

UNIVERGE VoIP Security Best Practice

Vol. I

Principles of VoIP Security **(Version: 1.0)**

NEC

NEC Corporation

Liability Disclaimer

NEC Corporation reserves the right to change the specifications, functions, or features, at any time, without notice.

NEC Corporation has prepared this document for the exclusive use of its employees and customers. The information contained herein is the property of NEC Corporation and shall not be reproduced without prior written approval from NEC Corporation.

UNIVERGE is a registered trademark of NEC Corporation.

Some of the NEC products identified in this document may not be available in certain regional markets. Please contact your NEC representative for availability.

© 2005 NEC Corporation

MS-DOS, Microsoft, Windows and Windows NT are registered trademarks of Microsoft Corporation.

All other brand or product names are or may be trademarks or registered trademarks of, and are used to identify products or services of, their respective owners.

Contents

1. Introduction	1
1.1 Overview.....	1
1.2 Audience	2
1.3 Author 2	
2. Threats on IP Telephony Systems and their Sources	3
2.1 Major Threats to IP Telephony Systems	3
2.2 Potentially Vulnerable Components	4
2.3 Sources of Vulnerabilities	5
3. Design Fundamentals.....	7
4. Axioms for Secure IP Telephony.....	8
4.1 Develop Appropriate Network Architecture [I].....	8
4.2 Check Acceptability of Risk [I].....	9
4.3 Internet Node Necessity for IP Telephony Devices [O]	9
4.4 Patch Management is Mandatory [O]	9
4.5 Confidentiality of Phone’s Credential [A].....	10
4.6 Confidentiality of Calls and Voice [I]	11
4.7 Firewall and Other Protection Mechanisms [I].....	11
4.8 Wireless LAN [I]	12
4.9 Attention to Soft-Phones or PC-based IP Phones [I].....	13
4.10Attention to Application Servers for Voice [I]	14
4.11Physical Security around VoIP System [O].....	14
4.12Power Blackout Consideration [O].....	14
4.13Review Statutory Requirements with Legal Advisors [L]	14
4.14E-911 Consideration [L--US]	15
4.15Vulnerability of SIP and IP specification [O]	15
5. IP Telephony and Encryption	16
5.1 Fundamentals of Encryption.....	16
5.2 Targets of Encryption	17
5.3 Encryption in UNIVERGE IP Telephony Solution	18
6. References	19

6.1 RFCs and Drafts.....	19
6.2 Notes and Papers from Vendor Independent Parties.....	19
6.3 White Papers from Vendor	20
6.4 Miscellaneous References	20
7. Acknowledgments	21
8. Glossary of Network Security Terms.....	22

1. Introduction

1.1 Overview

Network security represents an apex of concern for every organization these days. Regulations are both vastly increasing coming to pass in most regions. Security breaches may damage reputations and loss of business opportunities; and, while the IP telephony solutions can produce a new style of office communication and reduce network costs, it adds complexity onto development and maintenance. Corporate networks are vastly impacted due to the unique network nature of IP telephony systems and the coexistence of data traffic and voice traffic. The purpose of the “UNIVERGE VoIP Security Best Practices” series is to illustrate basic guidance for secure deployment and maintenance of UNIVERGE telephony systems.

This document is Volume I of a series of security blueprints for designing and implementing secure IP telephony systems. Volume I provides an overview and outlines general principles of overall security design for IP telephony systems.

Within the framework of VoIP system, NEC uses a defense-in-depth approach to network security design, which serves as a guide to network designers considering the security requirements of their respective networks. This particular design focuses on expected threats and their methods of mitigation, resulting in a layered approach to security where the failure of one security system is less inclined to compromise of the rest of the network. Although this document is product-agnostic, its proof-of-concept is based on products from NEC and its partners.

This document focuses on threats encountered in enterprise environments. Network designers who understand these threats can better decide where and how to deploy mitigation technologies. Without this understanding, deployments tend to be incorrectly configured, too focused on security devices, or lacking in threat response options. By taking the threat mitigation approach, this document provides network designers with information for making sound network security choices.

1.2 Audience

The UNIVERGE VoIP Security Best Practices series is intended for network and system managers. Although this document is essentially technical, it can be read without understanding network and system details.

This document is composed of volumes intended to provide proper information in proportion to your purpose. If you would like to understand the security overview, please refer to both Volume I and Volume II. If you are interested in integrating a secure VoIP system, refer to both Volume II and Volume III.

Since comprehensive security for a corporate network includes too many aspects to cover in this series, we focus on basic issues tailored to IP telephony systems. For example, we presume that your organization already maintains a security policy. NEC does not recommend deploying any security technology and device without first establishing the security policy.

1.3 Author

Toshio Miyachi is the primary author of this white paper.

2. Threats on IP Telephony Systems and their Sources

IP telephony systems must ensure the same dependability as traditional telephony systems, especially in business and services environments. This assurance poses a great technical challenge for several reasons. This chapter describes major threats to IP telephony systems and explores major sources of such threats. In effect, this chapter will categorize some key points which IP telephony system designers must consider.

2.1 Major Threats to IP Telephony Systems

(1) Unauthorized access to IP Telephony Systems

Unauthorized access to an IP telephony system can be tried from local and remote networks.

(2) Interception/Eavesdropping

The attacker might situate between the two end points of a communications link intending to monitor, record, or even manipulate the data stream of IP telephony. In some cases, the attacker may usurp control of the link, or intercept all streaming data by secretly breaking into the communication path. This form of interception is also known as the “man in the middle attack.”

