

H.323 Mediated Voice over IP: Protocols, Vulnerabilities & Remediation

Dr. Thomas Porter 2004-06-01

An IBM executive was once quoted as saying, "Our goal is to make the computer as easy to use as the telephone". Our goal, now, is the reverse: to make using an IP telephone at least as easy and secure as using a computer on the Internet.

Voice over IP (VoIP) can be a complex subject. Network security professionals may find the terminology foreign, and VoIP vulnerabilities are often misunderstood. This paper provides an overview of the H.323 protocol suite, its known vulnerabilities, and then suggests twenty rules for securing an H.323-based network.

VoIP Protocols

VoIP protocols can be classified according to their role during message transmission. H.323 and SIP are signaling protocols and thus, they are involved in call setup, teardown, and modification. RTP (real-time transport protocol) and RTCP (real-time transport control protocol) are media transport protocols, and are involved in end-to-end transport of voice and multimedia data. TRIP, SAP, SRP, OSP, et. al. comprise a group of VoIP-related support protocols. Finally, because H.323 mediated VoIP relies upon the underlying transport layer to move data, more traditional protocols that security professionals are familiar with, such as TCP/IP, DNS, DHCP, SNMP, RSVP, and TFTP, may be required.

H.323 Overview

H.323 (which is implemented primarily at versions 3 and 4 as of the time of this writing) is a byzantine international protocol, published by the [ITU](#), that supports interoperability between differing vendor implementations of telephony and multimedia products across IP-based networks. H.323 entities provide real-time audio, video and/or data communications. Support for audio is mandatory, while support for data and video is optional.

Media gateways

The primary components of an H.323 network include: endpoints, gateways, gatekeepers, and MCUs (Multipoint Control Units). Endpoints (telephones, softphones, IVRs, voicemail, video cameras. etc.) are the devices typically used by end-users in the

normal use of the system. Gateways (gateways and controllers) handle signaling and media transport, and typically serve as the interface to other types of networks such as ISDN, PSTN and or other H.323 systems. Gateways which focus primarily on converting between IP and other forms of media (such as PSTN) are termed *Media Gateways*. Gatekeepers are the logical entity with which endpoints register and are administered. They also manage call setup, teardown, and status and can assist in address resolution. MCUs are designed to support multi-party conferencing.

The dissection of H.323-mediated VoIP from a security point-of-view, is complex - the plethora of associated protocols & the large number of vendor implementations (at least 40 separate vendor implementations) has resulted in further complicating the interactions and different security features of the vendor implementations. SIP also has a number of security issues, but those will be addressed in another article.

As will be seen below, a surfeit of attacks against the H.323 protocols can be envisioned. The only existing vulnerabilities that we are aware of at this time take advantage of ASN.1 parsing defects in the first phase of H.225 data exchange. More vulnerabilities can be expected for several reasons: the large number of differing vendor implementations, the complex nature of this collection of protocols, problems with the various implementations of ASN.1/PER encoding/decoding, and the fact that these protocols, alone and in concert, have not endured the same level of scrutiny that other, more common protocols have been subjected to.

H.323 Signaling protocols

H.323 allows dissimilar communication devices to communicate with each other by using a standardized communication protocol. H.323 defines a common set of call setup and negotiating procedures and basic data-transport methods - the most common in VoIP applications being H.225.0, H.235, H.245, H.248, and the Q.900 signaling series. In addition, for VoIP communications H.323 specifies a group of audio codecs - the G.700 series. The following is an overview of these major protocols.

- **H.225** describes standards for Call Signaling Protocols (CSPs) and Media Stream Packetization. H.225/Q.931 call signaling is used to initiate connections between H.323 endpoints, over which the real-time data can be transported. H.225 messages are in binary ASN.1 PER (Packed Encoding Rules) format. The signaling channel is opened between an endpoint-gateway, a gateway-gateway, or gateway-gatekeeper prior to the

establishment of any other channels. Although the H.225.0 signaling channel may be implemented on top of UDP, all entities must support signaling over TCP port 1720. H.225 also defines messages used for endpoint-gatekeeper and gatekeeper-gatekeeper communication - this part of H.225 is known as "RAS" (Registration, Admission, Status), and, unlike CSPs, runs over UDP.

