

# VOICE OVER IP SECURITY

February 2008

© The Government of the Hong Kong Special Administrative Region

The contents of this document remain the property of, and may not be reproduced in whole or in part without the express permission of the Government of the HKSAR.

**Disclaimer:** Whilst the Government endeavours to ensure the accuracy of the information in this paper, no express or implied warranty is given by the Government as to the accuracy of the information. The Government of HKSAR accepts no liability for any error or omission arising from or related to the use of the information.

## TABLE OF CONTENTS

Summary .....	2
I. Introduction.....	3
What is VOIP? .....	3
II. VoIP Protocols.....	4
ITU-T H.323 .....	4
Session Initiation Protocol (SIP).....	5
Media Gateway Control Protocol (MGCP) .....	6
III. Security Threats to Voice Over IP.....	7
IV. Countermeasures and Best Practices .....	10
Encryption using IPSec, TLS and S/MIME .....	10
User and Device Authentication .....	10
Controlling the Interaction between the Voice to Data Segments .....	11
Packet Inspection .....	12
Anti-virus Software and Security Patching.....	12
Best Practices for VoIP Phone End-Users .....	13

## **SUMMARY**

Voice over Internet Protocol (VoIP) technology unites the worlds of telephony and data, by enabling the transfer of voice content (both phone calls and faxes) over the Internet, an Intranet or other packet-switched network. Using either a dedicated IP telephone set, or a networked computer running VoIP software, home and business users can use VoIP technology for direct voice communication. While cost effectiveness is one of major incentives for corporations in moving from a traditional telephony system to VoIP, VoIP infrastructures allow for new value-added voice and multimedia services. This creates tremendous revenue potential for service providers and has increased its deployment in the marketplace. This paper discusses a number of relevant security issues related to the use of VoIP technology, and recommends measures that can be taken to safeguard VoIP networks and systems.

## I. INTRODUCTION

### WHAT IS VOIP?

In VoIP technology, the voice signal is first separated into frames, which are then stored in data packets, and finally transported over IP network using voice communication protocol. Currently, most VoIP systems use either one of two standards; H.323<sup>1</sup> or the Session Initiation Protocol (SIP)<sup>2</sup>, although a few still use proprietary protocols like SCCP<sup>3</sup>.

---

<sup>1</sup> <http://www.itu.int/rec/T-REC-H.323/e>

<sup>2</sup> <http://www.ietf.org/rfc/rfc3261.txt>

<sup>3</sup> <http://www.javvin.com/protocolISCCP.html>

## II. VOIP PROTOCOLS

The two most widely used protocols for VoIP are the ITU standard H.323 and the IETF standard SIP. Both are signalling protocols that set up, maintain and terminate a VoIP call. In addition, the Media Gateway Control Protocol (MGCP)<sup>4</sup> provides a signalling and control protocol between VoIP gateways and traditional PSTN (Public Switched Telephone Network) gateways<sup>5</sup>.

### ITU-T H.323

H.323 is a comprehensive protocol under the ITU-T specifications for sending voice, video and data across a network. The H.323 specification includes several sub-protocols:

1. H.225 for specifying call controls (e.g. call setup and teardown);
2. H.235 for specifying the security framework for H.323 and the call setup;
3. H.245 for specifying media paths and parameter negotiations such as terminal capabilities;
4. H.450 for specifying supplementary services such as call hold and call waiting.

---

<sup>4</sup> <http://tools.ietf.org/rfc/rfc3435.txt>

<sup>5</sup> <http://tools.ietf.org/rfc/rfc3525.txt>

H.235 also provides security features such as authentication, integrity, privacy and some non-repudiation support in H.323 communications. It is designed to operate seamlessly with other protocols like H.245 and H.225.

A call setup is secured through Transport Layer Security (TLS). Once established, a call control is initiated so that encryption and media channel information can be negotiated. H.323 utilises RTP (Real-time Transport Protocol) / RTCP (Real-time Transport Control Protocol) as its transport protocol, which rides on top of UDP. Encryption is performed within the RTP packet by third party hardware, or at the network layer.

Authentication under H.323 can be either symmetric encryption-based or subscription-based. For symmetric encryption-based authentication, prior contact between the communicating entities is not required because the protocol uses Diffie-Hellman key-exchange to generate a shared secret identity between the two entities. With reference to the H.235 recommendation, a subscription-based authentication requires a prior shared secret identity, and there are three variations of this; (1) password-based with symmetric encryption, (2) password-based with hashing, and (3) certificate-based with signatures.

## **SESSION INITIATION PROTOCOL (SIP)**

SIP is a text-based application layer protocol that addresses the signalling and session management within a packet telephony network. It is defined in RFC 3261<sup>6</sup>. SIP uses a “request-response” model similar to the HTTP protocol. In SIP, authentication and authorisation are handled “*either on a request-by-request basis with a challenge/response*

---

<sup>6</sup> <http://tools.ietf.org/rfc/rfc3261.txt>

*mechanism, or by using a lower layer scheme*<sup>7</sup>. As SIP is a lightweight protocol, its security capabilities are very limited. SIP requests and responses cannot be end-to-end encrypted because message fields such as the request and route need to be visible to proxy servers that are present in many network architectures to ensure SIP requests are routed correctly. Voice data is transmitted in clear text over UDP and TCP.

