Voice Over Internet Protocol (VoIP) and Security

Greg Tucker
Voice Over Internet Protocol (VoIP) and Security

Greg S. Tucker
October 26, 2004

GIAC Security Essentials Certification (GSEC)
Practical Assignment
Version 1.4c, Option 1
Table of Contents

Abstract .............................................................................................................................. 1

Background ...................................................................................................................... 1

Quality of Service ......................................................................................................... 1
  Latency ......................................................................................................................... 2
  Jitter ............................................................................................................................. 2
  Packet Loss .................................................................................................................. 3

VoIP Components ......................................................................................................... 3
  End-user equipment ...................................................................................................... 3
  Network components .................................................................................................. 3
  Call processor ............................................................................................................. 4
  Gateways ..................................................................................................................... 4
  Protocols ..................................................................................................................... 5
    H.323 ......................................................................................................................... 5
    H.323 Security Concerns ......................................................................................... 6
  SIP ............................................................................................................................... 6
    SIP Security Concerns ............................................................................................. 7

NAT and VoIP ................................................................................................................. 8

Denial of Service ............................................................................................................ 8

Other Concerns ............................................................................................................. 9

Conclusion ..................................................................................................................... 10

References ..................................................................................................................... 11
Abstract
This paper will describe Voice Over Internet Protocol (VoIP) to a level that allows discussion of security issues and concerns. Business concerns of implementing VoIP, components of a VoIP system, and relevant security issues and concerns as they apply to the topics, are explored. The business concerns will be those that affect Quality of Service (QoS). VoIP components will include end-user equipment, network components, call processors, gateways and two of the more common architectures: H.323 and Session Initiation Protocol (SIP). Denial of service will be discussed, and, encryption and network address translation (NAT) will be discussed emphasizing how they impact the implementation of VoIP.

Background
Before VoIP, telecommunications occurred over a public switched telephone network (PSTN), that is, voice data traversed circuit switched connections. The cost savings of VoIP, both in dollars and bandwidth, compared to that of circuit switched networks (CSN), is encouraging companies to move to VoIP. VoIP deployment has brought with it many security concerns, motivating the need for security solutions to deal with the many issues.

VoIP security is complicated by the requirement of multiple components, in most cases, more components than traditional CSNs, and the fact that it is normally deployed on the current data network. Often, normal deployment requires co-existence of the CSN until VoIP functions have replaced those of the CSN. The security approach taken should address CSN and VoIP for as long as both exist.

VoIP is normally thought of as telephone communication. That is, a device that looks like a telephone accepts a telephone number and someone speaks into a handset to someone at the other end. VoIP, however, is much more than that. Kuhn, Walsh and Fries point out in the National Institute of Standards and Technology (NIST) draft document “Security Considerations for Voice Over IP Systems,” “Just about any computer is capable of providing VoIP.” Applications such as Microsoft’s NetMeeting, Apple Macintosh’s iChat & those for Linux all provide a variation of VoIP at the desktop. With these variations come additional vulnerabilities that should be addressed.

Quality of Service
VoIP has introduced requirements for data packets to reach their destination in a more restricted time frame than other internet protocol (IP) applications. Many applications are somewhat tolerant of packet delay and it may be imperceptible to the client using the application. Packet delay in VoIP, however, can reduce the functionality to unusable. Management of this delay is Quality of Service (QoS). From Chapter 49, Quality of Service Networking, of Cisco’s “Internetworking Technology Handbook”, “Quality of Service (QoS) refers to the capability of a network to provide better service to selected network traffic over various technologies, including Frame Relay, Asynchronous Transfer Mode (ATM), Ethernet and 802.1 networks, SONET, and IP-
routed networks that may use any or all of these underlying technologies.\textsuperscript{3} There are three main QoS issues that are affected by security measures of VoIP. They are latency, jitter and packet loss.\textsuperscript{2}

**Latency** is the time it takes for data to get from the source to the destination. The source is the person speaking into the phone and the destination is the listener at the other end. This is one-way latency. This time added to the one-way latency back to the originating client is round trip latency. PSTNs have a round trip latency of under 150 ms in the US.\textsuperscript{4} Mehta & Udandi point out in *Overview of Voice Over IP*:

The 1996 ITU Recommendation G.114 for one-way end-to-end transmission limits is:
- Under 150 ms: acceptable for most user applications
- 150 to 400 ms: acceptable provided that administrators are aware of the transmission time impact on the transmission quality of user applications
- Over 400 ms: unacceptable for general networking purposes \textsuperscript{4}

This gives us a range of latency that is tolerable of 75 ms to 400 ms for one-way. Since, users of telephone systems in the US have grown to expect less than 150 ms, we will assume 150 ms as the maximum cumulative latency.

