and stores it from the mediation device.

The PacketCable LI Architecture functions in the following manner. The LEAs have three types of surveillance which may be authorized: (1) "pen register," which records call-identifying information for calls originated by the subject, (2) "trap and trace," which records call-identifying information for calls received by the subject, and (3) "interception," which allows LEAs to intercept the call content as well as the call identifying information of the conversations of the subject. This type of intercept is referred to as a Full Content or Title 3.

It is of particular interest to note that the PacketCable LI architecture adheres to the PacketCable Electronic Surveillance Specification [10]. In that document the PacketCable Telecommunications Service Provider (PC/TSP) is not responsible for the decrypting, or ensuring a LEA's ability to decrypt any intercepted information unless the encryption was provided by the PC/TSP and the PC/TSP has access to the information necessary to decrypt the intercepted communication. CALEA does not prohibit a carrier from providing an encryption service for which it does not keep the ability to decrypt intercepted communications for LEAs [10].

When initiating an intercept a warrant is issued which the LEA physically delivers to the service provider. The service provider uses the LI administration function to enable LI of the specified target. The mediation device configures the tap for the authorized start date and time. Once the intercept is provisioned on the mediation device the individual intercepts are initiated automatically. The expiration date is configured on the mediation device. At the time of expiration the mediation device automatically removes the configuration for the warrant.

Cisco Service Independent Intercept Architecture Version 2.0 was developed to support the LI legislation requirements imposed on service providers (SP) and internet service providers (ISP). It is similar in concept to Cisco's **PacketCable Lawful Intercept Architecture for BTS Architecture for BTS Versions 4.4 and 4.5.** The document describes the implementation of an LI architecture on a Cisco IP network using version 1.0 of Cisco LI Management Information Base (MIB) for Voice over IP (VoIP) and IP data intercepts. The key requirements of the SII architecture regarding the authorization and access to the LI process and of LI information is the same for both architectures.

Cisco's approach incorporates a number of features to meet the previously listed LI intercept requirements:

- Use of standard access list technology to provide the intercept

- Encapsulation of the entire intercepted and replicated packet so that the original source and destination addresses are available for intercept purposes.

- Use of a control plane for intercept that is different from call control, preventing network operations personnel from detecting active intercepts in the network.

- An approach that limits the intercept activity to the router or gateway that is handling the target's IP traffic and only when the target is accessing the network.

- No LI-related command line interface commands that could allow for the detection of intercept activity.

- LI-related MIBs and traps sent only to the (third-party) components controlling the intercept.

- Support for multiple encapsulation and transport formats.

The over all function of the SII architecture is very similar to that described in the PacketCable architecture. Figure 3 illustrates the interface function between the components in the SII architecture. The LI administration and mediation device functions are the same as the PacketCable architecture. The remaining functions are similar also but change somewhat do to the variation in network components.

The Intercept-Related Information Intercept Access Point would in addition to voice and data intercepts would be the authentication, authorization, and accounting (AAA) server.

The CCIAP as in the previous architecture would be located as close the source as possible to minimize the number of monitored calls. An edge router, a trunking gateway, or an access server can be the CC IAP. The collection function would remain the same.

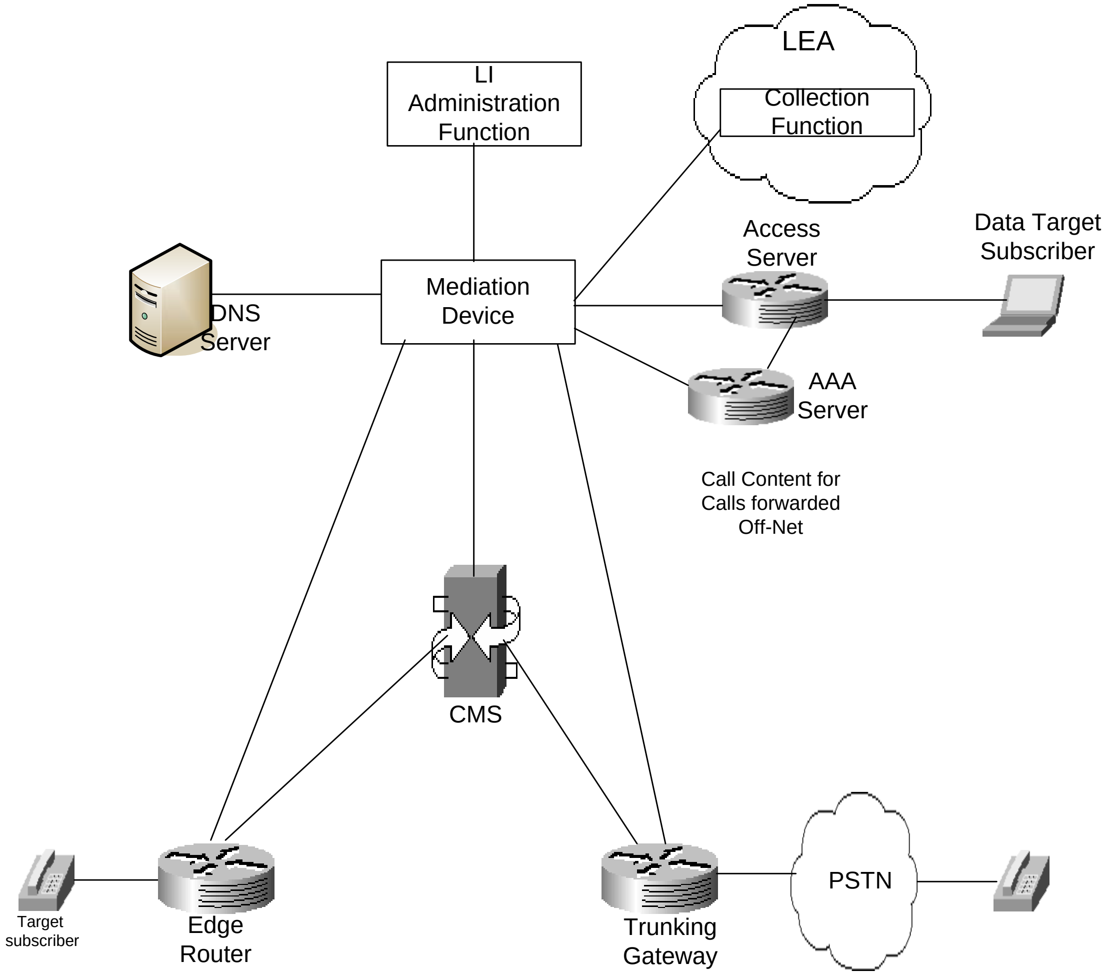

Figure 3. Cisco SII Voice Intercept Device Interfaces

Security is a major consideration as the integrity and privacy of intercept data transferred from the mediation device to the LEA collection function component. The following threats need to be defended against:

- Impersonation of LEAs and mediation devices
- Privacy and confidentiality breaches
- Message forgery
- Replay attacks

The proposed architectures apply various security mechanisms to protect the information obtained through LI. Any integration of end-point encryption key management provided by an ISP must incorporate adequate security protection. The MD and LEA servers should be hardened against attackers. All the interfaces must be able to provide strong cryptographic authentication. The authentication should match the function that is being performed. The interfaces must perform message integrity checking (such as Hash-based Message Authentication Code [HMAC]-Message Digest [MD5].

The two LI architectures PacketCable and SII formed the basis for a generic informational RFC published by the IETF as RFC 3924 [11]. In keeping with the IETF's position regarding not being a forum for a wiretap standard the RFC contains a disclaimer. The disclaimer clearly states that it is not an Internet Standard and makes no claims as to its compatibility with existing protocols.

An additional component suggested as useful in securing VoIP sessions is the Session Border Controller (SBC) [24]. The SBC functions at the signaling layer of the protocol stack and would be an additional component directly between the edge router and target subscriber. The SBC's ability to interpret signaling protocol such as SIP would make it an ideal device to participate in a CALEA network to help resolve encryption challenges posed by S/MIME and MIKEY yet without undermining the user's desire for security.

2.2.3 Security Mechanisms

As noted previously there are a number of security mechanisms applicable to VoIP SIP signaling and call content. Also CALEA has a number of Lawful Intercept process protection and access protection requirements that place authentication, message privacy, and integrity requirements on the overall LI process as mentioned in the review of Cisco's SII architecture. The following paragraphs present an overview of the security mechanisms applicable to VoIP and LI.

2.2.3.1 SIP Authentication

As part of the overall architecture the service provide must allow the SIP authentication mechanisms to function unimpeded, although the exchange could provide useful Intercept-Related Information to be collected in a LI. The SIP authentication function takes two primary forms, user agent authentication by another user agent, and second, user agent by proxy server [12]. SIP borrows digest authentication from HTTP [13]. Briefly, a UA sends 401 Unauthorized response to another UA. The 401 response contains a WWW-Authenticate header field which contains specific information regarding the digest challenge. A parameter in the Authenticate header field is Quality of protection (qop). The qop may either be set to auth for just authentication or auth-int for authentication and integrity. Another parameter supplied is a nonce which is a time sensitive random string. The responding UA generates a digest response calculated by applying the MD5 hash function to a concatenation of username, password, nonce, SIP request type (or method), and the request-URI. MD5 is a mathematical function that produces a one way fixed length digest of a message. SIP digest authentication protects against replay attacks. A replay attack is one in which an attacker captures a SIP request and retransmits it posing as the calling UA. Similarly a proxy server may require that a UA authenticate itself to the proxy server, in the form of a 407 proxy authentication required response, before a request will be processed. The response by the UA to an authentication challenge is similar to that of the UA to UA authentication challenge with some additional parameters as part of the concatenation.

In addition to authentication message integrity can also by applied. When qop=auth-int the response includes the MD5 message digest of the message-body which SIP is transporting. As mentioned previously one of transported protocols is SDP which contains information on the origin IP address, port number, media type, media encoding scheme [22] as well as additional information which would be of importance in IRI.

2.2.3.2 SIP and S/MIME

To provide for integrity and confidentiality in SIP messaging Secure/Multipurpose Internet Mail Extensions is employed [14]. SIP messages exchanged during session setup contain information including:

- Both parties' SIP URI's and IP addresses
- Two parties have established a call:
- The IP addresses and port numbers associated with the media (call content)

- Keying material carried in the SDP for Secure Real Time Protocol (SRTP)

This content is precisely the information which a Lawful intercept Intercept related information would contain.