(3) Fraud

Toll fraud is unauthorized access that uses resources without paying for them.

(4) Denial of service

Denial of Service covers actions and events that prevent systems from providing the agreed-upon levels of service to authorized users. A “Load-based DoS” involves bombarding a server with millions of requests. A “Malformed Request DoS” is a sophisticated protocol request that exploits a vulnerable area. Both attacks impact the availability of resources and could lead to degraded Quality of Service (QoS).

(5) Manipulation

Manipulation involves the unauthorized modification of information (including program code), typically caused by computer viruses and worms.

(6) Protocol attack

A protocol attack exploits vulnerabilities within VoIP protocols, such as SIP or H.323.

(7) SPIT

Spam (unsolicited messages) for IP telephony, (SPIT) is a nuisance to voice mail users. It may, block resources, resulting in a lower Quality of Service.

2.2 Potentially Vulnerable Components

Every area of an IP Telephony system might be the target of attack, if any vulnerability exists.

The first potentially vulnerable component is the communication platform since its IP PBX core plays a critical business role. The second component is the IP network infrastructure. Other components of IP telephony systems include the application platform, clients, the IP network infrastructure, and the management system (see Figure 1).

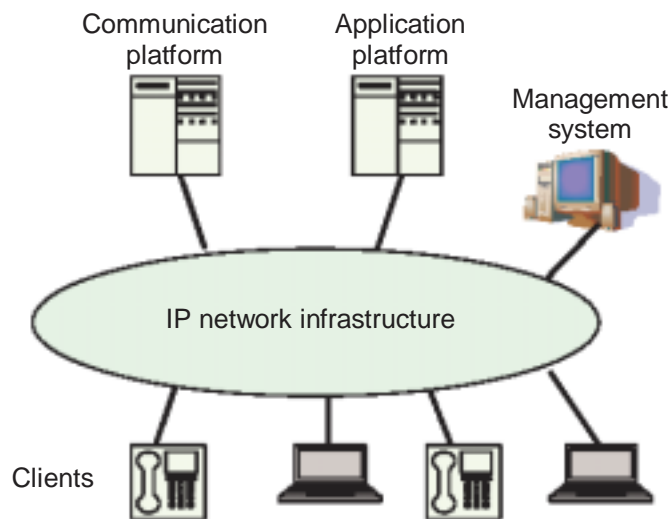


Figure 2-1 IP Telephony System

(1) Communication platform

An operating system is the main target for attackers, since it is common with servers and clients for data services. An attack typically originates on regular computer servers (e.g.: the initial target may filter through the data side. Other attacks result in DoS and protocol attacks or attacks via remote access (often at the administrator level). In addition, access points and gateways are vulnerable, and the remedy is similar to that of communication platforms.

(2) Application platform and Applications

Critical vulnerabilities of real-time applications include voice mailbox interception, service abuse through registration hijacking or toll fraud, and manipulation of statistic or accounting data.

(3) Clients

IP hard-phones and soft-phones should be hardened to make sure that data, voice streams, and authentication cannot be accessed or altered by unauthorized parties. The obvious objective is to retain baseline telephony in the event of an attack. Communications protocols must be secured by encryption. For IP phones, authentication on layer 2 using IEEE 802.1x is highly recommended.

(4) Management system

Local and remote administration access (including logon) must provide strong authentication and traffic must be encrypted. Software distribution modules must be carefully protected against fraud and manipulation.

(5) IP network infrastructure

Network devices such as routers and switches have to be protected as the base line of VoIP security.

2.3 Sources of Vulnerabilities

(1) Open Environment Relatively Easy to Attack [O]

IP telephony systems typically work in an open network with a lot of threats, while traditional telephony systems run in a rather closed and telephone-only network. In an open network, specification of most communication protocols is open in public and a number of networks of various security levels are mutually interconnected into a logically single IP network. Malicious users gather enough knowledge to conduct wiretapping and its use through Internet searches for protocol specifications.

In particular, IP telephony signaling protocol SIP (Session Initiation Protocol) packets are easily analyzed to procure an IP telephone's credentials and secretly gather non-encrypted calling records (because SIP is a text-based protocol). The IP network on which IP telephony systems are based may provide an opportunity for malicious users to break down the IP telephony services. We should evaluate these threats and decide if some mitigation measures are needed.

(2) Real-time Is Essential for Voice [R]

The traffic for IP telephony systems is real-time in nature. The voice call is more severely affected by packet loss, processing delay, and jitter than ordinary data network applications. Extraordinarily heavy load, which are inappropriately generated by malicious entities such as computer virus, may cause sound noise, interruption, or even dropout of some calls.

(3) Authentication of Devices [A]

Within a traditional telephony system, each phone is identified by the line interface of a wire by which the phone is connected to a PBX. Within an IP telephony system, each phone is identified and authenticated by an identity and credential stored in itself. That said, device authentication is highly important and the credentials embedded within each device, including a phone, must be securely protected.

(4) Data and Voice Integration [I]

Ordinary data traffic contains vastly different characteristic from IP telephony traffic; therefore, it is desirable to separate the two, if possible, to maintain stable service quality. However, such network separation is often not practical due to computers with soft phones or the general pursuit of cost merit from data and voice integration.