Although SIP supports S/MIME-based encryption using digital certificates, certain header fields used in requests and responses cannot be encrypted. The SIP protocol relies on transport layer security mechanisms such as TLS or IPSec to provide the required security for the whole message.

## **MEDIA GATEWAY CONTROL PROTOCOL (MGCP)**

MGCP is published by the Media Control Working Group as RFC 3435<sup>8</sup>. It “*expect[s] that MGCP messages will always be carried over secure Internet connections as defined in the IP security architecture as defined in RFC 2401, using either the IP Authentication Header, defined in RFC 2402, or the IP Encapsulating Security Payload, defined in RFC 2406.*”<sup>9</sup>. This allows for data origin authentication, connectionless integrity and optional anti-replay protection of messages passed between the Media Gateway (MG), which converts circuit-switched traffic to packet-based traffic, and the Media Gateway Controller (MGC) that dictates the service logic of the traffic.

---

<sup>7</sup> <http://tools.ietf.org/rfc/rfc3261.txt>

<sup>8</sup> <http://tools.ietf.org/rfc/rfc3435.txt>

<sup>9</sup> Same as above.

### III. SECURITY THREATS TO VOICE OVER IP

VoIP systems rely on a data network, which means security weaknesses and the types of attacks associated with any data network are possible. For example, in a conventional telephone system, physical access to the telephone lines or a compromise of the office private branch exchange (PBX) is required for in order to conduct activities such as wire-tapping. But for VoIP, voice is converted into IP packets that may travel through many network access points. Therefore the data is exposed to many more possible points of attack that could be used for interception by intruders. In fact, all the security risks associated with IP, such as computer viruses, Denial of Service and man in the middle attacks, are also dangerous to VoIP systems.

In particular, PC-based IP Phone hosts are more susceptible to attacks due to the prevalence of attack techniques pinpointing PC systems. These include operating system vulnerabilities, application vulnerabilities, service vulnerabilities, worms, viruses, and so on. A PC-based IP Phone is also at risk from any attack aimed at the entire data segment upon which it is residing.

Since voice communication protocols are session control protocols, IP addresses and TCP / UDP port information is enclosed in packets. In networks that use a Network Address Translation (NAT) technique, the IP address and port information in the packets cannot be encrypted because NAT devices require such information to perform the translation. This imposes another security constraint to these protocols.

The H.323 protocol is secured by using TLS, where a pre-defined TCP port 1300 must be used for the establishment of the Call Connection Channel and where no other security

mechanism is available for the first connection. This fixed and well-known port can be a threat to the protocol.

For SIP, encryption is based on the use of S/MIME. Only certain headers in the message are encrypted and critical headers, such as “To”, “From” and “Call-ID” fields, are not encrypted.

An "Uncontrolled barge-in" is a well-known security threat to MGCP. Since voice packets can be directed to a gateway through the appropriate UDP port, an attacker might be able to listen in on voice communications, unless protection is in place. To mitigate this threat, a gateway should only accept data from a pre-defined IP address and UDP port. The down side is that this adds processing overhead, and IP addresses can be spoofed. To counter spoofing, the MGCP can be configured to obtain a “remote session description” from the initiating gateway and pass this to the destination gateway for verification. However, this increases the call setup time.

VoIP services are also subject to spamming, known as a SPam over Internet Telephony (SPIT) attack<sup>10</sup>. SPIT leaves unsolicited marketing voice messages on target IP phones. As voice messages are generally larger in data size than email messages, the impact of SPIT on the network is much higher than the typical SPAM email counterpart.

Attackers have also tried exploiting VoIP technology to hijack identities and steal money<sup>11</sup>. This category of attack is similar to an email-based phishing attack, and hence its name “vishing” (or VoIP phishing). A victim will receive an email or be contacted with a phone call that directs him or her to a customer service number where they go

---

<sup>10</sup> <http://www.voip-news.co.uk/2007/11/01/spam-over-internet-telephony-threat-grows/>

<sup>11</sup> <http://www.fbi.gov/page2/feb07/vishing022307.htm>

through a number of voice prompted menus, in an attempt to steal account numbers, PINs, and other critical information.

## **IV. COUNTERMEASURES AND BEST PRACTICES**

### **ENCRYPTION USING IPSEC, TLS AND S/MIME**

Encryption is a means of preserving the confidentiality of transmitted signals. As SIP is an application-layer protocol, encryption mechanisms could be used at lower layers of the protocol stack, such as at the network and transport layers, by means of IPSec and TLS respectively. For the body of SIP messages, S/MIME can be used. However, encryption and decryption are CPU-intensive and take time to process, with a corresponding impact on performance. If the overall latency of the VoIP call is greater than 250 milliseconds, the quality of the call will be noticeably affected. In addition, SIP requests and responses cannot be completely encrypted for end-to-end security, because certain header fields (e.g. IP address, port numbers, etc) must be visible to proxies for the purpose of data routing.