In order to implement S/MIME the SIP UAs must support the digital signature algorithm RSA. A digital signature is a message hash value that has been encrypted using that person's private key in the asymmetric (public key) cryptography architecture. RSA is the most common algorithm for this process. The UAs must also support SHA-1 as a message digest hash, and advanced encryption suite (AES) as a message encryption algorithm. AES is specified for SIP in RFC 3853 [15]. In order for the receiving UA to validate the digital signature the receiving UA will verify the identity via a Certificate Authority to obtain and validate the sending UA's public key. The end-user address that the certificate asserts is valid is a concatenation of the "userinfo""'@"and "domainname" portions of a SIP or SIPS URI [12]. Although SIP requests and responses cannot be completely encrypted end-to-end in their entirety because the message fields pertaining to routing must be visible to the proxies to route correctly, S/MIME allows SIP UAs to secure message bodies end-to-end without affecting message headers providing end-to-end privacy, integrity and mutual authentication. This presents a challenge to LEAs when obtaining complete IRI or desiring to decrypt the call content if the key is encrypted within the SDP message body via S/MIME.

2.2.3.3 Transport Layer Security

SIP utilizes Transport Layer Security (TLS) [16] to secure signaling messages on a hop by hop basis. Each hop is a separate TLS session. The end points of each hop independently negotiate a separate secure TLS tunnel. Because of the separate hop by hop negotiation the connection information is exposed at each server in the negotiation path. TLS uses TCP as its reliable transport protocol. Port 5061 is used for SIP requests. TLS provides for privacy and data integrity between these endpoints using two TLS layers: the TLS Record Protocol and the TLS Handshake Protocol.

The TLS Record Protocol provides security having two basic properties: privacy through encryption such as DES or RC4. The keys for the symmetric encryption are unique for each connection and based on a secret negotiated by TLS Handshake Protocol. Message integrity is protected using a keyed Message Authentication Code (MAC). Possible hash functions are SHA and MD5 as well as others. The TLS Record Protocol encapsulates the high level protocol TLS Handshake Protocol.

The TLS Handshake Protocol provides security for the connection between server and client having three basic properties:

- Authenticating peer's identity using public key, cryptography (RSA)
- Negotiation of a shared secret unavailable to eavesdroppers (used by the TLS Record Protocol for data encryption)
- The negotiation is reliable so that no attacker can modify the negotiation without being detected.

Finally to guarantee that TLS on TCP will be applied to every hop on the signaling path from the endpoint caller to the domain of the called party a URI scheme sips (Secure SIP) may be specified during a SIP request [12].

2.2.3.4 Media Security Secure Real-time Transport Protocol (SRTP)

The previous paragraphs have discussed the issue of signaling security in VoIP. The following paragraphs discuss the security provisions for call content.

A VoIP media stream uses is encapsulated in Real-Time Transport Protocol [17 which utilizes User Data Protocol (UDP). RTP is the standard protocol used to transport the actual voice conversation. RTP Control Protocol (RTCP) provides some session control and information over RTP sessions.

The protocol used to secure the call content is the protocol Secure Real Time Protocol (SRTP) [19]. There are a number of challenges to providing media stream security [18] among these are:

- Completing a real time exchange of keys, crypto suite, and parameters without clipping the start of the conversation;
- Performing encryption, decryption, and authentication without introducing significant media latency or extra bandwidth;
- Rekeying without interrupting or adding delay to the session.

Figure 4 illustrates the structure of a Secure Real-Time Protocol packet. The shaded first 12 bytes of the packet are the original part of an RTP packet.

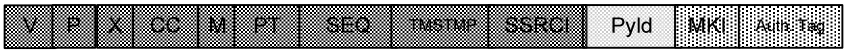

Figure 4. Secure Real-Time Protocol Packet

The fields in the first 12 byte header field are described as follows:

- Versiion (V). A 2-bit field set to 2, the current version of RTP

- Padding (P). If this bit is set, additional padding octets placed on the end of the packet to make the packet a fixed length for block ciphers

- Extension (X). If this bit is set, and additional extension of 4 bytes is added to the header to accommodate certain payload types.

- CSRC count (CC). This 4-bit field contains the number of content source identifiers (CSRC) present following the header (only used for mixers of multiple RTP streams.

- Marker (M). This bit indicates the start of a new frame in video or in silence-suppressed speech, a talk-spurt

- Payload type (PT). A 7-bit field defining the codec in use. This number matches the profile number in the SDP.

- Sequence number (SEQ). A 16-bit field that is incremented for each RTP packet sent. It is used to detect missing/out of sequence packets.

- Timestamp (TMSTMP). A 32-bit field indicating the relative relative time when the payload was sampled. It is used to help a receiver remove jitter when playing back stored packets.

- Synchronization source identifier (SSRCI). A 32-bit field identifying the sender of the RTP packet. A SSRC number is chosen randomly at the start of a session by each participant.

- Contributing source identifier (CSRC) list. This field can be absent or as many as 15 32-bit fields in the header. The number is set by the CSRC count header field. This is only present when the RTP packet is being sent by a mixer.

The last two fields in the header Master Key Index (MKI) and the authentication tag (Auth Tag) are extensions added by SRTP to the original RTP header. These two fields will be discussed in the following paragraphs.

Each RTP stream established by SIP has a different UDP receiving port and a unique SSRC identifier. The UDP port number and SSRC of the sender uniquely index a media stream in on direction so that conversation is made up of two RTP sessions. SRTP uses symmetric keys and ciphers to encrypt the payload but does not provide any key management or generation. SRTP relies on an out of band key management and exchange as mentioned above in the discussion of SDP.

SRTP defines how session keys are generated from the agreed out of band establishment of master keys and how the session keys are utilized or refreshed during a session. SRTP uses AES in counter (CTR) mode to encrypt the media stream. The cipher text is produced by exclusive ORing (XOR) an RTP key stream with the media plain text. The key stream is an encryption of a counter initialized with an initialization vector (IV). The IV is generated using salting key, the SSRC, and the SRTP SEQ number plus a rollover counter

(ROC). For message authentication a HMAC SHA-1 hash is performed over the packet and added to the end.

A master key and master salt are inputs to a function defined within SRTP. This function generates the session encryption key used to drive the AES encryption algorithm. The session salt key also produced from the SRTP function is used as input to the IV. The session authentication key is used for the HMAC SHA-1 message authentication function. The RTP header is not encrypted by SRTP however, the RTP payload is encrypted. The master key index (MKI) field added by SRTP is used by key management for rekeying or refreshing keys. The authentication tag provides integrity protection over the RTP header and payload. The MKI is not protected by the authentication tag. When authentication is used encryption is applied to the payload first followed by the authentication. Figure 5 illustrates the concept of master key/master salt to session keys/session salt keys.

As mentioned in the above discussion SRTP makes use of master keys established by some key management system prior to their use in the SRTP encryption function. There are a number of methods to exchange the master keys.

Preshared keys is one approach in which the UAs have previously exchange a secret key to use in an algorithm. Since the keys are not carried in the SIP signaling, the signaling protocol does not need to be secured. This approach is only effective in small groups and is subject to cracking.

Another method referred to in previous paragraphs the public key encryption approach where the sender encrypts the RTP media packets with the public key of the receiver. The receiver uses his private key to decrypt the media stream. This approach is inefficient and computationally intensive and means that keys are not actually refreshed. It is a better approach for the authenticated parties to securely negotiate a session key at the time of the session setup. The Diffie-Hellman public key exchange could be used to generate the same master key without passing a secret key between the two negotiating parties.

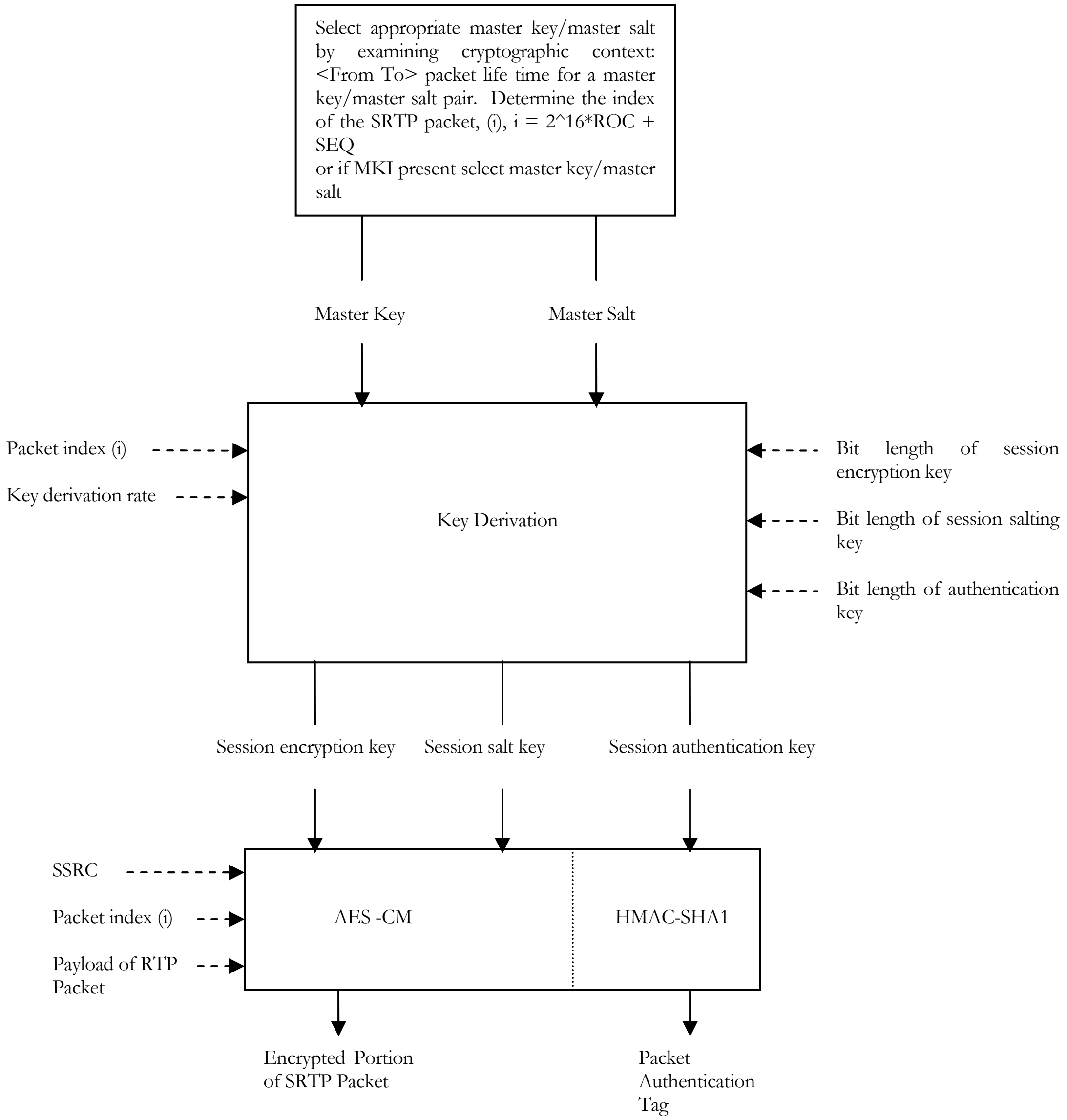
Select appropriate master key/master salt by examining cryptographic context: <From To> packet life time for a master key/master salt pair. Determine the index of the SRTP packet, (i), i = 2^16*ROC + SEQ
or if MKI present select master key/master salt
Master Key
Master Salt
Packet index (i)
Key derivation rate
Key Derivation
Bit length of session encryption key
Bit length of session salting key
Bit length of authentication key
Session encryption key
Session salt key
Session authentication key
SSRC
Packet index (i)
Payload of RTP Packet
AES -CM
HMAC-SHA1
Encrypted Portion of SRTP Packet
Packet Authentication Tag

Figure 5. Master Key/Master Salt to Session Keys/Session Salt and SRPT Encryption

Authenticated end parties may establish a secure signaling session as mentioned in the secure SIP signaling discussions. The authenticated parties can exchange session keys or material to derive session keys. Secure SIP may also be used to protect each hop in which the SDP message body carries the keying material secured by TLS. However the keying material is exposed in every proxy server.