Latency is introduced from several sources including the various components of the VoIP system and the network. The constraint put on the time to deliver packets impacts the security solution. Most security solutions are a combination of methods that apply to different parts of the VoIP system and will usually add delay. Of major concern is encryption. Places to implement encryption in VoIP will be discussed later. Encoding and decoding encryption can have significant impact on latency. This is largely due to the size of the key and the complexity of the algorithm. In general, the larger the key, the more secure the data is and the more time it takes to encode and decode.\textsuperscript{5} A balance must be made between the desired security and the desired perceived quality. One of the biggest barriers to securing a VoIP solution properly is being able to implement the security solution with a minimum of latency to ensure that QoS meets the standards set by an organization. Security is often sacrificed for better quality.

**Jitter** is introduced when packets have different latencies. Jitter can cause long delays before packets arrive and can cause packets to get out of sequence. Since most VoIP communication is done with user datagram protocol (UDP) (rfc 768), packets that arrive out of order cannot be reassembled at the protocol level, however sequence numbers and timestamps allow packets to be reassembled at the application level if the application supports it. There is a cost though. Reassembling the packets takes time and packets that take longer to arrive may cause packets to be dropped, holding up the delivery of received packets. A method used to reduce jitter is to create a buffer that resequences packets as they come in and passes them on to the application.\textsuperscript{2} Using 150 ms as the maximum latency, the buffer would need to be cleared every 150 ms minus other latencies. Since some packets may be delayed for longer than this difference, the buffer may get cleared before a group of packets could be resequenced, leaving a gap, leading to packet loss. In general, jitter is reduced as the size of the
buffer increases from zero, however, there will be a size that, when exceeded, will cause more packet loss. Another method of reducing jitter is to use the QoS feature of some routers, switches and firewalls.  

**Packet Loss** is when packets do not arrive at their destination or arrive too late to be processed. Packet loss is usually perceived as gaps in the communication. Defense against packet loss may include sending redundant information and can be reduced with encoding schemes.  

Additionally, network tuning mechanisms can help reduce packet loss.

**VoIP Components**

The components of VoIP include: end-user equipment, network components, call processors, gateways and protocols.

**End-user equipment** is used to access the VoIP system to communicate with another end point. Connection to the network may be physically cabled or may be wireless. The end-user equipment may be a phone that sits on a desk or a softphone that is installed on a PC. Functions include voice and possibly video communication, and may contain instant messaging, monitoring and surveillance capabilities.

Though end-user equipment is often deployed on an internal, protected network, it is usually is not individually protected by other devices (firewalls) and may be threatened if the equipment has vulnerabilities. The threat, of course, is also dependent on the level of security that exists on the internal network. If the device is allowed to reach or can be reached from a public or unprotected network, there may be threats that are not normally found on the internal network. Softphone software may have vulnerabilities, there may be vulnerabilities in the operating system (OS) it is running on, and there may be vulnerabilities of other applications running on the OS. Patching OSs, softphone software and those other applications can help mitigate the risk of any threats that are present. Additionally, some end-user equipment may have firmware upgrades that can be applied or may be able to obtain updated software during registration.

For OS based VoIP solutions, consideration should be given to virus detection and host-based firewalls as well as host-based intrusion detection. Centralization of management of these security components is best, allowing the users of the solution to focus on their duties instead of security details, increasing productivity.