Even if encryption is used, physical access to VoIP servers and gateways may allow an attacker to perform traffic analysis and derive call information from encrypted messages<sup>12</sup>. Therefore, adequate physical security should be in place to restrict access to key VoIP network components.

### **USER AND DEVICE AUTHENTICATION**

---

<sup>12</sup> [http://www.freeswan.org/freeswan\\_snaps/CURRENT-SNAP/doc/ipsec.html](http://www.freeswan.org/freeswan_snaps/CURRENT-SNAP/doc/ipsec.html)

Some call processing servers have an automatic phone registration feature that “bootstraps” an unknown phone with a temporary configuration and then allows it to interact with the network. This feature is useful for bulk deployment of IP phones, but it also poses a threat, because rogue devices can start unauthorised services or launch attacks against other devices after they are attached to the network. Device authentication using the MAC address of an IP phone is one solution to this problem. The automatic registration function of the call-processing server should also be disabled. If a phone with an unknown MAC address attempts to download a network configuration from the call-processing server, the request will be rejected and the rogue IP phone will not be granted a network configuration. They will not be able to connect to the network because the server does not recognise the MAC address. User authentication (such as with a user ID and password, or a personal identification number (PIN)<sup>13</sup>) is also an effective measure to avoid call masquerading, because it provides certain level of non-repudiation and certainty to the caller’s identity.

## **CONTROLLING THE INTERACTION BETWEEN THE VOICE TO DATA SEGMENTS**

IP-based telephony provides a platform for telephone calls over an existing IP data network. However, in order to maintain quality of service (QoS), scalability, manageability, and security, voice and data should be separated using different logical networks as far as possible. Segmenting IP voice from a traditional IP data network greatly enhances the mitigation of VoIP attacks.

Given that voice and data segments should be separated, technologies such as virtual LANs (VLANs), access control, and stateful firewalls, can provide the type of Layer 3

---

<sup>13</sup> [http://www.cisco.com/en/US/docs/voice\\_ip\\_comm/cucm/admin/3\\_3\\_2/ccmcfg/b07user.html](http://www.cisco.com/en/US/docs/voice_ip_comm/cucm/admin/3_3_2/ccmcfg/b07user.html)

segmentation necessary to keep voice and data segments separate at the access layer. Stateful firewalls should be placed at locations on the network where segments are allowed to interact. Layer 3 switches can also be used to control access between data and voice segments via access control and filtering.

To further protect a VoIP infrastructure, the organisation might consider a closed VoIP system, in which qualified users within the organisation have access to VoIP services for internal communication purposes only. This would completely segregate VoIP services from the Internet and minimise threats from external network attacks.

## **PACKET INSPECTION**

Certain Network Intrusion Prevention Systems (NIPS)<sup>14</sup> can parse and analyse VoIP protocols. For VoIP systems that have NIPS implemented, attacks targeting VoIP services can be detected and prevented.

## **ANTI-VIRUS SOFTWARE AND SECURITY PATCHING**

Computers which use software for VoIP connections, as well as all other VoIP servers and gateways, should be protected with a personal firewall, along with anti-virus and malicious code software that are up to date with the latest virus signature and/or malicious code definitions. This provides basic protection against attacks on the data segment that could be transferred to the voice segment. In addition, security patches for PC-based IP phones as well as VoIP servers and gateways should be kept up-to-date.

---

<sup>14</sup> [http://www.icsalabs.com/icsa/topic.php?tid=aeb3\\$347927fc-1ef76edd\\$c11f-5218d1b1](http://www.icsalabs.com/icsa/topic.php?tid=aeb3$347927fc-1ef76edd$c11f-5218d1b1)

Because of the prevalence of virus threats to client PCs, a VoIP dedicated IP hardware telephone is preferable to a software-based IP phone (that is, a computer running VoIP software). A software-based IP phone is at more risk, because VoIP software might be able to tunnel data through a corporate firewall. If the software is mis-configured or subject to vulnerabilities, it may enable disallowed data packets to bypass the firewall. In addition, computers are more vulnerable to viruses, worms, and other threats due to vulnerabilities from other system components such as web browsers. However, hardware IP telephones do sometimes have their own security problems if, for example, there is a design flaw in the phone itself. .

## **BEST PRACTICES FOR VOIP PHONE END-USERS**

Users of VoIP phones need to be aware of and plan for contingencies when making voice calls if a VoIP system should fail. To protect from vishing attacks, users should not release sensitive information such as credit card data, bank account, or login credentials to strangers during a VoIP communication or conversation.

For users of software-based IP phones, their computers must be protected with a personal firewall, plus anti-virus / malicious code repair software updated with the latest virus signatures and malicious code definitions. A consistent patch management procedure should be enforced to ensure all software components, including IP phone software, installed on computers are patched and updated appropriately.

Sensitive data should not be stored in an IP phone, as there are no encryption systems built into IP phones to protect data. Also other people, such as administrators, might be able to gain access to sensitive data in an IP phone remotely via TFTP.