The use of the a=crypto attribute in SDP could be used to carry SRTP keying and configuration information [21]. This information would contain the cipher set and the master key and master salt to be used by the cipher set as well as MKI information. The use of S/MIME as mentioned in the discussion of SDP security provides end-to-end protection of keying material carried in the SDP message body.

Multimedia Internet Keying (MIKEY) [20] is a protocol for use with SRTP. MIKEY may be used with SDP. The SDP message body does not need to be encrypted as MIKEY supports is own encryption and integrity protection. A number of key exchange methods may be used with MIKEY including; preshared key, public key and Diffe-Hellman key generation. The output of the MIKEY process once the key exchange method is applied are the master keys, traffic encryption key (TEK), and master salt used by the SRTP function to generate the session encryption keys, session salt keys, and the session authentication keys as previously discussed.

2.3 New architecture requirements

A new architecture which will meet the CALEA requirements and provide LEAs with encryption keys will need to posses these attributes:

- When an end user attempts to establish an encrypted signaling or media session over SIP using keys not available to the ISP that session will be rejected and an appropriate error message response.

- The end user will then be offered a negotiated process to establish signaling and media session encryption keys in an authenticated secure manner to protect the end user's message privacy.

- If LI is not required for that signaling and media session the keying information will only be retained as appropriate to support the current session and discarded there after.

- If LI intercept of the signaling or media information is required the timestamped signaling and media session keys will be stored and correlated with the signaling and media information as required within the established CALEA compliant intercept architecture.

To establish a new architecture the issue of the RSA digital signature algorithm and its use of the public key infrastructure must be accommodated. S/MIME, Secure SIP (utilizing TLS), and MIKEY make use of the RSA digital signature/ pubic key encryption in the cipher suite. Figure 6 illustrates the protocol and related encryption layers for SIP.

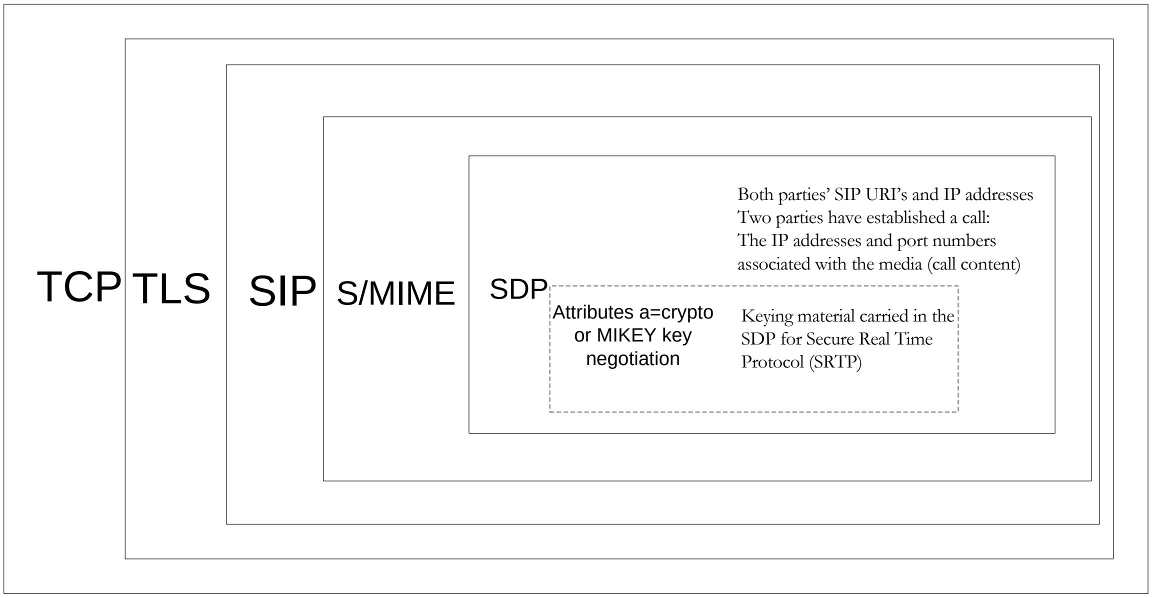
TCP
TLS
SIP
S/MIME
SDP
Both parties' SIP URI's and IP addresses
Two parties have established a call:
The IP addresses and port numbers
associated with the media (call content)
Attributes a=crypto
or MIKEY key
negotiation
Keying material carried in the
SDP for Secure Real Time
Protocol (SRTP)

Figure 6. SIP Protocol and Related Encryption Layers

CHAPTER 3

Approaches to a New Lawful Intercept Architecture

This chapter briefly describes two approaches to solving the security issues that were highlighted in the above discussions. The key point being that the use of the public key infrastructure with RSA makes the decryption of relevant Lawful Intercept information contained in the SIP/SDP protocol and the SRTP keying information very difficult without access to the appropriate private keys. Two approaches are then proposed to solve this difficulty.

To accomplish either of these approaches the SIP proxy server function must be modified and or a Session Border Controller (SBC) must be functioning as an intermediary to establish the SIP call signaling. A SBC is utilizes a back-to-back user agent B2BUA to act as a user agent server to establish a border between a private and a public VoIP network. The B2BUA responds to SIP requests from UAs via the public network and after applying policies forwards the reformatted requests to the client UAs on the private side. The B2BUA functions in reverse when calls are initiated by its client UAs.

3.1 The LI Mediation Device Initiating the Acquisition of the Private Key

Upon the receipt of a LI request by a LEA the Mediation Device will determine the CA and obtain the private keys of the relevant parties. For this transaction there must be an absolute guarantee of authentication and privacy between the communicating devices. An encrypted call signaling session would proceed as follows:

- The UA initiates a SIP signaling call with a desired cipher suite (ie. S/MIME)

- The B2BUA device (SBC) if configured by the ISP Mediation device to intercept the call the SBC will query the Mediation device to verify that the private key has been obtained.

- If the response is affirmative it informs the mediation device that a call is being established and sends IRI (containing keying material) to the intercept point and allows the call to continue.

- If the query reveals that the call being established is not the subject of a LI B2BUA will wait an appropriate amount of time before allowing the call to proceed as normal. This is done so that all calls appear the same whether subject to intercept or not.

The major disadvantage to this approach is the exposure of the UA's private key and the understandable reluctance of the internet community to implement such a scheme.

3.2 Session Border Controller Intermediary Security Negotiation

The second approach is for the UA to establish all security negotiations with the SBC. The initiating UA will use the public key of the SBC for the S/MIME security encryption. The SBC will use its private key for encryption when establishing S/MIME or other security requirements as required.

The SBC will also assume all of the MIKEY key negotiation functions for SRTP between the initiating UA and the public internet essentially posing as the called UA. The remote SBC will perform the same functions for the called UA. The functional process is illustrated in figure 7.

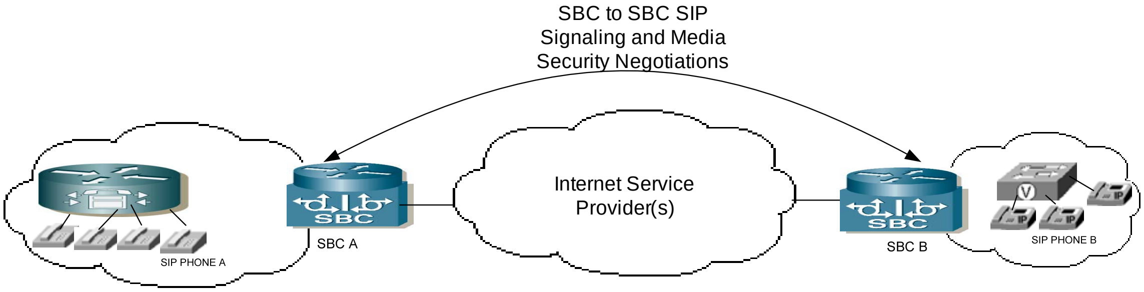

Figure 7. SBC to SBC Security Negotiations

The process will proceed in the following steps:

- Initiating UA establishes SIP signaling session with local SBC. Session may include specific security suites establish for SDP authentication, integrity, and privacy (S/MIME and MIKEY SRTP key negotiation).

- If the UA is subject to a LI the SBC obtains required IRI and forwards this to the Mediation Device after establishing a mutually authenticated secure connection using a defined cipher suite for encryption and integrity.

- At the conclusion of key negotiation for SRTP the engaged SBCs will forward the appropriate keys to the MD using the mutually authenticated secure connection and defined cipher suite as above.

- If the UA is subject to a LI and message content is required to be captured the MD will instruct the SBC to replicate the packets and forward over secure connection to the MD.

- The Service Provider Mediation Device will forward all relevant IRI and CC (if required) information to the LEA collection device.

An important issue exists regarding the encryption keys for the SRTP. It is possible to share a secret key out of band (apart from the SIP protocol) which will then circumvent the SBC to SBC security negotiation process. To remedy this situation the SBC will read the security description parameters in the SDP body describing the session. The media parameter m=audio 42850 RTP/SAVP signals a secure audio/video session with the device listening to port 42850. The SBC will recognize the secure media requirement. If the SRTP encryption keys or their negotiation through MIKEY has not been obtain by the SBC. The SBC will terminate the session and require the UA to send or negotiate the required keys within the SDP or MIKEY negotiations.

