ITL Bulletin

ADVISING USERS ON INFORMATION TECHNOLOGY

SECURING VOICE OVER INTERNET PROTOCOL (IP) NETWORKS

By Thomas J. Walsh and D. Richard Kuhn
National Institute of Standards and Technology

Voice over IP (VOIP) - the transmission of voice over traditional packet-switched IP networks - is one of the hottest trends in telecommunications. As with any new technology, VOIP introduces both opportunities and security challenges. Lower cost and greater flexibility are among the promises of VOIP for the enterprise, but security administrators will face significant issues. Administrators may assume that since digitized voice travels in packets, they can simply plug VOIP components into their already-secured networks and expect a stable and secure voice network. Unfortunately, many of the tools used to safeguard today's computer networks, namely firewalls, Network Address Translation (NAT), and encryption, don't work "as is" in a VOIP network.

VOIP systems take a wide variety of forms. Just about any computer is capable of providing VOIP, and most users don't realize that they already have basic VOIP applications. Microsoft's NetMeeting, or the newer Windows Messenger, which come with Windows platforms, provides voice and video services, and Linux platforms have a number of VOIP applications from which to choose. In general, though, the term Voice Over IP is associated with equipment that provides the ability to dial telephone numbers and communicate with parties on the other end who may have either another VOIP system or a traditional analog telephone. Demand for VOIP services has resulted in a broad array of products, including:

- Traditional telephone handset - Usually these products have extra features beyond a simple handset with dial pad. Some of these units may have a "base station" design that provides the same convenience as a conventional cordless phone.
- Conferencing units - These provide the same type of service as conventional conference calling phone systems, but since communication is handled over the Internet, they may allow users to coordinate traditional data communication services, such as a whiteboard that displays on computer monitors at both ends.
- Mobile units - Wireless VOIP units are becoming increasingly popular, especially since many organizations already have an installed base of 802.11 networking equipment. Wireless VOIP products present particularly acute security problems, given the well-known weaknesses of the 802.11 family of protocols.
- PC or "softphone" - With a headset, software, and inexpensive connection service, any PC or workstation can be used as a VOIP unit, often referred to as a "softphone."

In addition to end-user equipment, VOIP systems include specialized components beyond those found on an ordinary IP network: call managers and media/signaling gateways. Call managers are required to set up calls, monitor call state, handle number translation, and provide basic telephony services. Call managers also handle signaling functions that coordinate with media gateways, which are the interface between the VOIP network and the public switched telephone network (PSTN). Depending on the system, gateway functions may be implemented as a board or dedicated appliance, or may be provided through a distributed system of servers and databases.

Current VOIP systems use one of two protocols, H.323 or the Session Initiation Protocol (SIP). SIP is the Internet Engineering Task Force (IETF) specified protocol for initiating a two-way

Continued on page 2

ITL Bulletins are published by the Information Technology Laboratory (ITL) of the National Institute of Standards and Technology (NIST). Each bulletin presents an in-depth discussion of a single topic of significant interest to the information systems community. **Bulletins are issued on an as-needed basis** and are available from ITL Publications, National Institute of Standards and Technology, 100 Bureau Drive, Stop 8900, Gaithersburg, MD 20899-8900, telephone (301) 975-2832. To be placed on a mailing list to receive future bulletins, send your name, organization, and business address to this office. You will be placed on this mailing list only.

Bulletins issued since August 2003

communication session. It was designed to be simpler than H.323, but has become increasingly complex, as the standard has evolved. SIP is text based; its messages are similar to e-mail message formats. Also, SIP is an application level protocol, that is, it is decoupled from the protocol layer it is transported across. Unlike H.323, SIP uses only one port in the call setup process. The architecture of a SIP network also differs from the H.323 structure. A SIP network is made up of end points, a proxy and/or redirect server, location server, and registrar. In the SIP model, a user is not bound to a specific host. Instead, users initially report their location to a registrar, which may be integrated into a proxy or redirect server.

H.323 is the International Telecommunication Union (ITU) specification for audio and video communication across packetized networks. H.323 acts as a wrapper for a suite of media control recommendations by the ITU incorporating several other protocols, including H.225 and H.245. Each of these protocols has a specific role in the call setup process, and all but one make use of dynamic ports. An H.323 network is made up of several endpoints (terminals) that are normally bound to a specific address, a gateway, and possibly a gatekeeper, multipoint control unit, and back end service. The gateway serves as a bridge between the H.323 network and the outside world of (possibly) non-H.323 devices, including SIP networks and traditional PSTN networks.