- **Q.931** is originally a Layer 3 protocol of ISDN. A subset of this standard is used in H.323 in the primary call signaling channel. It carries PER-encoded H.225 call signaling messages as a payload.
- **H.235** recommends an assortment of messages, procedures, structures and algorithms for securing signaling, control and multimedia communications under the H.323 architecture.
- **H.245** describes a set of call control protocols. After a connection has been set up via the call signaling procedure, H.245 messages (there are many of these) are used to resolve the call media type, to exchange terminal capabilities, and to establish the media flow before the call can be established. H.245 messages also manage call parameters after call establishment. H.245 messages are encoded in ASN.1 PER syntax. The messages carried include notification of terminal capabilities, and commands to open and close logical channels. The H.245 control channel is permanently open, unlike the media channels.
- **H.248** (also known as Megaco) is the international standard for media gateway control. H.248 addresses the relationship between the Media Gateway (MG), which converts circuit-switched voice to packet-based voice (the MG handles the media) , and the Media Gateway Controller (or softswitch), which dictates the service logic of that traffic - that is, it manages call signaling and other non-media-related functions
- Other protocols are also defined within the H.323 standard. These include, but are not limited to: the H.260 series, the H.450-series, the H.460 series, the T.120 series, T.140, and the H.320 series.

ASN.1 (Abstract Syntax Notation One) is commonly misunderstood. It is not a programming language, but it is a flexible notation that allows one to define a variety of data types. ASN.1 encoding rules are sets of rules used to transform data specified in the ASN.1 language into a

standard format that can be decoded on any system that has a decoder based on the same set of rules. The H.323 protocol family is compiled into a wire-line protocol using PER (Packed Encoding Rules). PER is a compact binary encoding that is used on limited-bandwidth networks. It is designed to optimize the use of bandwidth, but the tradeoff is complexity: decoding PER PDUs (protocol data units) has led to problems due to a number of factors including issues with octet alignment, integer precision, and unconstrained character strings.

H.323 Messaging sequence

H.323 signaling exchanges typically are routed via the gatekeeper or directly between the participants as chosen by the gatekeeper. Media exchanges are normally directly routed between the participants of a call.

Normally, the first message components used to initiate an H.323 exchange are Gatekeeper Discovery packets. Establishing a call between two endpoints requires two TCP connections between the endpoints: one for call setup (Q.931/H.225 messages), and one for capabilities exchange and call control (H.245 messages). First, an endpoint initiates an H.225/Q931 exchange on a TCP well-known port (TCP 1720) with another endpoint. Successful completion of the "call" results in an end-to-end reliable channel supporting H.245 messaging.

H.245 negotiations usually take place on a separate channel from the one used for H.225 exchanges, but newer applications support tunneling of H.245 PDUs in the H.225 signaling channel. There is no well-known port for H.245. The H.245 transport address is always passed in a call-signaling message. The media channels (those used to transport voice and video) are similarly dynamically-allocated. As an aside, this use of dynamic ports makes it difficult to implement security policy on firewalls, NAT, and traffic shaping.

H.323 data communications utilizes both TCP and UDP. TCP ensures reliable transport for control signals and data, because these signals must be received in proper order and cannot be lost. UDP is used for audio and video streams, which are time-sensitive but are not as sensitive to an occasional dropped packet. Consequently, the H.225 call signaling channel and the H.245 control channel typically run over TCP, while audio, video, and RAS channel exchanges rely on UDP for transport.

RTP and RTCP

Real-time transport protocol (RTP) is an application layer protocol that provides end-to-end

delivery services of real-time audio and video. RTP provides payload identification, sequencing, timestamping, and delivery monitoring. UDP provides multiplexing and checksum services. RTP can also be used with other transport protocols.

The actual media, such as voice, first needs to be encoded using an appropriate codec. The encoded audio stream is then passed to RTP, which is used to transfer the real-time audio/video streams over the Internet. RTCP provides status and control information for the use of RTP.

Real-time transport control protocol (RTCP) is the counterpart of RTP that provides control services. The primary function of RTCP is to provide feedback on the quality of the data distribution. Other RTCP functions include carrying a transport-level identifier for an RTP source, called a canonical name, which can be used by receivers to synchronize audio and video.

RTP runs on dynamic, even-numbered, high ports (ports above 1024), while RTCP runs on the next corresponding odd numbered, high port.