3.3 Flow Diagram and Timing Considerations

A possible scenario of a SIP signaling session with an SBC-to-SBC intermediary CALEA function is presented below and illustrated in figure 8. The intent of the signaling session is to establish a SRTP media stream using the a=crypto attribute in SDP with the SRTP keying information. The master key and master salt key will be protected by S/MIME. Using figure 6 as the architectural concept the source SBC (SBC A will be directly involved with the user agent SIP Phone A in the SIP/SIMME security negotiations and the SBC A will be responsible for obtaining the encryption keys and communicating with the Mediation Device.

3.3.1 SIP Signaling Session with SBC-to-SBC Intermediary CALEA Function

1. SBC A establishes a secure connection to mediation device

2. SIP Phone A initiates a session towards SIP Phone B via SBC A
    - Assuming that Phone A has the public key of SBC A and has authenticated SBC A
      Phone A sends SDP body (Description of Media Session) encrypted with pkcsy-mime (S/MIME), the message body is digitally signed with Phone A's private key to authenticate the message

    a. SBC A examines LI subject list to ascertain whether Phone A is on list
    - **No** Proceed to step 5
    - **Yes** SBC A determines if the intercept requires IRI or both IRI and Call Content Intercept (CCI)
        - **IRI only**
        - **IRI and CCI**

3. If VoIP Conversation from Phone A is subject to a Lawful Intercept SBC A begins decryption of SDP message body to determine if media session is to be encrypted
    - If SIP Phone subject of LI
        - **IRI only**

    a. SBC A sends IRI information to Mediation Device

4. If VoIP Conversation from Phone A subject to a Lawful Intercept and the media session is to be encrypted (m = audio *port* RTP/**SAVP**)

    a. SBC A decrypts crypto graphic context information

5. SBC A sends message to SIP Phone A based on results of master key/master salt deduction and determination of session keys if required
    - **If not subject of LI** : (after time delay) keys validated session continued
        - SBC A merely determines if master key and master salt are present. No attempt is made to determine session keys this delay is important so that all clients whether subject to LI or not have equal delays
    - **If subject to LI and session keys determined** : keys validated session continued
    - **If subject to LI and session keys not determined** : keys not validated session discontinued

- If VoIP Conversation from Phone A not subject to Lawful Intercept or session keys have been successfully determined, signaling message is rebuilt with S/MIME using SBC B public key and forwarded to SBC B

- If VoIP Conversation is subject to Lawful Intercept and session keys are not successfully determined an error message is sent to SIP Phone A and session is discontinued. Phone A must start SIP signaling initiation again.

6. SBC B Accepts signaling initiation and rebuilds S/MIME SDP message body if required.
    a. SBC B forwards signaling initiation information to SIP Phone B encrypted with SIP Phone B's public key and signed with SBC B's private key.
7. SIP Phone B responds with S/MIME SDP message body agreeing to call setup

8. SBC B rebuilds SIP Phone B response and forwards to SBC A.

9. SBC A rebuilds response and forwards to SIP Phone A
    a. Response is encrypted with SIP Phone A's public key and digitally signed with SBC A's private key

10. Secure Real Time Media Session begins
    a. SIP Phone a sends SRTP packets to SBC A
    b. SBC A copies packets for decryption
    c. Forwards encrypted packet to SBCB
    d. SBC B forwards encrypted packet to SIP Phone B
    a. SBC A forwards decrypts SRTP packet and forwards decrypted SRTP packets to Mediation Device over secure connection
    Mediation Device RTP packet contains:
        i. Complete IP and UDP header
        ii. Complete RTP header
        iii. Decrypted RTP payload

11. Session Ends
    a. SBC A forwards end of session message to SBC B
    b. SBC B forwards end of session message to SIP Phone B
    c. SBC A deletes all remnants of stored LI information for completed session.
    d. Ends secure connection with Mediation Device if necessary.

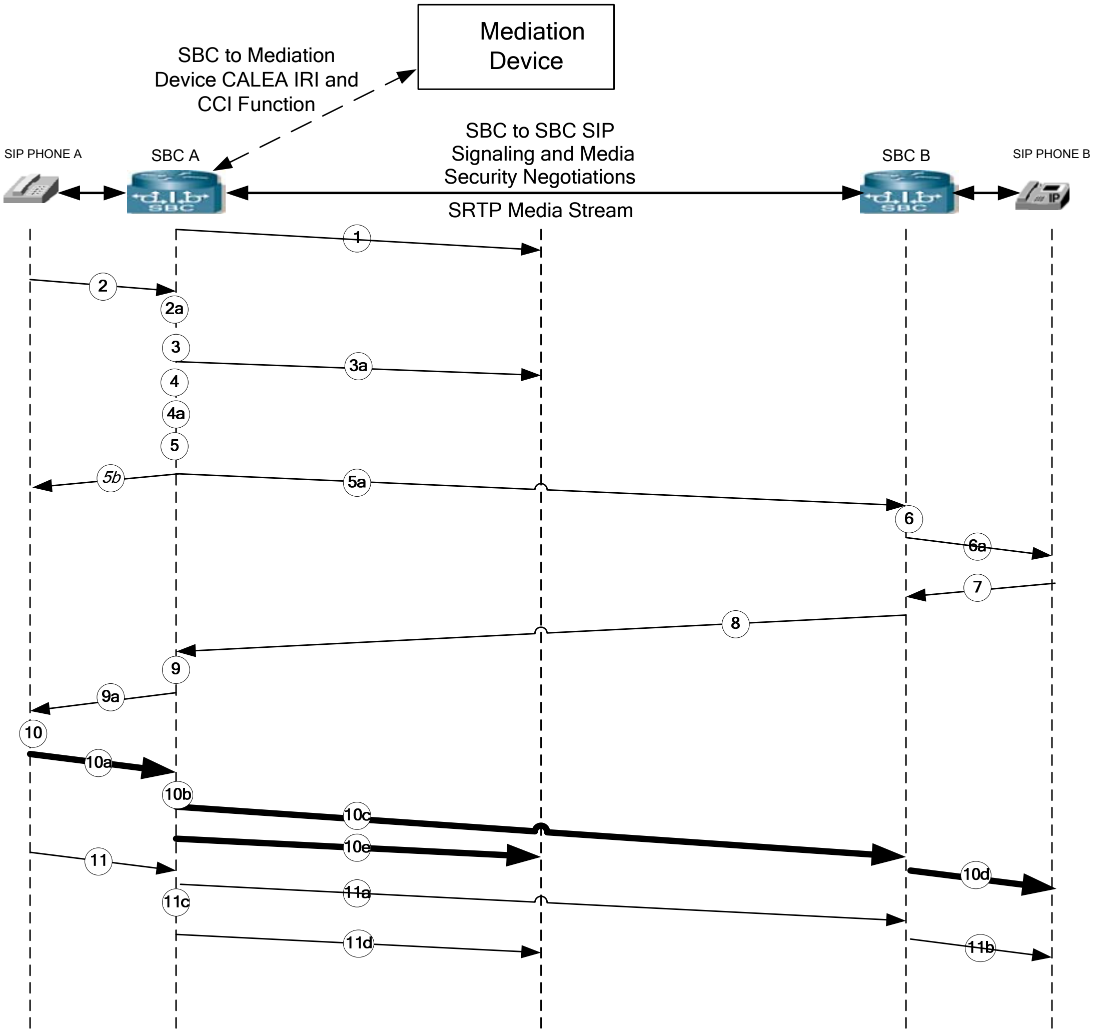

Figure 8. Flow Diagram of SIP Session with SBC-to-SBC Intermediary CALEA Function

3.4 SBC-to-SBC Intermediary Architecture Advantages

The SBC-to-SBC intermediary architecture has a number of advantages over an ISP managed key infrastructure.

- No new security infrastructure is created which the ISP would have to manage
- The session encryption keys are distributed in the sense that only an SBC with an active LI intercept has encryption keys stored limiting the damage of an attack
- The user agent is free to choose any of the standard available encryption protocols to secure his conversation giving the user agent a greater sense of security

3.5 Considerations In Deriving An Accurate Architecture Model

There are several reasons why an accurate mode should be determined for the SBC-to-SBC Intermediary architecture.

- As noted in the CALEA compliant architecture requirements a user agent subject to a Lawful Intercept must not be aware of the intercept. Therefore as suggested in paragraph 3.3.1 and the accompanying figure 8 any additional time necessary to process an unsecured/secured VoIP session that is subject to an intercept must be included in the processing time of a non-intercepted unsecured/secured VoIP session.

- The scalability of this proposal must be established by modeling as accurately as possible the number of Lawful Intercepts an SBC may process at any given time without introducing unacceptable latency in the media streams.

CHAPTER 4

The Session Border Controller Timing Model

This chapter presents the SBC timing model in detail with comments on the various functions of the SBC relating to an SBC lawful intercept of a SIP VoIP call. The proper timing model will be a combination of processes that could run prior to a lawful intercept call setup and processes that must run real time and in the background through the duration of the VoIP call.

4.1 IPsec Tunnel from SBC to Mediation Device

A function of the SBC which is referenced in Figure 8 and may be initiated and maintained prior to a User Agent initiated phone call that is subject to a lawful intercept is the secure IPsec communication link between the SBC A and the Mediation Device which is part of the Lawful intercept architecture. Equation 4.1 shows the generic timing components of Security Associations necessary to establish an IPsec tunnel [26].

IPsec tunnel between the SBC A and Mediation Device:

$$t_{ke} = t_{ke1} + t_{ke2} + t_{ke3} + t_{ke4} + t_{ke5} + t_{ke6} + t_{ke7} + t_{ke8} + t_{ke9} \tag{4.1}$$

Where:

t_{ke1} = ISAKMP SA Negotiation

t_{ke2} = ISAKMP SA Negotiation Final

t_{ke3} = Exchange Keying Material (source)

t_{ke4} = Exchange Keying Material (destination)

t_{ke5} = Exchange Identity Information (source)

t_{ke6} = Exchange Identity Information (destination)

t_{ke7} = IPsec SA Negotiation

t_{ke8} = IPsec SA Negotiation final

t_{ke9} = IPsec Hash

With the IPsec communication established between the SBC and Mediation Device the Mediation Device can send call intercept information to the SBC. This Lawful Intercept update time (t_{LIu}) could be an on going process, occurring at various times as new Lawful intercept information is received from a LEA..

Lawful Intercept Update time:

$$Update = t_{LIu} \tag{4.2}$$

4.2 VoIP call setup delay via SIP

The VoIP call setup time via SIP signaling will be a function the SBC A will need to respond to in a reasonable time to avoid application timeouts. However since this is the signaling process to setup the real time media stream precise setup time is not critical.