Data networks usually include some firewalls and other network gateway nodes for protecting an internal network or separating it into multiple, logical sub-networks. When an IP telephony system is integrated within a data network, some gateway configurations might be modified in order to move IP telephony-related traffic through them. Meanwhile, the modification must not lower the security level of the data network or introduce any threats to the data network.

Communication among IP telephony system components is carried symmetrically or in a peer-to-peer fashion; though most applications on data networks function asymmetrically, based upon the client server model. Due to this difference integration of voice with a data network, either special IP telephony gateways or network architecture redesign is required prior to data and voice integration.

3. Design Fundamentals

The security consists of confidentiality (C), integrity (I) and availability (A), referred to as “CIA”. That is, secure IP telephony systems should retain system confidentiality, user information, and phone dialog. It should also realize end-user and system administrator intentions, whenever a phone call is requested, to prevent malicious activities conducted within the networks. A designer of IP telephony systems understands numerous threats from device failures to malicious activities and prepares some measures to mitigate them unless a user accepts corresponding risks.

Some user organizations establish their security governance according to the framework of [ISO/IEC 17799 \(Code of Practice for Information Security Management\)](#). IP telephony system designers should understand it and support user organizations to implement it. You may give an account of the following sections of the international standard:

Chapter 6	Personnel Security
Chapter 7	Physical and Environmental Security
Chapter 8	Communications and Operations Management
Chapter 9	Access Control
Chapter 10	Systems Development and Maintenance

In the U.S. market, we should take into account the [NIST Special Publication 800-53 \(Recommended Security Controls for Federal Information Systems\)](#) as well.

4. Axioms for Secure IP Telephony

This chapter represents overarching design considerations for secure IP telephony systems. These general principles should be implemented in every installation of IP telephony systems.

4.1 Develop Appropriate Network Architecture [I]

Establishing a secure network for VoIP and data is a complex process that requires a much greater effort than is required for data-only networks, due to the integration of voice and data into a single network. IP-based telephony provides telephony over the existing IP data networks. However, for reasons including QoS, scalability, manageability, and security, deployment of IP telephony devices and IP data devices should occur on two logically disparate segments. Segmenting IP voice from the traditional IP data network greatly increases attack mitigation capabilities and allows use of the same access, core, and distribution layers. Although the segments should be separated, deployment of two physical IP infrastructures is not recommended. Technologies such as virtual LANs (VLANs), access control, and a stateful inspection firewall, provide the Layer 3 segmentation necessary to keep the voice and data segments separate at the access layer. A stateful firewall, which is deployed at specific locations in the network where the segments are allowed to interact, is also useful for providing host-based DoS protection against connection starvation and fragmentation attacks, dynamic per-port-granular access, spoof mitigation, and general filtering. For each customer, we should develop appropriate network architecture and ensure the following points, even if IP telephony is introduced into an existing IP data network:

- i) Identify where voice and data share the same network from both physical and logical viewpoints.
- ii) Ensure the security of the voice gateway system to other VoIP systems and PSTN.
- iii) Ensure that voice crosses smoothly and securely over boundaries such as a firewall and NAT of data sub-networks.

This axiom corresponds to the NIST Recommendation 1 in SP 800-58.

4.2 Check Acceptability of Risk [I]

We should explain the risk introduced into network environments containing a VoIP system to our customer and examine their acceptability of the risk. When their network demands highly stable and continuous operation and the risk with VoIP system deployment is not acceptable, we have to give up the integration of the network and voice. This axiom corresponds to the NIST Recommendation 2 in SP 800-58.

4.3 Internet Node Necessity for IP Telephony Devices [O]

All IP telephony system devices work as nodes of the Internet, so they are potential targets of network attacks, including computer viruses and network worms, which vastly differ from the case of traditional non-IP PBXs. For example, a maintenance login account with a well-known ID and password is very dangerous. An operation log should be gathered and retained for an appropriate time period so that you can proactively locate abnormal behavior and analyze an entry point and game trick of an attacker. You can learn further details in the following document:
IETF: "Site Security Handbook", RFC 2196, <http://www.ietf.org/rfc/rfc2196.txt>

4.4 Patch Management is Mandatory [O]

Since VoIP system software is very complex, we must assume that it may include some, possibly many, unknown vulnerabilities. Especially, on Linux OS and Windows OS, which are included in some of VoIP devices, almost regularly quite a number of vulnerabilities are reported. Some of them are potential attack vectors or make up an infection path of computer viruses and network worms. If they are the case, it is serious risk for the user to leave it with neither workarounds for mitigation nor fundamental fix measures. VoIP users should be strongly encouraged to buy maintenance service so that they can get information and patches for vulnerabilities as soon as possible.

4.5 Confidentiality of Phone's Credential [A]

While a PBX in a conventional switching system identifies each phone by the cable line through which it is connected to the PBX, IP-PBX identifies each IP-phone with a SIP ID and its credential. If a SIP ID and credential of an IP-phone, say X, are leaked, someone can configure another phone as if it is X, and disturb normal operation of X. So we have to maintain the confidentiality of pairs of SIP IDs and credentials.

We can mitigate many attacks against the IP telephony network if users and devices are authenticated properly. The primary method for the device authentication of IP phones is the MAC address. If a phone with an unknown MAC address attempts to access the IP-PBX, and it has no knowledge of the IP phone's MAC address, then automatic registration of that IP phone should be disabled. However, this is hardly a complete solution against attacks, since MAC address spoofing is rather easy.