**Network components** include cabling, routers, switches and firewalls. Usually the existing IP network is where a new VoIP system is installed. The impact on the IP network is greater than merely adding more traffic. The added traffic has more of an urgency to reach its destination than most of the data traffic that is already supported. Switches, routers and firewalls will need to recognize and act on VoIP data in order to keep latency down. Additional security measures, addressed later, will complicate this process.
Performance can be gained by separating the data traffic from the VoIP traffic by putting them on different virtual local area networks (VLAN). This allows management of the data to be segregated so it can be handled based on data type. Since the VoIP data must have a higher QoS level, isolation of the data types via VLAN can help increase the performance at the cost of that on other VLANs. This cost may be very low to the other applications. Although VLANs should not be relied on alone, they will add a layer of security. The ability to listen to, or sniff, the network, potentially allows the hacker to monitor calls and manipulate the VoIP system. It is generally more difficult for a hacker to sniff or interfere with the voice traffic from the data VLAN when the voice traffic is on its own VLAN, but it can be done by manipulating the routing of the network. Encryption can also help defend against sniffing.1

Another IP network concern is network slow downs that might increase latency, jitter or packet loss. Slow downs can be caused for many reasons including configuration issues, denial of service (DoS) attacks or high bandwidth utilization by other systems on the network. Configuration issues are probably best addressed with education and checking mechanisms, such as having a co-worker verify configurations. DoS attacks are difficult to defend against, but may be reduced by filtering the traffic that can communicate on the network to be only that which is allowed. This may prove difficult due to the use of random ports by VoIP. Regular network bandwidth analysis can help with tuning of a network and helps with capacity planning. Being aware of bandwidth growth trends helps network administrators know when bandwidth needs to be addressed.

VoIP suffers from most of the same IP network vulnerabilities as other systems. A well-secured internal network is the first step to protecting the VoIP system as it was for the pre-existing IP network. Care must be taken to ensure security solutions keep latencies low or the security solution itself may prove to be a DoS.

Call processor functions can include phone number to IP translation, call setup, call monitoring, user authorization, signal coordination, and may help control bandwidth.6 Call processors are usually software that runs on a popular OS. This leaves it open to network attacks for the vulnerabilities of the given OS, the vulnerabilities of the application and other applications running on the OS.

Gateways can be categorized into three functional types: Signaling Gateways (SG), Media Gateways (MG) and Media Controllers. In general, they handle call origination and detection and analog to digital conversion. Signaling gateways manage the signal traffic between an IP network and a switched circuit network, while media gateways manage media signals between the two. Media Gateway Controllers manage traffic between SGs and MGs. The most common gateway protocols are MGCP (rfc 2705) and Megaco. Both are composites or derivations of previously but now less used protocols.6

Vulnerabilities can exist between the internal IP network and the "gated", circuit switched network. Care should be taken to ensure any vulnerabilities are mitigated. Gateway communication should be secured with IPSec to prevent interference with
calls and to prevent unauthorized calls from being setup. The gateway itself is vulnerable to IP based attacks and can be mitigated by using IPSec and by removing any unnecessary services and open ports, as should be done with any server.

Protocols
There are several protocols used for VoIP but two are most common. They are H.323 and Session Initiation Protocol (SIP).

H.323
H.323 is a protocol suite specified by the International Telecommunications Union (ITU) that lays a foundation for IP based real-time communications including audio, video and data. H.323 allows for different configurations of audio, video and data. Possible configurations include audio only, audio & video, audio & data and, audio, data and video. H.323 does not specify the packet network or transport protocols.

This standard specifies four kinds of components: Terminals, Gateways, Gatekeepers and Multi-point Control Units (MCU). Terminals are the end-user equipment discussed above. Gateways handle communication between unlike networks with protocol translation and media format conversion. Gatekeepers provide services such as addressing, authorization and authentication, accounting functions and call routing. MCUs handle conferencing.

The ITU defines the H.323 zone that consists of terminals, gateways, MCUs, and a gatekeeper. The gatekeeper manages the zone.

H.323 uses different protocols to manage different needs. There are audio codecs and video codecs that encode and decode the audio and video data.

H.225 covers registrations, admissions & status (RAS) and call signaling. RAS handles various functions between the endpoints and the gateway, including registrations and admission control as its name implies. It also manages changes in bandwidth and disengage procedures. A RAS channel is opened, prior to opening other channels, between the gateway and endpoint whereby RAS messages are passed. Call signaling channels are opened between endpoints and between an endpoint and a gatekeeper. They are used to set up connections. Call setup and termination uses Q.931. H.245 is for channel negotiations such as flow controls and general commands and H.235 specifies security.