Most VOIP components have counterparts used in data networks, but the performance demands of VOIP mean that ordinary network software and hardware must be supplemented with special VOIP components. One of the main sources of confusion for those new to VOIP is the assumption that because digitized voice travels in packets just like other data, existing network architectures and tools can be used with little or no change. Unfortunately, VOIP adds a number of complications to existing network technology, and these problems are compounded by security considerations.

What's Different About VOIP Security?

To understand why security for VOIP isn't the same as data network security, we need to look at both the unique constraints of transmitting voice over a packet network, and at characteristics shared by VOIP and data networks. Packet networks depend on a large number of configurable parameters: IP and media access control (MAC) (physical) addresses of voice terminals, addresses of routers and firewalls. VOIP networks add specialized software such as call managers and other programs used to place and route calls. Many of the network parameters are established dynamically every time a network component is restarted, or when a VOIP telephone is restarted or added to the network. Because there are so many places in a VOIP network with dynamically configurable parameters, intruders have as wide an array of potentially vulnerable points to attack as they have with data networks. But VOIP systems have much stricter performance constraints than data networks, with significant implications for security.

Quality of Service (QoS) is fundamental to the operation of a VOIP network. A VOIP application is much more sensitive to delays than its traditional data counterparts. If one downloads a file, a slowdown of a few seconds is negligible. In contrast, a delay of merely 150 *milli*seconds is enough to turn a crisp VOIP call into a garbled, unintelligible mess. In the VOIP vernacular, this is termed the *latency* problem.

Latency turns traditional security measures into double-edged swords for VOIP. Tools such as encryption and firewall protection can help secure the network, but they also introduce a significant amount of delay. Latency is not just a quality of service issue, but a security issue as well, because it increases the system's susceptibility to a Denial of Service (DoS) attack. For a DoS attack to succeed in a VOIP network, it need not completely shut down the system. It must only delay voice packets for a fraction of a second. The necessary impediment is even less when latency-producing security devices are slowing down traffic.

Another QoS issue, *jitter*, refers to non-uniform delays that can cause packets to arrive and be processed out of sequence. Real-time Transport Protocol (RTP), the protocol used to transport voice media, is based on the User Datagram Protocol (UDP), so packets received out of order cannot be reassembled at the transport level, and therefore must be reordered at the application level, introducing a significant overhead. Even when packets manage to arrive in order, high jitter causes them to arrive at their destination in spurts. This scenario is analogous to uniform road traffic coming to a stoplight. As soon as the stoplight turns green (bandwidth opens up), traffic races through in a clump.

Infrastructure issues become significant with a change to VOIP. With conventional telephones, eavesdropping requires either physical access to tap a line or penetration of a switch. Attempting physical access increases the intruder's risk of being discovered, and conventional private branch exchanges (PBXs) typically use proprietary protocols, specialized software, and have fewer points of access than VOIP systems. With VOIP, opportunities for eavesdroppers are multiplied. VOIP units share physical network connections with the data network, and in many cases, VOIP and data are on the same logical portion of the net-

Who we are

The Information Technology Laboratory (ITL) is a major research component of the National Institute of Standards and Technology (NIST) of the Technology Administration, U.S. Department of Commerce. We develop tests and measurement methods, reference data, proof-of-concept implementations, and technical analyses that help to advance the development and use of new information technology. We seek to overcome barriers to the efficient use of information technology, and to make systems more interoperable, easily usable, scalable, and secure than they are today. Our website is http://www.itl.nist.gov/.

work. Protocols are standardized, and tools to monitor and control packet networks are widely available. Attaching a packet sniffer, such as the freely available "voice over misconfigured internet telephony" (known by its unfortunate acronym "vomit"), to the VOIP network segment makes it easy to intercept voice traffic.

Like other types of software, VOIP systems have been found to have vulnerabilities due to buffer overflows and improper packet header handling. Exploitable software flaws typically result in two types of vulnerabilities: denial of service or disclosure of critical system parameters. In some cases, the system can be crashed, producing a memory dump in which an intruder can find IP addresses of critical system nodes, passwords, or other security-relevant information. Crashing a VOIP server may also result in a restart that restores default passwords or falls prey to a rogue server attack. In addition, buffer overflows that allow the introduction of malicious code have been found in VOIP software, as in other applications.