Enabling the encryption mechanism is recommended to keep confidentiality of phone's credentials, if supported. Since in the key exchange process for the encryption strong mutual authentication is implemented, the risk of phone's credential leak and impersonation is removed almost completely.

4.6 Confidentiality of Calls and Voice [I]

The combination of data and voice segmentation and a switched infrastructure makes call eavesdropping attacks dramatically easy. For instance, there is a tool known as “vomit” (Voice Over Misconfigured Internet Telephones), which converts a Cisco IP phone conversation into a wave file playable with ordinary sound players, even if the phones are not actually misconfigured. Rather, if someone obtains access to the IP data stream at any point in the network, they could eavesdrop. It is the case for not only Cisco IP phones but for all VoIP systems unless the encryption feature is enabled.

To a degree, keeping the segments separate thwarts devices within the data segment from listening to calls within the voice segment. An obvious way around this segmentation is to unplug an IP phone in the voice segment and plug in another device such as a workstation. By definition, using a switched infrastructure should thwart a device even in the same segment from call monitoring. However, tools such as “dsniff” effectively turn the switched medium into a shared medium. Thus, segmentation provides minimal attack mitigation on its own. The true value of segmentation is the ability to tune network intrusion detection systems (NIDSs) as outlined in the secure and monitor all voice servers and segments axiom.

One additional attack should also be noted. If the hacker has access to the local switched segment, the hacker could insert a phone into the voice segment with a spoofed Media Access Control (MAC) address, assume the target phone’s identity, and intercept a call. Mitigation techniques for this attack are discussed in the following rogue devices axiom.

Enabling the encryption mechanism is recommended to maintain the confidentiality of calls and voice, if supported.

4.7 Firewall and Other Protection Mechanisms [I]

While most communication over traditional data network is based on server-client model, VoIP communication model is peer-to-peer in terms of both signaling and media transport. On top of that, there exist some weak points of the SIP protocol specification, which might be exploited to tear down established call connection by malicious attackers. So some protection mechanisms specifically for a VoIP system should be implemented in order to mitigate attacks.

The peer-to-peer nature of VoIP systems impacts to firewall implementation. You should understand VoIP related features of firewalls and consider how to take advantage of firewalls.

VoIP-ready firewalls and other protection mechanisms should be deployed and made enable to mitigate the inherent vulnerabilities. This axiom comes from NIST Recommendation 6 in SP 800-58.

4.8 Wireless LAN [I]

When mobile units are to be integrated with the VoIP system, you should use products implementing WPA (WiFi Protected Access) rather than WEP (Wired Equivalent Privacy). WEP2 or WPA (WiFi Protected Access) requires TKIP (Temporal Key Integrity Protocol) and combines the temporal key with the client's MAC (Media Access Control) address (Ethernet NIC-Network Interface Card fixed 48-bit address) and then adds a relatively large 16-octet (128-bit) initialization vector to produce the key that will encrypt the data. TKIP ensures that each station uses different key streams to encrypt the data. This axiom comes from NIST Recommendation 8 of SP 800-58.

4.9 Attention to Soft-Phones or PC-based IP Phones [I]

In comparison, soft-phone hosts are more susceptible to attacks due to the number of vectors within the system. These include Operating System (OS), application, and service vulnerabilities, worms, and viruses. IP phones run on custom OSs with limited service support and are less likely to have these vulnerabilities. In addition, a soft-phone resides within the data segment and presents susceptibility to any attack against that entire segment, not just the host itself. The Code-Red and Nimda worms/viruses, for instance, bogged down soft-phone user systems and the segments they resided in that they were rendered unusable. No amount of QoS will prioritize voice traffic over data traffic in the data segment if the end system placing the call is unusable.

Many soft-phones support a data port, allowing PC to phone connection with only a single cable, providing data and voice connectivity to the user's workspace. In this case, follow the data/voice segmentation principle. Some IP phones only provide basic Layer 2 connectivity. Meaning the IP phone essentially acts as a hub when combining the data and voice segments. Some IP phones provide enhanced Layer 2 connectivity with the option to use VLAN technology, such as 802.1q, to place the phone and the data port in two different VLANs. This architecture assumes that the IP phones are deployed and support VLANs, separating the data and voice segments. Security designs should not rely solely on VLANs for network separation. Rather, they should follow layered security best practices and also rely on Layer 3 access control within the distribution layer into which the IP phone connects. This best practice is followed in all designs.

Because the deployment of soft-phones provides a path for attacks against the voice segment, we don't recommend using them unless a stateful firewall brokers the data-voice interaction. Soft-phones naturally reside in the data segment and require access to the voice segment in order to access call control, place calls to IP phones, and leave voice messages. Calls placed between IP telephony devices generally use dynamically assigned UDP (User Datagram Protocol) port numbers, requiring a stateful inspection device to allow pinpoint access between the segments. Without a stateful firewall brokering all connections between the data and voice networks, you would have to allow wide UDP port ranges. In most networks, it is not possible to secure all connections between the data and voice segments with a stateful firewall. Remember that multiple data and voice segments exist in an enterprise, most likely on the same switch. Stateful firewall segmentation would not be feasible here, nor would Layer 3 stateless filtering suffice.

If privacy or security are a concern, soft-phone systems or PC-based IP Phones, which implement VoIP using an ordinary PC with a handset and special software, should not be used (if practical). This axiom corresponds to NIST Recommendation 7 of SP 800-58.