Real-time Transport Protocol (RTP) is used to transport data, typically via UDP and provides a timestamp, sequence number, data type and ability to monitor delivery. Real-time Transport Control Protocol (RTCP) is used mainly to monitor quality and manage synchronization.

As mentioned above, the H.235 protocols of H.323 are for security profiles. These standards address authentication, integrity, privacy, and non-repudiation and are expressed as Annexes to H.235 Version 2. They are Annexes D, E & F as follows:
Annex D provides message integrity and/or authentication using symmetric keys. It also has a voice encryption option.

Annex E provides authentication, message Integrity and non-repudiation using asymmetric methods.

Annex F is a hybrid of Annex D and Annex E providing authentication, non-repudiation and message integrity.

The four security goals, authentication, integrity, privacy, and non-repudiation are accomplished with the four mechanisms: configuration, authentication, key exchange and encryption. During the initial stage of configuration, the device is authorized to the network and may be authenticated. Integrity and privacy are accomplished through encryption using symmetric or asymmetric keys. A signature is attached to gain the fourth goal of non-repudiation.

**H.323 Security Concerns**

Using H.323 to setup VoIP connections is a complicated process that is made more complex by adding security measures. Many of the protocols used with the H.323 suite use random ports causing problems securing through firewalls but may be able to be mitigated by using direct routed calls. Since the ports required for H.323 are not set, a filtering firewall would have to have all possibly needed ports left open. Therefore, the firewall would need to be H.323 aware allowing communication without opening up the firewall to other traffic. A stateful firewall and/or application firewall is required to ensure consistency of the characteristics of connections.

NAT is a problem for H.323 because the IP and port on the IP header do not match those in the messages. This may be mitigated with an H.323 aware firewall. Additionally, there will be restrictions in other security measures if NAT is involved.

**SIP**

Session Initiation Protocol (SIP) is a signaling protocol specified by the Internet Engineering Task Force (IETF) used to set up and tear down two-way communications sessions. SIP operates on the application level so can be used with several different protocols. Using TCP allows use of SSL/TLS providing more security whereas, UDP allows for faster, lower latency, connections.

Usual components of an SIP system are the user agent (UA), proxy server, registrar server, and the redirect server. The UA software contains client and server components. The client piece makes outgoing calls and the server is responsible for receiving incoming calls. The proxy server forwards traffic, the registrar server authenticates requests, and the redirect server resolves information for the UA client.

The endpoints begin by connecting with a proxy and/or redirect server which resolves the destination number into an IP address. It then returns that information to the
originating endpoint which is responsible for transmitting the message directly to the destination. A security advantage of SIP is that it uses one port.²

The main concerns for security of SIP are confidentiality, message integrity, non-repudiation, authentication and privacy. New security mechanisms were not created for SIP instead, SIP uses those provided by HyperText Transfer Protocol (HTTP) and Simple Mail Transfer Protocol (SMTP) as well as Internet Protocol Security (IPSec).

Signal confidentiality is best provided with full encryption, however, since some SIP message fields must be read and/or modified by some proxies, care must be taken and possibly other methods used. If however, the proxy can be trusted, then encryption at the transport and/or network layers may be the best solution. Security at the transport and networking layers accomplishes full packet encryption using IPSec. TLS had been used, but has been deprecated¹². Full encryption requires support of the encryption method at each end point where it is implemented.

HTTP authentication uses the 401 and 407 response codes and header fields. This provides a stateless challenge-base mechanism for authentication whereby the challenge and user credentials are passed in the headers. When a proxy or UA receives a request, it may challenge to ensure the identity of the sender. Once identity has been confirmed the receiver should also verify that the requester is authorized. Details of this “digest” method may be found in rfc 3261¹².

Secure/Multipurpose Internet Mail Extension (S/MIME) is an enhancement to Multipurpose Internet Mail Extension (MIME) that replaces Pretty Good Privacy (PGP). Since MIME bodies are carried by SIP, SIP may use S/MIME to enhance security, MIME contains components that can provide integrity and encryption for MIME data² and as rfc 2633 states S/MIME can be used for “authentication, message integrity and non-repudiation of origin (using digital signatures) and privacy and data security (using encryption).”¹² S/MIME is useful when full encryption of the packet is not feasible due to the need of network components to use data from the header fields.