Standard VoIP Call Setup:

$$Delay = t_{cs} \tag{4.3}$$

In the case of SIP

$$t_{cs} = t_{cs1} + t_{cs2} + t_{cs3} + t_{cs4} + t_{cs5} + t_{cs6} + t_{cs7} \tag{4.4}$$

and;

t_{cs1} = INVITE delay time response of SBC A

The SBC A will examine the source of the INVITE message containing the identifying Intercept Related Information such as the URI or IP address and

Check if user agent (SIP PHONE A) initiator is subject of LI. Every SIP INVITE Message is subject to this delay. The default SIP port is 5060.

t_{cs2} = Decryption time of SDP Body

If S/MIME encrypted SDP (application/pkcs7-mime) the SBC A

must decrypt the SDP packet session description information containing the crypto context of the media session, and determine if m=audio *port* RTP/SAVP (media is SRTP)

t_{cs3} = Valid Session Key derivation time

Time to determine master key and salt key and derive session

keys If user agent subject of LI and Key derivation not successful

SBC A sends 4XX CLIENT ERROR MESSAGE

t_{cs4} = INVITE Message Rebuild time

When Successful Session Key derivation obtained or initiator not subject to LI the INVITE message is rebuilt with S/MIME attributes, encrypted with SBC B public key and authenticated with SBC A private key

SBC A forwards rebuilt INVITE to SBC B

SBC A sends 100 TRYING response to SIP PHONE A, with delay of t_{cs3} if initiator not subject to LI.

t_{cs5} = 180 RINGING response delay

This indicates that the INVITE has been received by the user agent SIP PHONE B

This delay will have incorporated into the timing component forwarding and transport delays and the processing time of the remote end SBC B to rebuild the S/MIME portion of the invite message.

t_{cs6} = 200 OK delay

This is the 200 OK response from the user agent SIP PHONE B with the SDP parameters for the media session of SIP PHONE B plus the SBC B and SBC A S/MIME rebuild time with forwarding and transport delays included in the element.

SBC A forwards the S/MIME encrypted SDP body to SIP PHONE A

t_{cs7} = ACK delay SIP PHONE A to SIP PHONE B

The SRTP media session may now be established. SBC A will forward the encrypted packet stream to SIP PHONE B via SBC B. If the user agent is the subject of a LI SBC A will copy the packet stream decrypt and forward to the Mediation Device via the established IPsec tunnel. SBC A will monitor the session and change the session keys when and if required. It is probable that the copying process will not delay the real time packet stream significantly. However the process will utilize SBC A resources limiting the number of LI intercepts it may process.

The session ends when a BYE message is sent from the user agent initiating the termination. The responding user agent sends a 200 OK message in response. SBC A will delete all remnants of the stored LI information for the completed session and terminate the IPSec Tunnel with the Mediation Device if appropriate.

4.3 Real time Media Stream Intercept and Processing

Having obtained the session encryption and the session authentication key the SBC A is ready to authenticate and decrypt the SRTP packets as they arrive. As the media frames arrive they are immediately copied and stored. The original packet is framed and formatted for the outgoing interface of the SBC A usually at line speed.

The copied and stored packet is now available for authentication and decryption. The processing time for the packet is

Incoming SRTP Packet Processing Time :

$$\textit{Processing Time} = t_p \tag{4.5}$$

Where

$$t_p \;\; = t_{rp} + t_{ap} + t_{dp} + t_{sp} \tag{4.6}$$

and;

t_{rp} = Packet Recognition Time

This is the time it takes to recognize the SRTP packet as one
subject to a LI, ie;

Reads the IP source and destination address, udp port and SSRCI. This function is similar to an Intrusion Prevention function where each packet is examined to the port level. In the Lawful Intercept process the examination of the SRTP packet to the SSRCI component would be yadditional time above an

Intrusion Prevention function but would assure that all packets related to a specific real time session.

t_{ap} = Packet Authentication Time

This is proportional to the length of the authenticated portion of the packet and the complexity of the authentication algorithm (SHA-1- HMAC) and inversely proportional to the processing speed.

t_{dp} = Packet Decryption Time

This is proportional to the length of the encrypted portion of the packet and the complexity of the encryption/decryption algorithm (AES) and inversely proportional to the processing speed.

t_{sp} = Decrypted Packet Storage Time

4.4 IPsec Packet to Mediation Device Assembly Time

When constructing a packet for an IPsec tunnel the overhead process makes the construction of a large packet more efficient then a small packet. Therefore rather then process each decrypted incoming packet separately as a tunneled IPsec packet it is more efficient to concatenate as many small packets together to approach the maximum acceptable frame size as the outgoing interface will support. In the case of Ethernet one would want to approach 1500 bytes.

<u>IPsec Packet Assemby Time</u>:

$$IPsec\ Packet\ Assembly\ Time = t_t \tag{4.7}$$

Where

$$t_t = t_{rp} + t_{ap} + t_{dp} + t_{sp} + t_{tea} \tag{4.8}$$

The element t_{tea} is made up of several components. To approach the 1500 byte target one must first subtract out the portions of the IPsec tunnel packet which will not be encrypted. This will include the 20 byte IP header, the 4 byte Security Parameter Index field, the 4 byte Sequence Number field, and the 20 byte Authentication field at the end of the IPsec packet. This gives a total available payload of 1452 bytes to utilize. This 1452 bytes must also include a one byte pad-length field which specifies the number of bytes added to the encrypted data to bring the total to an even 16 byte boundary as a requirement of the Advanced Encryption Standard operating in block mode. There is also a one byte field at the end of the payload specifying the next IP which is a pointer pointing to the beginning of the payload data.

In addition the encapsulated header field of the Mediation device will include a 20 byte IP header and a 20 byte TCP header field. Subtracting this from the 1452 bytes gives a value of 1412 bytes.

With no compression applied to the RTP packet each packet header will have a 40 byte header consisting of 20 bytes, IP header, 8 bytes UDP header, and 12 bytes RTP header. The payload field size for a G.711 codec is 160 bytes, for a total of 200 bytes. This would imply that a total of 7 RTP packets may be included in the Ethernet frame making that assumption would give a total encrypted portion of the IPsec packet of 7 x 200 =1400 bytes plus 40 bytes for the Mediation device IP and TCP header fields. However 1440 bytes is precisely at a 16 byte field boundary. This would not include the 2 bytes required in IPsec for the pad-length and next-IP fields. Therefore the maximum number of RTP frames, without compression, that could be placed in a 1500 byte frame is 6.

The Ethernet payload size for an uncompressed IP/UDP/RTP header then becomes:

(6 x 200 bytes of RTP packets) + (20bytes Mediation device IP header) + (20 bytes Mediation device TCP header) + (6 byte pad for 16 byte boundary) + (2 bytes pad-length and next IP fields) + (4 byte IPsec packet sequence number) + (4 byte IPsec Security Parameter Index) +(20 byte IPsec IP header) + (20 byte IPsec Authentication Data field) = 1296 bytes

To improve efficiency of the IPsec payload and reduce the number of times that an IPsec packet assembly process is initiated the RTP header may be significantly reduced. If the first packet in the assembly is a full packet header the remaining RTP packets may have headers reduced to nominally 5 bytes. The above scenario in assembling an IPsec packet within an Ethernet frame may now be repeated.

The Ethernet payload size for an uncompressed IP/UDP/RTP header then becomes:

(1 x 200 bytes of RTP packets) + (7 x 165 bytes of RTP compressed head) + (20bytes Mediation device IP header) + (20 bytes Mediation device TCP header) + (11 byte pad for 16 byte boundary) + (2 bytes pad-length and next IP fields) + (4 byte IPsec packet sequence number) + (4 byte IPsec Security Parameter Index) +(20 byte IPsec IP header) + (20 byte IPsec Authentication Data field) = 1456 bytes

In an IPsec tunneled packet the payload is encrypted which would include the destination IP and TCP header of the Mediation device. In addition the Payload, Security Parameters Index and Sequence Number field are authenticated.

t_{tea} = IPsec Packet Assembly Time

This time element is proportional to the length of the encrypted portion of the packet and the complexity of the encryption/decryption algorithm (AES). It is also proportional to the authenticated length of the packet and the complexity of the authentication algorithm, HMAC-SHA-1, and inversely proportional to the processing speed.

Figure 9 illustrates the SIP signaling timing functions discussed above. Figure 9 also shows the timing elements in the real time processing of the VoIP media stream as presented in the timing model.