4.10 Attention to Application Servers for Voice [I]

Some application servers, such as a voice-mail system server, must be accessed from both data and voice segments. They are placed in the data segment in most cases and while simultaneously communicating with IP-PBX within the voice segment. A unified voice-mail system server uses the traditional e-mail store in the data segment for voice-message storage. It requires communication with the IP-PBX to notify users of voice mail. Generally, these services run over well-known TCP ports. It is appropriate to consider introducing a stateful firewall in order to keep necessary communication link and filter out DoS attack packets.

4.11 Physical Security around VoIP System [O]

Physical controls are especially important in a VoIP environment and should be deployed accordingly; so, you must implement and reinforce all types of physical security devices. This axiom corresponds to NIST Recommendation 4 of SP 800-58.

4.12 Power Blackout Consideration [O]

Most VoIP phones and soft phones require power; so, power injectors, backup power and generators must be added to all critical functions, including diverse wiring and switches, in order to keep them work during power outages. You should evaluate costs for additional power backup systems and requirements for ensuring continued operation. This axiom comes from NIST Recommendation 5 in SP 800-58.

4.13 Review Statutory Requirements with Legal Advisors [L]

You should carefully review statutory requirements regarding privacy and record retention with competent legal advisors. This axiom comes from NIST Recommendation 9 in SP 800-58.

4.14 E-911 Consideration [L--US]

In the U.S., E-911 emergency services communications must support the automatic location service, which supplies caller's location to E-911 call centers. You should either give special consideration for the automatic location service with enough accuracy (i.e., with street number) or persuade your customer to give up E-911 calls with VoIP systems. This axiom corresponds to NIST Recommendation 3 of SP 800-58.

4.15 Vulnerability of SIP and IP specification [O]

Some vulnerability exists within the SIP specification, which is difficult to completely remove due to the nature of a VoIP system as a distributed system. For example, a "Bye request" and "Cancel request" from an unauthorized third party (also known as "Tear Down attack") can not be detected, because we cannot distinguish the third party from legitimate SIP proxies.

The Internet Protocol or IP is also vulnerable in a broadband network because some protocol parameters (such as a field length designed in the 1970s) are no longer appropriate for the ultra high speed networks. The vulnerability could be exploited in order to launch a DoS attack to an IP-PBX or an IP phone.

Although exploitation prevention might not represent a general solution, consider taking full advantage of any information about individual network topology when accepting request packets. In addition, introduce mechanisms such as a Network Intrusion Detection System or NIDS to detect and record malicious packets.

5. IP Telephony and Encryption

Although encryption does not resolve all IP telephony security issues, it significantly reduces expected probability of losses and damage caused by attacks. There are several encryption schemes for IP telephony, each of which presents various pros and cons. You should select the most appropriate encryption scheme from available alternatives.

5.1 Fundamentals of Encryption

Encryption, by definition, is the algorithmic process of obscuring data to render it unreadable, without special knowledge. The original data is called as “**plaintext**”, and the encrypted form is called “**ciphertext**”. The process by which the original plaintext is recovered from a ciphertext is called **decryption**. The encryption and decryption process is parameterized by a piece of information called a “**key**”. The robustness of encryption depends upon the length (size) of a key.

Encryption algorithms are categorized into two types. One type is called symmetric encryption algorithms, and the others are asymmetric encryption algorithms.

The **symmetric encryption algorithms** use a common key for encryption and decryption. Their encryption and decryption process is generally fast. Two parties who exchange encrypted messages must share the common encryption key before beginning the communication process. A different key is selected for each communicating pair or for each session. DES, 3-DES and AES are well known algorithms in this category.

The **asymmetric encryption algorithms** use a pair of keys. While one key is used for encryption, the other is used for decryption. Usually, a pair of keys are assigned for each user or network node. One of them is publicly announced and the other is confidential. The former is called a **public key**, and the later is a **private key**. A system for formally announcing a public key for each user or node is called **public key infrastructure** or **PKI**. By encrypting with the public key of a receiver, secret messaging is realized. When a sender encrypts his message with his private key, a receiver can check the sender’s authenticity by decrypting the ciphertext with sender’s public key. In general, asymmetric encryption algorithms are more complex to process than symmetric encryption algorithms and use longer keys. So they are not necessarily appropriate to use in IP telephony systems until processors with high enough performance are adopted for IP phone devices. RSA is a well known algorithm of this category.

5.2 Targets of Encryption

In an IP telephony system, two types of messages are exchanged (see Figure 2). One message type is called **signaling**, which raises a request and controls and checks the status of other nodes within the systems. They are carried in the Session Initiation Protocol or SIP. The other type of messages, called **media streams**, carry voice over the Real-time Transport Protocol or RTP.