Tradeoffs between convenience and security are routine in software, and VOIP is no exception. Most, if not all, VOIP components use integrated web servers for configuration. Web interfaces can be attractive, easy to use, and inexpensive to produce because of the wide availability of good development tools. Unfortunately, most web development tools are built with features and ease of use in mind, with less attention to the security of the applications they help produce. VOIP device web applications have been discovered with weak or no access control, script vulnerabilities, and inadequate parameter validation, resulting in privacy and denial of service vulnerabilities. As VOIP gains in popularity, with implementations on devices of all types, it is almost inevitable that more administrative web applications with exploitable errors will be found.

ITL Bulletins Via E-Mail
We now offer the option of delivering your ITL Bulletins in ASCII format directly to your e-mail address. To subscribe to this service, send an e-mail message from your business e-mail account to listproc@nist.gov with the message **subscribe itl-bulletin**, and your name, e.g., John Doe. For instructions on using listproc, send a message to listproc@nist.gov with the message **HELP**. To have the bulletin sent to an e-mail address other than the From address, contact the ITL editor at 301-975-2832 or elizabeth.lennon@nist.gov.

What do the Special Characteristics of VOIP Mean for Security?

Meeting the security challenges of VOIP can require changes to a number of familiar security components. Firewalls are a staple of security in today's IP networks. Whether protecting a local-area network (LAN), a wide-area network (WAN), encapsulating a demilitarized zone (DMZ), or just protecting a single computer, a firewall is usually the first line of defense. Firewalls work by blocking traffic deemed to be malicious or potentially risky. Acceptable traffic is determined by a set of rules programmed into the firewall by the network administrator. These may include such commands as "Block all FTP traffic (port 21)" or "Allow all http traffic (port 80)." Much more complex rule sets are available in almost all firewalls. Firewalls also provide a central location for deploying security policies, the ultimate bottleneck for network traffic, because no traffic can enter or exit the LAN without passing through the firewall.

This situation lends itself to the VOIP network where firewalls simplify security management by consolidating security measures at the firewall gateway, instead of requiring all the endpoints to maintain up-to-date security policies. This takes an enormous burden off the VOIP network infrastructure. Unfortunately, this abstraction and simplification of security measures comes at a price. The introduction of firewalls to the VOIP network complicates several aspects of VOIP, most notably dynamic port trafficking and call setup procedures. Several commercial solutions are available to alleviate this including Application Level Gateways (ALGs), that make the firewall "VOIP-aware," and Midcom Controls, which allow the firewall to be traversed by allowing it to receive instruction from an application-aware agent. That is, they can understand the VOIP protocol data carried as a payload within an ordinary packet, making it possible to do stateful filtering of call packets. Attempting to implement a VOIP system on a legacy network without such devices is generally not feasible.

Firewalls, gateways, and other such devices can help keep intruders from compromising a network. However, these devices are no defense against an internal hacker and don't protect voice data as it crosses the Internet. Another layer of defense is necessary at the protocol level to protect the data itself. In VOIP, as in data networks, this can be accomplished by encrypting the packets at the IP level using Internet Protocol Security (IPsec). This way, if anyone intercepts VOIP traffic and is not the intended recipient (for instance, via a packet sniffer), such packets would be unintelligible. The IPsec suite of security protocols and encryption algorithms is the standard for securing packets against unauthorized viewers over data networks and will be supported by the protocol stack in IPv6. So it seems logical to extend IPsec to VOIP, encrypting the signal and voice packets on one end and decrypting them only when needed by their intended recipient. Unfortunately, the nature of the signaling protocols and the VOIP network itself make it necessary for routers, proxies, and other components to read the VOIP packets, so encryption is often done at the gateways to a network, rather than the endpoints. Such a scheme also allows the endpoints to be computationally simple and promotes scalability as new encryption algorithms can be overlaid on the network without upgrading the endpoints. Several factors, including the expansion of packet size, ciphering latency, and a lack of QoS urgency in the cryptographic engine itself, can cause an excessive amount of latency in the VOIP packet delivery. This leads to degraded voice quality, so once again there is a tradeoff between security and voice quality, and a need for speed.