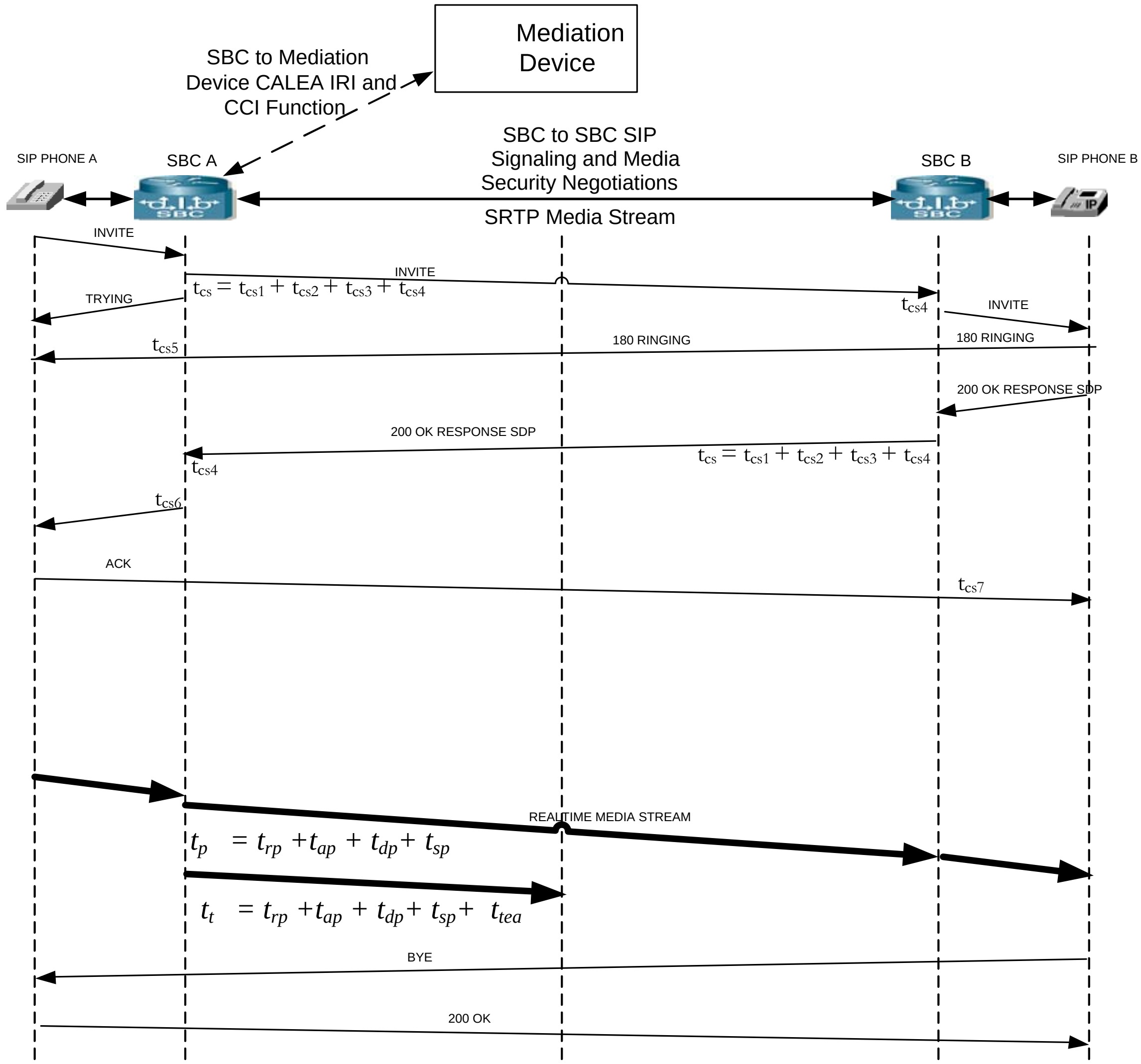
Mediation Device
SBC to Mediation Device CALEA IRI and CCI Function
SBC to SBC SIP Signaling and Media Security Negotiations
SRTP Media Stream
SIP PHONE A
SBC A
SBC B
SIP PHONE B
INVITE
INVITE
$t_{cs} = t_{cs1} + t_{cs2} + t_{cs3} + t_{cs4}$
TRYING
t_{cs4}
INVITE
t_{cs5}
180 RINGING
180 RINGING
200 OK RESPONSE SDP
200 OK RESPONSE SDP
$t_{cs} = t_{cs1} + t_{cs2} + t_{cs3} + t_{cs4}$
t_{cs4}
t_{cs6}
ACK
t_{cs7}
REALTIME MEDIA STREAM
$t_p = t_{rp} + t_{ap} + t_{dp} + t_{sp}$
$t_t = t_{rp} + t_{ap} + t_{dp} + t_{sp} + t_{tea}$
BYE
200 OK

Figure 9. SIP Signaling and Real Time Processing Model Timing Elements

4.5 Total Time Slot Processing Time

In a real time application such as VoIP the packets carrying the data must arrive at specific repetitive times to maintain the quality of the VoIP conversation. In the case of the codec G.711 mentioned earlier the packets must arrive every 20 msec or 50 packets per second to maintain the required codec bit rate (160 byte payload x 8 bits/byte x 50 packets/s = 64,000 bits/s).

Assuming that all the active VoIP conversations are using G.711 codecs for analog to digital conversion the 20ms time slots could be filled with a continuous stream of RTP packets. Within this 20 ms time slot the total processing time for the packets in the time slot would be modeled based on the timing elements presented in the above discussion. In each 20 ms time slot two kinds of processes are occurring for the packets that are the subject of a Lawful Intercept. The majority of LI packets are subject to the process according to equation 4.6 while a smaller number are undergoing the IPsec packet assembly process described by equation 4.8. If there are M number of packets undergoing the standard incoming SRTP process and N number of packets undergoing the IPsec Packet Assembly process. Then the total time for each process is given by the following equations:

Total Time to process M packets in standard incoming SRTP mode :

$$\textit{Total Processing Time} = t_{tp} \tag{4.9}$$

Where

$$t_{tp} = \sum_{i}^{M} (t_{rp} + t_{ap} + t_{dp} + t_{sp})_i \tag{4.10}$$

Total Time to assemble N packets for the IPsec tunnel:

$$\textit{Total Processing Time} = t_{tt} \tag{4.11}$$

Where

$$t_{tt} = \sum_{i}^{N} (t_{rp} + t_{ap} + t_{dp} + t_{sp} + t_{tea})_i \tag{4.12}$$

The total processing time occurring within the 20 msec time slot is:

$$\textit{Total Processing Time} = t_{ts} \tag{4.13}$$

Where

$$t_{ts} \quad = t_{tp} + t_{tt} \tag{4.14}$$

As the number of Lawful Intercept VoIP conversations increase the total processing time will increase. At some point the process will begin to fail. Either the CPU utilization will reach 100% and new incoming packets will be dropped or the total processing time t_{ts} will exceed the 20 msec window and following packets will begin to be stored until they can be processed. It is this second failure mode that is examined in more detail in the following sections.

4.6 **SIP Signaling Process Timing Elements**

As mentioned in the previous paragraphs the IPsec tunnel setup to the Mediation Device and the SIP Signaling processes are functions that need not occur in real time but may be performed in the background with nominal proficiency to prevent the application from timing out. The SIP Signaling process may be collocated in the same SBC or may be in a separate device with a defined protocol link between the SBC performing the real time processes and the signaling processes.

The following values may be used for the IPsec tunnel setup and the SIP signaling process:

t_{ke} = 30 milliseconds

t_{Lin} = 2 μ seconds

t_{cs1} = 2 μ seconds

t_{cs2} = 0.2 seconds

t_{cs3} = 50 μ seconds

t_{cs4} = 0.41 seconds

t_{cs5} = 80 milliseconds

t_{cs6} = 2.2 seconds

t_{cs6} = 40 milliseconds

4.7 Total Time Slot Processing Time Simulation

A program was written to determine when the total process times described by equation 4.14 would cause a failure and overwhelm the resources of the SBC. Assuming that the CPU is not overwhelmed first, this failure of total process time should be determined when there is insufficient time within the 20 ms time slots to process any packets and all arriving packets that are subject to a Lawful Intercept are stored. This should be observable as a constant increase in bytes stored as time progresses. Two Session Border Controllers were modeled with the following timing parameters shown in table 1.

TABLE 1

Session Border Controller Timing Parameters

Parameter	**SBC 1**	**SBC 2**
t_{rp}	**0.38 µs**	**0.23 µs**
t_{ap}	**13.4 µs**	**4.47 µs**
t_{dp}	**13.5 µs**	**4.5 µs**
t_{sp}	**1 µs**	**0.3 µs**
t_{tea}	**176.5 µs**	**64.7 µs**

SBC 1 simulation model was run with two different interfaces, first a fast Ethernet (100Mbs) for the incoming and outgoing packets. Both SBC 1 and SBC 2 were then run with Gigabit Ethernet interfaces (1Gbs) for the incoming and outgoing interfaces. Also each simulation was tested with the IPsec tunnel packet constructed with 6 uncompressed header RTP packets and also with 8 compressed header packets.

The simulated 20ms time slot windows were filled with 1s randomly over a period of time. The number of available positions within the 20 ms time slot was determined by the interface bandwidth and the length of the Ethernet frame. The length of an Ethernet frame in bytes is described below:

- 7 octets preamble
- 1 octet starting delimiter
- 6 bytes destination mac address
- 6 bytes source mac address
- 2 bytes Ethernet type
- 204 bytes payload (Uncompressed IP/UDP/RTP + 160 bytes data)
- 4 bytes CRC
- Interframe gap (0.96 µs for 100Mbs and 0.096 µs for Gbps)

These values give an Ethernet frame length of 19.36 µs for 100Mbs and 1.936 µs for Gigabit Ethernet. Dividing each frame length into 20ms gives a value of 1033 locations within a 20ms time slot and 10,331 for Gigabit Ethernet. The 1s represented a lawful intercept packet while 0 represented an empty location or a location that was not a lawful intercept packet. Thus a string of 0s would represent a longer Ethernet packet such as one carrying signaling information. In the random fill process the simulation counted the number of ones which represented intercepted SRTP VoIP media streams. The simulation also counted the number of IPsec process packets in the frame by counting the number of times a slot would have a one in it and then resetting the IPsec process counter to 0 when the count reached a 6 or 8. The simulation would store a packet when it was still operating on a previous process, either normal SRTP receiving or an IPsec packet assembly process. When it finished that process it would decrement the store counter and continue to the next process. The total number of stored bytes was accrued in a dynamic function varying in an up and down manner as it approached the failure point and began a dramatic constant increase. Conceptually all packets arriving at the inbound SBC interface would be forwarded out the outbound SBC interface. Packets not subject to a Lawful Intercept would not go through the modeled processes. The following figures show the simulation results with some comments.

4.8 Simulation Results

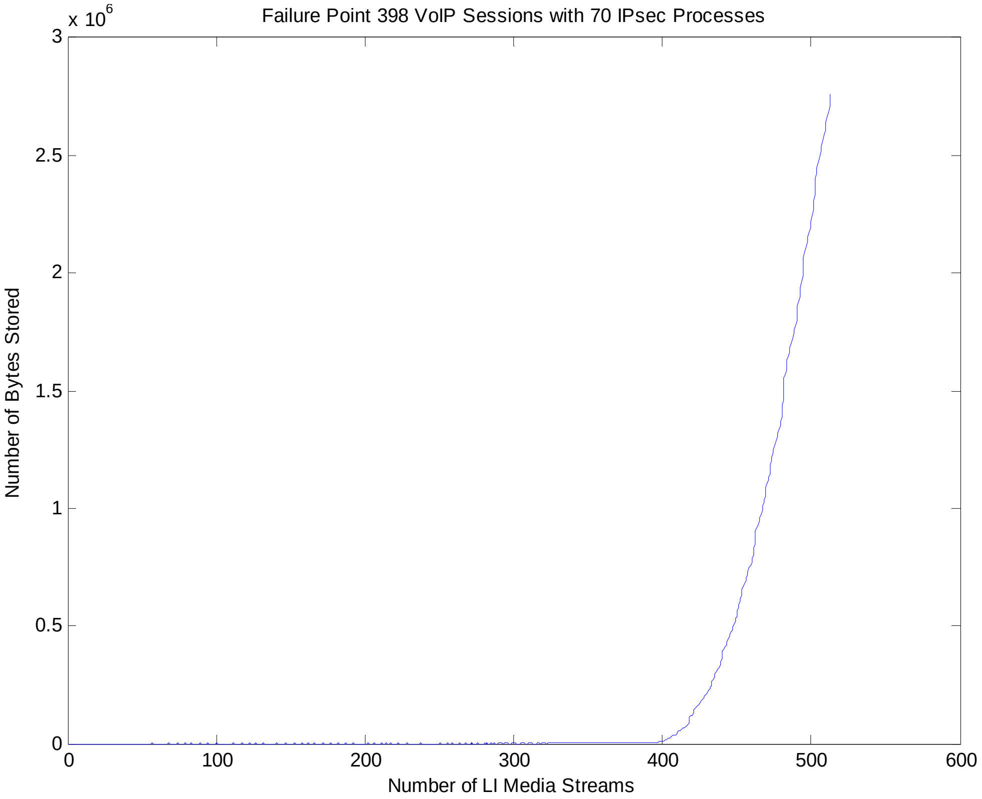

Figure 10. SBC 1 100Mbs Interfaces 6 Count IPsec Process

In the first simulation described by figure 10 398 VoIP sessions is nearly 39% of the available bandwidth which would make it difficult for large packets to be added into the time slot. The IPsec packet assembly process collect 6 RTP packets and concatenated them into a tunneled IPsec packet.