User identification is done via certificate belonging to the user that is compared to the header information. Integrity of the message is verified by matching the information in the outside header with that of the inside header. Normally, S/MIME is used to encrypt Session Description Protocol (SDP) but there may be requirements to encrypt certain header components. SIP can provide header privacy by encapsulating the entire message using MIME type message/sip. If used for anonymity the message will need to be decrypted before the certificate can be identified and consequently validated.¹²

**SIP Security Concerns**

HTTP Digest does not provide the best integrity. Without S/MIME, spoofing of the header would not be difficult.¹³

S/MIME requires a public key infrastructure. Since certificates are associated with users, moving from one device to another may be difficult. With S/MIME there may be
issues with firewalls or other proxy devices that may require viewing and/or changing SIP bodies (i.e. SDP).\textsuperscript{13}

There is information in SIP headers that may be considered sensitive, i.e. an unlisted phone number. Consideration may need to be given to providing per-user options that allow protection of this information.\textsuperscript{13}

SIP and H.323 both use protocols that may use random ports requiring that the firewall be able to open and close ports as required. An H.323 or SIP aware firewall may be required.\textsuperscript{2}

As with H.323, NAT presents problems for SIP.

**NAT and VoIP**

Network Address Translation (NAT) allows one network address to be translated at a gateway between two networks into another address so that the packet will have a valid source address on the network it is on. Most commonly, NAT (rfc 1631) is used to change private IP addresses into public, Internet routable, IP addresses. Ports may also be translated. NAT traversal is usually only a concern if end-user devices connect directly with an external network or if they connect to the internal network from an external network.\textsuperscript{2}

NAT is a layer of security because it hides the real addresses on the internal network from the public network.\textsuperscript{2} NAT can however, be a problem, because the routing device does not know the actual IP address of the device. The information defining the endpoint is in the header. The routing device must be able to read the header and in some cases (i.e. with proxy firewalls) change it.\textsuperscript{2} This is hampered when encryption is used.

The best solution is to not use NAT if at all possible. By removing the issue, the problem disappears, though another problem may present itself. When NAT is required, care must be taken to select application and proxy firewalls that handle the implementation or, alternatively, consider a service offered by the public networks.

**Denial of Service**

Denial of Service (DoS) is caused by anything that prevents the service from being delivered. A DoS can be the result of unavailable bandwidth or VoIP components being unavailable. Many things can cause a DoS including: a network getting congested to a level that it cannot provide the bandwidth needed to support the application; servers not capable of handling the traffic; extraneous services may be running that reduce the available resources to the server; malicious programs such as viruses and trojan horses; other malicious programs with the purpose of causing DoS; or hacking activity.\textsuperscript{6}

If DoS is caused by bandwidth constraints, potential solutions are increasing the bandwidth and/or isolating the VoIP traffic so that it gets service first. Various methods of ensuring servers don’t stop working, such as failover methods like clustering, can
help reduce DoS from failing components. Each component of the VoIP system offered by the vendor, should be evaluated, removing those that are unnecessary. Server size should be planned such that all desired vendor services and expected traffic can be supported, adding some percentage for expected growth.

Defending against malicious programs and activity is more difficult but should begin with applying appropriate patches in a timely manner, and installing virus protection with frequent updates. In addition, installation designers should consider a host based firewall, intrusion detection and/or intrusion prevention.

Defense against DoS attacks of public servers can best be done by locating the device with the public available IP addresses behind a firewall or other device that only allows communication from trusted sources. Also, harden the operating systems in use, removing all unnecessary services and applications from the servers and workstations, patching, etc.

Other Concerns

Additional concerns of a VoIP system that need to be considered are databases, web servers, additional VoIP services offered by the vendor, protocol stacks, access to public or unknown networks, physical security and electrical power.

Databases are needed at some point of the VoIP implementation to store and retrieve information as needed to accommodate various functions of the system. Database security principles should be applied including changing the default administrator password, patches as they become available, and best practices concerning access to the database, especially from sources other than the VoIP system.