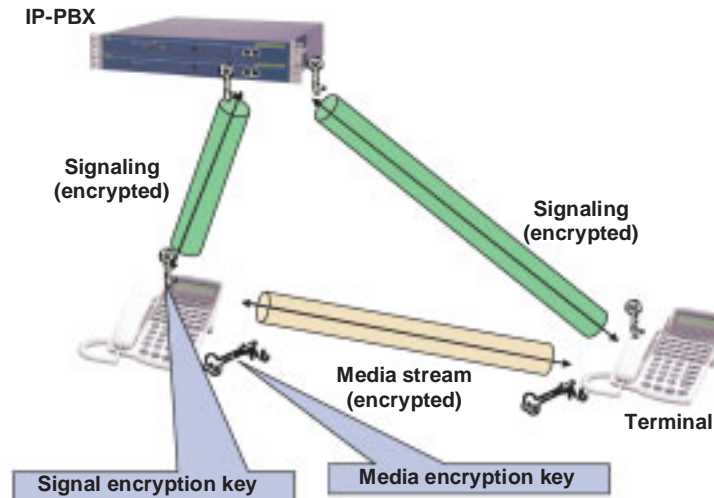


Figure 2 Two Types of Messages and Encryption

To protect signaling messages from abnormal SIP message injection, the SIP protocol is encrypted. To protect voice messages from eavesdropping, the RTP protocol is encrypted. The encryption scheme for RTP is standardized as Secure Real-time Transport Protocol or SRTP. There are two approaches for SIP encryption. One approach encrypts entire SIP messages, and the other encrypts only specific fields of SIP messages. Although the SIP standard mentions TLS and S/MIME as mechanisms for implementing the two, they are not widely adopted on a product base, since they employ an asymmetric encryption algorithm demanding high performance processors. In the meantime, a proprietary extension for SIP encryption is necessary.

Prior to beginning an encrypted communication with a symmetric encryption algorithm, a common encryption key is established and shared securely between two communicating parties. This process is called **key exchange**. The Internet Key Exchange or IKE is a well known protocol for key exchange, but it is inappropriate because it is based on an asymmetric encryption algorithm. The key exchange for SRTP may be realized with a secured SIP protocol.

Another approach is to encrypt whole IP packets with the IPsec protocol. It is widely used for implementing virtual private networks or VPNs linking between network sites. However, the application of IPsec to encrypt inter-node communications leads to the implementation issue of secure key exchange.

5.3 Encryption in UNIVERGE IP Telephony Solution

With the UNIVERGE solution, signaling (SIP) messages between IP-PBX and nodes supporting encryption, including IP phones, media gateways, and voice stream (RTP) messages between nodes, the supporting encryption can be encrypted. The encryption algorithm is AES. The key exchange for SRTP — encrypted RTP — is provisioned by SIP messages for each session. Since the key provisioning must be confidentially performed, the SIP encryption should be enabled whenever SRTP is selected.

There are two options for SIP encryption: the full signaling encryption and the partial signaling encryption. The full signaling encryption encrypts all parts of SIP messages, so no one can view them except two signaling end points. Regarding the partial encryption, only sensitive fields of SIP messages are encrypted, and other fields, which SIP application level gateways or SIP-ALGs must interpret, remain in plaintext. The full signaling encryption is a more robust solution than the partial signaling encryption against some attacks. However, the partial encryption must be chosen when some signaling messages pass through SIP-ALGs such as VoIP-ready firewalls.

The key exchange for the SIP encryption is performed every time a node sends a register request SIP message to IP-PBX based on a secret shared key. The secret shared key is automatically generated and set when the system is installed or a new node is added into the system.

6. References

6.1 RFCs and Drafts

For a listing of the H.323 standard and associated components, refer to the primer and www.itu.int.

RFC 2196:

“Site Security Handbook”, September 1997, <http://www.ietf.org/rfc/rfc2196.txt>

RFC 3261:

“SIP: Session Initiation Protocol”, June 2002, <http://www.ietf.org/rfc/rfc3261.txt>

RFC 3665:

“Session Initiation Protocol (SIP) Basic Call Flow Examples”, December 2003, <http://www.ietf.org/rfc/rfc3665.txt>

RFC 2976:

“The SIP INFO Method”, October 2000, <http://www.ietf.org/rfc/rfc2976.txt>

RFC 3515:

“The Session Initiation Protocol (SIP) Refer Method”, April 2003, <http://www.ietf.org/rfc/rfc3515.txt>

RFC 3702:

“Authentication, Authorization, and Accounting Requirements for the Session Initiation Protocol (SIP)”, February 2004, <http://www.ietf.org/rfc/rfc3702>

RFC 3435:

“Media Gateway Control Protocol (MGCP) Version 1.0”, January 2003, <http://www.ietf.org/rfc/rfc3435.txt>

RFC 3711:

“The Secure Real-time Transport Protocol (SRTP)”, February 2004, <http://www.ietf.org/rfc/rfc3711.txt>

6.2 Notes and Papers from Vendor Independent Parties

NIST CSD: Special Publication 800-58 “Security Considerations for Voice Over IP Systems”, January 2005, <http://csrc.nsl.nist.gov/publications/nistpubs/800-58/SP800-58-final.pdf>

6.3 White Papers from Vendor

Siemens Whitepaper: "Security in Real-Time IP Communications",
<http://www.webtorials.com/main/resource/papers/siemens/paper11/Real-TimeIPSecurity.pdf>

3Com Whitepaper: "IP Telephony Security — A Double-edged Sword?",
http://www.3com.com/other/pdfs/legacy/en_US/3Com-503152.pdf

Cisco Systems Whitepaper: "Securing IP Voice",
http://www.cisco.com/application/pdf/en/us/guest/netsol/ns165/c654/cdccont_0900aec80240249.pdf

6.4 Miscellaneous References

Dynamicsoft Inc.: "[SIP Security](#)", May 2000

7. Acknowledgments

I thank Mr. Sam Safa and other members of this document series development project team for their warm support and many comments on draft review.

8. Glossary of Network Security Terms

ACL (Access Control List or Access List)

A rule set that is typically seen in networking devices like routers and layer 3 switches, and defines whether or not each access to a service or host is permitted. Usually based upon a source IP address or port. In more general context, ACL is a matrix that defines the permission levels from a list of users who access the resources or services.