In figure 11 the simulation was modified to process an 8 count IPsec packet assembly. Note again that 427 VoIP Sessions occupy 41 percent of the available locations in the 20ms time slot. This would make it difficult on the outgoing interface to add any large packets or have a significant number of packets that were not part of a Lawful Intercept. In order to present a more realistic scenario the inbound and outbound interfaces were changed to Gigabit Ethernet. This would provide 10 times the number of packet locations in a 20ms frame.

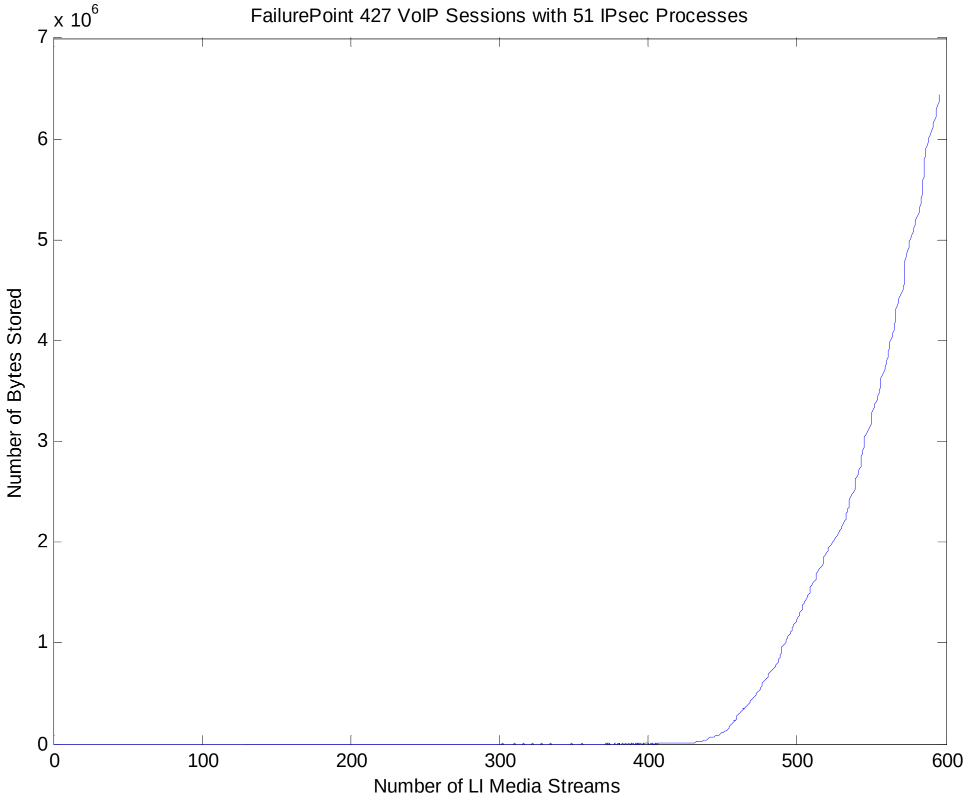

Figure 11. SBC 1 100Mbs Interfaces 8 Count IPsec Process

.

In the results of the simulation shown in figure 12, having Gigabit Ethernet interfaces, the intercept VoIP session count of 776 is much more reasonable in terms of total available time slot locations of 10,331. However, the CPU on SBC 1 has probably been at the 100%

utilization point before reaching the 776 session limit. Note also that the plot line has a great many rapid variations. This would be indicative of the number of stored bytes varying rapidly because prior to a dramatic memory consumption increase there are many more packet locations not subject to a Lawful Intercept then in the previous 100Mbit/s interface simulation. This larger number of 0 time slot locations gives the current process opportunity to be completed and then move to a stored packet to begin processing thus decrementing the stored bytes counter.

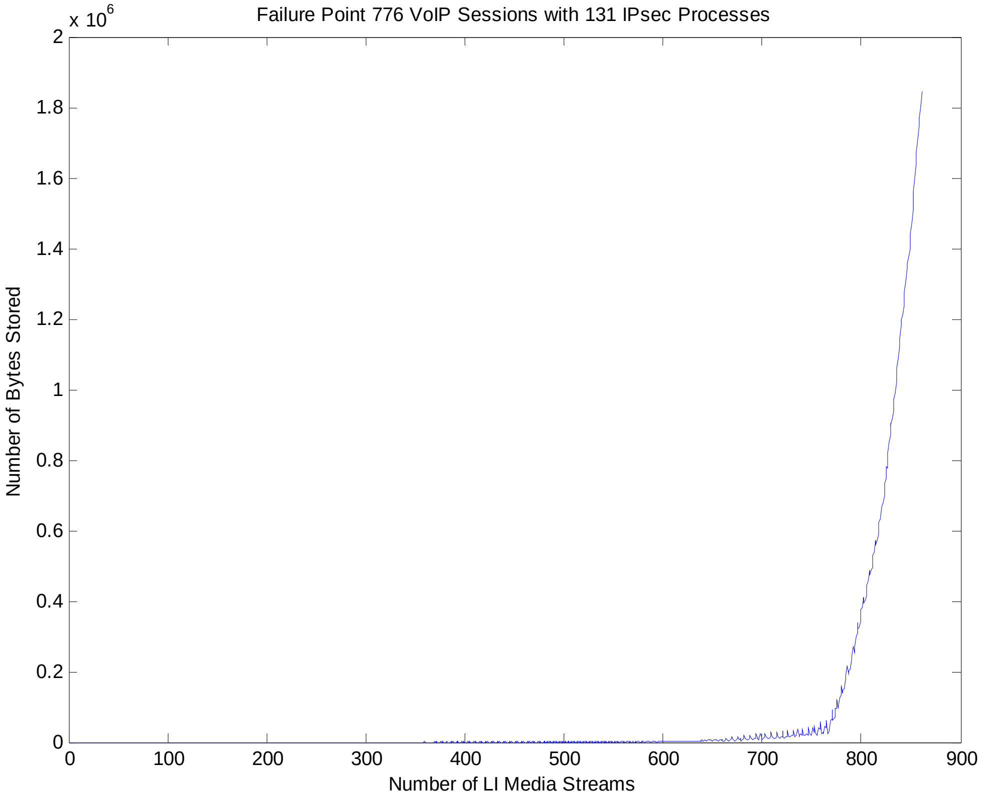

Figure 12. SBC 1 with GE Interfaces 6 Count IPsec Process

Figure 13 shows the results of the simulation with an 8 count IPsec assembly process. The increase in VoIP sessions before dramatic failure is clearly visible when comparing the results in figure 12 and 13. This would be expected as it takes running only three 8 count IPsec processes to be equivalent to four 6 count IPsec processes.

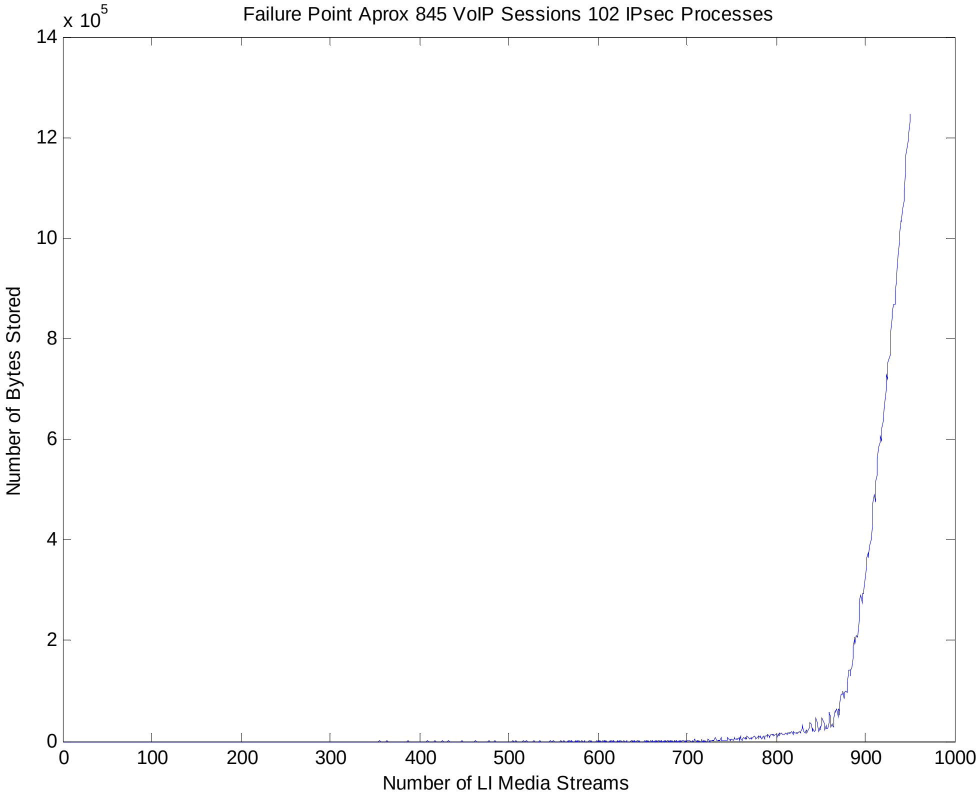

Figure 13. SBC 1 with GE Interfaces 8 Count IPsec Process

Figures 14 and 15 demonstrate that the faster process times of SBC 2 increase the number of VoIP sessions before the process failure point is reached, indicated by a dramatic increase in number of bytes stored. As with the previous simulations a 6 and 8 count IPsec packet assembly process was simulated with the expected increase in the total number of VoIP sessions before the process began to fail for the 8 count IPsec packet assembly process.