A common feature of end-user equipment is a web browser, the purpose of which is to provide additional functionality and increased productivity. A VoIP server may have a web browser interface allowing management. If supported, patch the device when the patch becomes available and use as strong authentication as can be supported.

Each vendor, having their own implementation of VoIP, may require any number of services to run on a server to support their product. As mentioned before, keep patches up to date and turn off all unneeded services. If the risk is great enough, consider encryption and/or protection by another device such as a firewall. The voice application and the OS have similar vulnerabilities and should be patched as well.

If the VoIP system stays within a secured network and only connects to the public network through a gateway, the gateway is a vulnerability that needs addressing. Deploy the hardened gateway behind an appropriate firewall, i.e. one that is aware of the protocols used.

VoIP must process the protocols that it supports so it needs to have some implementation of a network stack. Stack implementations are written by the vendor or purchased from another vendor. With the latter, all vendors that purchased a specific
vendor’s stack will share the same vulnerabilities.\textsuperscript{1} Patch if necessary, when patches become available.

Ensure that the components are physically secure. Access to the box allows ownership. There are many methods of compromising a device, depending on the device and the underlying OS, with physical access. Good security practices include removing the a-disk and the CD-ROM from the boot list and password protect the configuration.

If a component is unavailable, then there is a denial of service. Planning should include separate power sources and uninterruptible power supplies (UPS) for the event of a power loss.

**Conclusion**

Security for a VoIP system should begin with solid security on the internal network. It should be protected from the threats of attached hostile networks and the threats of the internal network. The security policy should include any specific VoIP needs. The load of the VoIP system should be accommodated by the network and the servers involved, ensuring that proper resources are in place and available. Conducting a risk analysis of each component and process will identify the vulnerabilities and threats. This will provide the information needed to determine proper measures. Striking a balance between security and the business needs of the organization is key to the success of the VoIP deployment.
References


Other References of Interest


## Upcoming SANS Training

Click here to view a list of all SANS Courses

<table>
<thead>
<tr>
<th>Course Name</th>
<th>Location</th>
<th>Dates</th>
<th>Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>SANS Amsterdam August 2020 Part 1</td>
<td>Amsterdam, NL</td>
<td>Aug 03, 2020 - Aug 08, 2020</td>
<td>Live Event</td>
</tr>
<tr>
<td>SANS FOR508 Canberra August 2020</td>
<td>Canberra, AU</td>
<td>Aug 17, 2020 - Aug 22, 2020</td>
<td>Live Event</td>
</tr>
<tr>
<td>SANS Amsterdam August 2020 Part 2</td>
<td>Amsterdam, NL</td>
<td>Aug 17, 2020 - Aug 22, 2020</td>
<td>Live Event</td>
</tr>
<tr>
<td>SANS Virginia Beach 2020</td>
<td>Virginia Beach, VAUS</td>
<td>Aug 31, 2020 - Sep 05, 2020</td>
<td>Live Event</td>
</tr>
<tr>
<td>SANS Philippines 2020</td>
<td>Manila, PH</td>
<td>Sep 07, 2020 - Sep 19, 2020</td>
<td>Live Event</td>
</tr>
<tr>
<td>SANS London September 2020</td>
<td>London, GB</td>
<td>Sep 07, 2020 - Sep 12, 2020</td>
<td>Live Event</td>
</tr>
<tr>
<td>SANS Baltimore Fall 2020</td>
<td>Baltimore, MDUS</td>
<td>Sep 08, 2020 - Sep 13, 2020</td>
<td>Live Event</td>
</tr>
<tr>
<td>Threat Hunting &amp; Incident Response Summit &amp; Training 2020</td>
<td>New Orleans, LAUS</td>
<td>Sep 10, 2020 - Sep 17, 2020</td>
<td>Live Event</td>
</tr>
<tr>
<td>SANS OnDemand</td>
<td>OnlineUS</td>
<td>Anytime</td>
<td>Self Paced</td>
</tr>
<tr>
<td>SANS SelfStudy</td>
<td>Books &amp; MP3s OnlyUS</td>
<td>Anytime</td>
<td>Self Paced</td>
</tr>
</tbody>
</table>