AES (Advantage Encryption Standard)

A symmetric cryptography algorithm for a block cipher that the U.S. Department of Commerce's National Institute of Standards and Technology (NIST) adopted to new replace DES (Data Encryption Standard) and 3DES ("triple" DES). These are two national standards which are easily breakable. AES uses keys of 128-, 192-, or 256- bit lengths. AES is also known as "Rijndael."

Authentication

The process used to ensure a user's identity. A call controller, IP phone, media gateway, and other components verify the authenticity of a user or call before access is permitted to a resource or device. Password checking is the most common method of authentication, but considered insecure in a hostile environment.

Buffer overflow

Programming errors that write more data than storable within a data storage area of a buffer. Adjacent data is overwritten or corrupted, which causes unexpected program behavior, including system shutdown. If the overflow portion of the data is interpreted as instructions and injected by the attacker, the attacker could execute arbitrary instructions.

Certificate (Digital Certificate)

A structured record issued by a trusted Certificate Authority (CA) to declare that a public key is owned by a specific user (or other entity), including an IP phone and application.

CERT/CC (Computer Emergency Response Team/Coordination Center)

A major reporting center for Internet security problems within the U.S. Part of the Software Engineering Institute of Carnegie Mellon University, established in 1988 and federally funded. CERT/CC staff coordinates responses to security breaches and offers technical assistance. Visit the web site <<http://www.cert.org/>> and <<http://www.us-cert.gov/>>. Its versions in other areas include [JPCERT/CC](http://www.jpccert.or.jp/) <<http://www.jpccert.or.jp/>> in Japan, [NISCC](http://www.niscc.gov.uk/) <<http://www.us-cert.gov/>> in Britain and [AusCERT](http://www.auscert.org.au/) <<http://www.auscert.org.au/>> in Australia.

Clear-text

Non-encrypted data. Clear-text is also known as "plaintext". Encrypted data is known as "cipher-text".

Denial of Services (DoS)

Network or system attacks that disturb normal service support. A typical method is to flood with too many requests or bogus packets. Another method is to send malicious or ill-formed messages to a target node, causing the node stop completely or behave abnormally.

Digital Signature

An electronic signature allowing a recipient to verify that the message was sent by the person claiming the certificate. A sender uses a “hash function” to compute a small digest of his message. He then encrypts the digest with his private key and turns it into a digital signature. In order to authenticate the message, the recipient ensures that the digest of a received message matches the digital signature decrypted with the sender’s public key.

Encryption

An algorithmic process that transforms data from plaintext to cipher-text with a key; this prevents third parties from viewing the original data.

Flood

A large number of messages or packets appearing more rapidly than manageable. “SYN flood” is an example, which is launched to attack a TCP-enabled node.

Hardening

Renders a host or server harder to attack by closing security holes and minimizing vulnerabilities within the operating system, application software and configuration. Removing unnecessary default accounts and shutting down unnecessary services are typical hardening activities.

IDS (Intrusion Detection System)

A system that detects attacks and anomalous activities on a network or computer system. A Network IDS (NIDS) is designed to detect by gathering and analyzing network packets, whereas a Host IDS (HIDS) detects illegal actions by analyzing logs from the OS and applications.

IPsec

A security protocol (of the Internet protocol family) that provides authentication and encryption. IPsec works at layer 3 and inter-operates with any application level protocol over IP.

IPS (Intrusion Prevention System)

A system that detects and blocks intruders on a network or computer system. See *IDS (Intrusion Detection System)*.

L2 Authentication

Network access control implemented within the Data link layer (or layer two). IEEE802.1x is a common standard.

Lockdown

By definition, to restrict the functionality of a system. If network administrators lock down a client desktop, users can perform only certain operations. If network administrators lock down a node, then mapping the node's identity via its MAC, IP address or other characteristics, restricts network access from the node and into a specific physical port on a Layer 2 and/or Layer 3 switch.

Man-in-the-Middle Attack

Various attacks conducted by interjecting into the path of secure communications or key exchange. The attacker is able to read, insert and modify messages between two parties without either party knowing that the link between them has been compromised. The attacker must be able to observe and intercept messages moving between the two victims. This situation is especially serious in the context of key exchange.

NAT (Network Address Translation)

A function that rewrites the source and/or destination addresses of IP packets when they pass through a router or firewall. NAT is most commonly used to enable multiple hosts on a private network to access the Internet using a single public IP address. According to specifications, routers should not act in this way, but it represents a convenient and widely-used technique. NAT introduces complications in IP telephony systems.

NAT Traversal

NAT Traversal allows connection establishment across a gateway device, whether or not it uses NAT. Network applications that consider an IP address of the other party, including IPsec VPN and IP telephony, do not work without NAT Traversal. Many techniques exist, but none are perfect. Examples include UDP hole-punching and STUN.

Network Analyzer

A specialized hardware device or software on a desktop or laptop computer that captures packets passing through a network for inspection and problem detection. Also known as a "sniffer", "packet sniffer", "packet analyst", "traffic analyst" or "protocol analyst," the network analyst plugs into a port on a network hub or switch and decodes one or more protocols into a human-readable format for the network administrator. It can also store packets on disk for further analysis later on. Network analysts can also analyze the packets in real-time to alert the administrator about problems. Hardware network analysts can detect voltage and cable problems, which software-only analysts cannot.