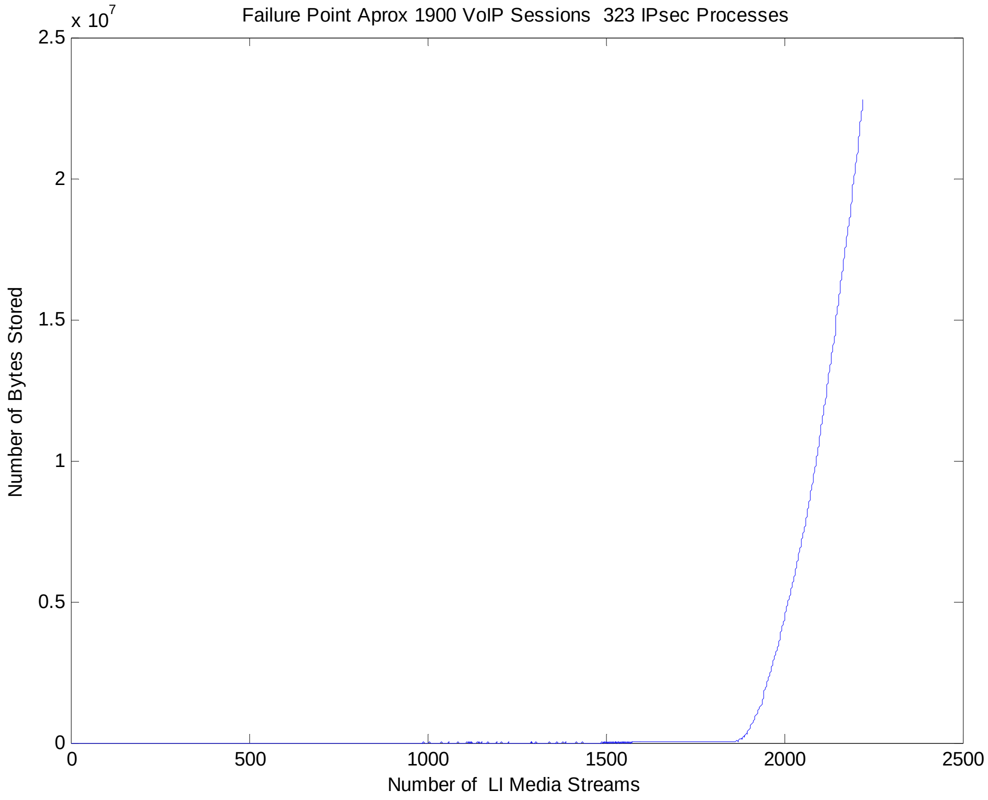

Figure 14. SBC 2 with GE Interfaces 6 Count IPsec Process

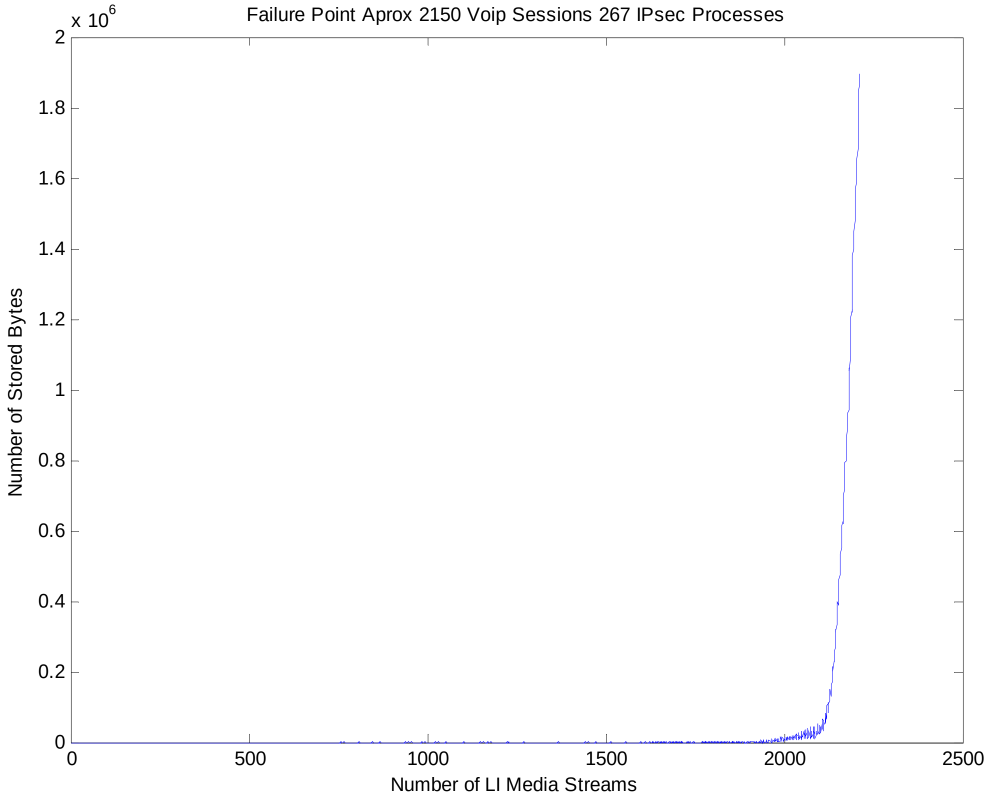

Figure 15. SBC 2 with GE Interfaces 8 Count IPsec Process

CHAPTER 5

Conclusions and Future Work

The results of the modeling simulation have demonstrated that the use of a Session Border Controller as an intermediate device in a Lawful Intercept process is a reasonable approach to providing the desired security by a user agent and the accessibility of Call Content or Intercept Related Information to Law Enforcement Agencies. The intermediate functions of the Session Border Controller in particular overcome the challenge of a user agent which desires a high level of security such as provided by S/MIME or other protocols for the signaling and Secure Real Time Transport Protocol media sessions for legitimate reasons and the need for Law Enforcement Agencies to be able to obtain decrypted information in a timely manner.

The simulation has demonstrated that the number of Lawful Intercept VoIP sessions that may be processed is significantly large before the process starts to fail by consuming large amounts of memory storage. In reality the number of Lawful Intercepts processed on a single Session Border Controller should be quite small.

Significant future work needs to be undertaken to obtain performance in a Session Border Controller configured to perform the signaling and simulated real time intercept of a VoIP media stream. Other protocols such as Multimedia Internet Keying (MIKEY) need to be added to the capabilities of the SBC and tested. Also how much of the SIP Proxy server functions need to be incorporated into the SBC functionality need to be investigated.

These complex security protocols, ie; AES, S/MIME, etc. were not a consideration that standard VoIP phones had the processing capability to achieve in the past. However, soft phones are being advertised currently with these capabilities. With the increase in processing capability of VoIP phones growing rapidly these capabilities will not be far behind in a standard VoIP phone. These products make this line of research a viable direction of development.

REFERENCES

List of References

[1] S. Landau, "Security, wiretapping, and the Internet," *IEEE Security & Privacy Magazine*, vol. 3, issue 6, pp. 26-33, Nov-Dec 2005.

[2] A. Gidari and P. Cole, "Designing the right wiretap solution: setting standards under CALEA," *IEEE Security & Privacy Magazine*, vol. 4, issue 3, pp. 29-36, May-June 2006.

[3] G. Goth, "Internet Wiretapping Decision Headed for the Courts," *IEEE Internet Computing Magazine*, vol. 10, issue 1, pp. 8-10, Jan-Feb 2006.

[4] H. Alvestrand et al., "IETF Policy on Wiretapping," RFC 2804, May 2000

[5] A. Milanovic et al, "Methods for lawful interception in IP telephony networks based on H.323," *The IEEE Region 8 EUROCON 2003,Computer as a Tool*, vol. 1, 22-24, pp. 29-36, Sept. 2003

[6] A. Milanovic et al, "Distributed system for lawful interception in VoIP networks," *The IEEE Region 8 EUROCON 2003,Computer as a Tool*, vol. 1, 22-24, pp. 203-207, Sept. 2003

[7] E. Miller, F. Andreasen, G. Russel, "The Emergence of Integrated Broadband Cable Networks The Packet Cable Architecture ," *IEEE Communicatins Magazine*, vol. 39, issue 6, pp. 90-96, June 2001.

[8] Cisco Systems, Inc, "PacketCable Lawful Intercept Architecture for BTS Versions 4.4 and 4.5," ver. 2, March 15, 2006

[9] Cisco Systems, Inc, "Cisco Service Independent Intercept Architecture Version 2.0," ver. 2, March 15, 2006

[10] B. Marshall et al, "PacketCable™ Electronic Surveillance SpecificationPKT-SP-ESP-I02-030815, Second Issued Release, August 15, 2003

[11] F. Baker et al., "Cisco Architecture for Lawful Intercept in IP Networks," RFC 3924, Oct. 2004

[12] J. Rosenberg et al., "SIP: Session Initiation Protocol," RFC 3261, June 2002

[13] J. Franks et al., "HTTP Authentication: Basic and Digest Access Authentication," RFC 2617, June 1999

[14] B. Ramsdell, "Secure/Multipurpose Internet Mail Extensions (S/MIME) Version 3.1 Message Specification," RFC 3851, July 2004

[15] J. Peterson, "S/MIME Advanced Encryption Standard (AES) Requirement for the Session Initiation Protocol (SIP)," RFC 3853, July 2004

[16] C. Allen and T. Dierks, "The TLS Protocol Version 1.0," RFC 2246, Jan. 1999

[17] H. Schulzrinne et al, "RTP: A Transport Protocol for Real-Time Applications," RFC 3550, July 2003.

[18] A. Johnston and D. Piscitello, *Understanding Voice over IP Security,* Norwood, MA Artech House, 2006

[19] M. Baugher et al, "The Secure Real-time Transport Protocol (SRTP)," RFC 3711, March 2004.

[20] J. Arkko et al, "MIKEY: Multimedia Internet KEYing, " RFC 3830, August 2004.

[21] F. Andreasen et al, "Session Description Protocol (SDP) Security Descriptions for Media Streams," RFC 4568, July 2006.

[22] A. Johnston, *Understanding the Session Initiation Protocol second edition,* Norwood, MA Artech House, 2004

[23] USA Patriot Act, H.R. 3162, One Hundred Seventh Congress of the United States of America, October 2001

[24] F. Cao, S. Malik, "Vulnerability Analysis and Best Practices for Adopting IP Telephony in Critical Infrastructure Sectors ," *IEEE Communicatins Magazine*, vol. 44, issue 4, pp. 138 - 145, April 2006.

[25] V. Prevelakis, D. Spinellis, "The Athens Affair," *IEEE Spectrum Magazine*, vol. 44, number 7, pp. 26 - 33, July 2007.

[26] C. Goodrich, "CALEA Compliant Voice Over IP", *Unpublished Thesis* Wichita State University , May 2005.