Virtual private network (VPN) tunneling of VOIP has also become popular recently, but the congestion and bottlenecks associated with encryption suggest that this solution may not always be scalable. Although great strides are being made in this area, the hardware and software necessary to ensure call quality for encrypted voice traffic may not be economically or architecturally viable for all enterprises considering the move to VOIP.

What are the Prospects for Securing a VOIP Network?

Thus far, we have painted a fairly bleak picture of VOIP security. The construction of a VOIP network is an intricate procedure that should be studied in great detail before being attempted. Integrating a VOIP system into an already congested or overburdened network could be disastrous for an organization's technology infrastructure. There is no easy "one size fits all" solution to the issues discussed in this bulletin. The use of VPNs, versus ALG-like solutions and the choice of SIP or H.323 are decisions that must be made based on the specific nature of the current network and the VOIP network to be. However, the technical problems are solvable, and the establishment of a secure implementation of VOIP is well worth the difficulty associated with these solutions. To implement VOIP securely today, start with these general guidelines, recognizing that practical considerations may require adjustments for the organization:

- Put voice and data on logically separate networks. Different subnets with separate RFC 1918 address blocks should be used for voice and data traffic, with separate DHCP servers for each, to ease the incorporation of intrusion detection and VOIP firewall protection.
- At the voice gateway, which interfaces with the PSTN, disallow H.323, SIP, or Media Gateway Control Protocol (MGCP) connections from the data network. Use strong authentication and access control on the voice gateway system, as with any other critical network management component.
- A mechanism to allow VOIP traffic through firewalls is required. There are a variety of protocol-dependent and independent solutions, including ALGs for VOIP protocols, Session Border Controllers, or other standards-based solutions. Stateful packet filters can track the state of connections, denying packets that are not part of a properly originated call.
- Use IPsec or Secure Shell (SSH) for all remote management and auditing access. If practical, avoid using remote management at all and do IP PBX access from a physically secure system.
- Use IPsec tunneling when available instead of IPsec transport because tunneling masks the source and destination IP addresses. This secures communications against rudimentary traffic analysis (i.e., determining who is calling each other).
- If performance is a problem, use encryption at the router or other gateway, not the individual endpoints, to provide for IPsec tunneling. Since some VOIP endpoints are not computationally powerful enough to perform encryption, placing this burden at a central point ensures all VOIP traffic emanating from the enterprise network has been encrypted. Newer IP phones are able to provide Advanced Encryption Standard (AES) encryption at a reasonable cost.
- Look for IP Phones that can load digitally (cryptographically) signed images to guarantee the integrity of the software loaded onto the IP Phone.
- "Softphone" systems, which implement VOIP using an ordinary PC with a headset and special software, should be avoided, if possible, where security or privacy are a concern. In addition to violating the separation of voice and data, PC-based VOIP applications can be vulnerable to worms and viruses that are all too common on PCs, and may infect other parts of the network.
- Consider methods to "harden" any VoIP platform based on common operating systems such as Windows or Linux. This includes disabling unnecessary services and possibly using host-based intrusion detection methods.
- Be especially diligent about maintaining patches and current versions of VOIP software.
- Analyze the impact of VOIP adoption on the rest of the organization's infrastructure, including issues such as backup power, E-911 emergency location, and records retention policies or other legal issues.

VOIP can be done securely, but the path is not smooth. It will likely be several years before standards issues are settled and VOIP systems become a mainstream commodity. Until then, organizations should proceed cautiously and not assume that VOIP components are just more peripherals for the local network. Above all, it is important to keep in mind the unique requirements of VOIP, acquiring the right hardware and software to meet the challenges of VOIP security. For more information on securing VOIP systems, see draft NIST Special Publication 800-58, *Security Considerations for Voice Over IP Systems*, at http://csrc.nist.gov/publications/nistpubs/index.html.

Disclaimer: *Any mention of commercial products or reference to commercial organizations is for information only; it does not imply recommendation or endorsement by the National Institute of Standards and Technology nor does it imply that the products mentioned are necessarily the best available for the purpose.*

U.S. DEPARTMENT OF COMMERCE
National Institute of Standards and Technology
100 Bureau Drive, Stop 8900
Gaithersburg, MD 20899-8900