Open port

TCP/IP port is a number assigned to user sessions and server applications. The port number resides in the TCP header as well as in the UDP header for applications such as voice over IP (VoIP) and videoconferencing. The software that responds to a port number is said to be "listening" for its packets. "Open ports" of a host are a set of ports listened by an application on the host. "Open" means a service is running and is ready to process request.

Port scan

A process to search a network host for open ports. Typically automated tools, called “port scanners” are used. While administrators conduct a port scan to check the security of their system, attackers also try port scans to locate easily exploitable system vulnerabilities.

Radius (Remote Authentication Dial-In User Service) Server

A database server for authentication, authorization, accounting or AAA. It stores the user name, password and other authentication information. It also replies to authentication queries from clients and gathers accounting information from those clients.

Registration

When a VoIP endpoint (IP hard or soft-phone, for example) joins an IP telephony network, it informs the IP-PBX that it is present and available for calls with some additional attributes by a register request of the SIP protocol. Some form of authentication usually precedes, or is a part of, this process.

Resource starvation

A form of service denial where the attacker repeatedly asks for a resource, such as a memory space, a TCP connection or a DHCP address lease. With a large volume request, an attacker can overwhelm the victim until servicing legitimate requests from other nodes is no longer viable. In more serious cases the victim might fall into deadlock.

Security scan

A test to search vulnerabilities on a system or network. A security scan does not attempt to break into the system, but rather it tries to find areas of vulnerability. A security scan uses a variety of automated software tools, typically performing hundreds of routine tests and checks. Security experts recommend that a security scan be undertaken at least quarterly. Nessus is one of commonly used tools for automated security scans.

Signaling

Control signals sent back and forth in order to start and stop a transmission or other operation. The signals are the commands that request an operation to be performed such as establishing, monitoring, or releasing IP-telephone calls. It is also referred to as call control. H.323 and SIP are common VoIP signaling protocols.

SIP ALG (Application Level Gateway)

A gateway with an application-level proxy server for SIP. In general, an application level gateway or ALG provides all the basic proxy features and also provides extensive packet analysis. In an IP telephony solution, two application level protocols, SIP and RTP (or SRTP) are used. The port number for RTP (or SRTP) packets is dynamically assigned. SIP ALG can check it to pass through only RTP (or SRTP) packets with legitimate IP addresses and port numbers. It is not the case for IP-layer packet filtering gateways.

SRTP (Secure Real-Time Transport Protocol)

A security-enhanced version of the Real-Time Transport protocol (RTP), which is ubiquitously employed for voice delivery within an IP telephony network. SRTP, specified in RFC3711, adds confidentiality, message authentication, and replay attack protection

Spoofing

Network-based attacks which involves altering the source address of a computer to disguise the attacker and exploit weak authentication methods. IP Spoofing is used to bypass access control lists or firewalls as well as obscuring the real identity or location of an attacker.

SSL (Secure Socket Layer)

A security protocol developed by Netscape. It is widely used with the HTTP protocol for the World Wide Web or WWW. It can validate the identity of a site and create an encrypted connection for protecting sensitive data. It can also be used to secure other TCP based protocol than HTTP. SSL is currently at version 3. See TLS.

Stateful inspection

A firewall technology that ensures that all inbound packets are the result of an outbound request. It is also called “stateful packet inspection” (SPI), designed to help prevent an attacker from sending harmful or bogus packets into the network. Stateful inspection causes problems with videoconferencing and VoIP, in which a user outside the network wants to initiate a communication with a user inside the network.

TLS (Transport Layer Security)

A security protocol based on the Secure Sockets Layer (SSL) 3.0 protocol developed by Netscape. It is standardized in RFC 2246 by IETF and replacing SSL. In the same way as SSL, TLS uses digital certificates to authenticate the client as well as authenticate the server, but TLS supports a wider array of cryptography algorithms than SSL.

Trojan horse, Trojan

A program that appears legitimate, but performs some illicit or unwanted activity when it is run. It may be used to capture password information or make the system more vulnerable to future entry or simply destroy programs or data on the hard disk. It is similar to a virus, except that it does not replicate itself. It stays in the computer doing its damage or allowing somebody from a remote site to take control of the computer. It often sneaks in with a free game or other utility software. The word “Trojan” comes from Greek mythology, in which the Greeks battled the people of Troy.

UDP (User Datagram Protocol)

A protocol within the TCP/IP protocol suite that is used in place of TCP when a reliable delivery is not required. UDP is widely used for IP telephony, streaming audio and video, and videoconferencing, because there is no time to retransmit erroneous or dropped packets.

Virus, Computer virus

Self-replicating, malicious programs that spread by inserting copies of themselves into other executable code or document files. They are one of the several types of malware to malicious software in the narrow sense. However, they often represent a wider category, including other types of malware, such as worms. Anti-virus vaccines are only effective for viruses in the former sense.

VLAN, Virtual LAN

A logical sub-network within a local area network that is created via software rather than manually moving cables in the wiring closet, segregating traffic sent over a common physical network infrastructure into the logically discrete sub-networks (virtual LANs). VLANs are implemented in port switching hubs and LAN switches.

Worm

Self-replicating, malicious program that replicates itself through a network by taking advantage of vulnerabilities within a network service. When trying to spread, worms often clog network traffic, corrupt data or crash infected systems. The Morris worm is the first historical example of worm, which broke out on November 2, 1988. Other infamous worms include Sasser, Nimda, and Blaster.