Vulnerabilities

H.225 (denial of service; execution of code)

Recently, the University of Oulu Secure Programming Group (OUSPG) tested the effects of sending modified Setup-PDUs (see below) to a number of differing vendor H.323 implementations. Modified Setup-PDUs are TCP/IP packets that carry the H.225/Q.931 initial signaling information (protocol identifier, source address, called number, etc.) encoded according to ASN.1 PER (Packed Encoding Rules). The H.225 Setup-PDU is an excellent test candidate for several reasons: The Setup-PDU contains many information elements, whose length and type are variable; The Setup PDU is normally the first packet exchanged during H.323 communication; and affected systems can be quickly rebooted for additional testing. OUSPG prepared a test suite containing approximately 4,500 modified Setup-PDUs, and fed these to each tested H.323 device. They found that many systems that implement H.323 are vulnerable to one or more of these malformed PDUs -- affected devices typically crashed or experienced 100% CPU utilization.

These failures result from insufficient bounds checking of H.225 messages as they are parsed and processed by affected systems. These errors are primarily due to problems in low-level byte operations with vendor ASN.1 PER/BER PDU decoders, as mentioned

earlier. Depending upon the affected system and implementation, these attacks result in system crash and reload (DoS), or in the case of systems that parse these data (such as Microsoft ISA server), execution of code within the context of the security service.

Additionally, we have found that flooding multiple, malformed GRQ (Gatekeeper Request) packets to the gatekeeper results in the disconnection of a number of vendor's IP phones.

H.245

Another obvious candidate for vulnerability testing is the first of the many H.245 messages - the Terminal Capability Set (TCS) message. The TCS message occurs early in the H.245 exchange so that the calling party can determine the version and capabilities of the corresponding H.245 endpoint. Work is in progress in this area. Vulnerabilities to H.245 have not been announced yet but are anticipated.

Issues with remediation

The H.323 protocol suite is complex, but provides a great deal of flexibility. Protocol-specific problems will be addressed in a similar manner as problems with traditional protocols -- through testing and independent audit followed by remediation. Unfortunately there are still issues that cannot be easily addressed using traditional security devices found in typical organizations.

Firewall issues

As you have seen, many H.323 protocols (particularly those involving the data streams) are made up of dynamic IP address/port combinations. Each terminal-terminal conversation requires, at a minimum, 5 channels to be opened - 2 control channels (one H.225 and one H.245) per endpoint, plus one bidirectional voice channel. Three of these will be on dynamically-allocated ports. In addition, users naturally expect to be able to make both inbound and outbound calls. Since H.323 relies heavily on dynamic ports, packet filtering firewalls are not a viable solution, as every port above 1024 would have to be opened. Therefore, most firewall solutions supporting H.323 must at least disassemble the control stream packets (H.245, H.225.0) and then dynamically (and ideally, statefully) open up the firewall as needed. Additionally, H.323 contains embedded IP address information that is not re-written by most NAT implementations.

All of these features make the implementation of H.323 security complex.

Data encryption

Payload or data encryption is another important piece of the H.323 security puzzle, but in most cases, the ability of an attacker to access one or more control channels will yield information about a call that is almost as valuable as the data content. Forty years ago, phreakers whistled or yelled into telephones and compromised the signaling channel in order to make free phone calls. Today, analyses of a signal channel, for instance, could allow an attacker to gather information regarding the duration, endpoints and other parameters of incoming and outgoing calls. Lastly, data encryption has no effect on the security of VoIP-related infrastructure devices or their cognate applications. No amount of encryption can protect against a single bad password.

Twenty rules for securing H.323 networks

In light of the above issues, there are still numerous steps that can be taken to secure VoIP infrastructure. The key to securing H.323 networks is to use and enforce the security mechanisms already deployed in data networks (firewalls, encryption, etc.). This allows one to emulate the security level currently enjoyed by PSTN network users.

1. Update existing security policies, practices, and procedures to reflect the new requirements of converged networks.
2. Distribute, communicate and enforce these policies.
3. Ensure that all networked systems are patched, and virus scanners are up to date.
4. Install and monitor IDS, IPS, and Honeypots.
5. Exercise diligence in analyzing logs from intrusion detection systems, firewalls, routers, servers and other network devices.
6. Secure off-site backups and develop disaster recovery plans.
7. Encrypt H.323 traffic.
8. Utilize VPNs wherever possible.
9. Employ H.323-ready firewalls and other appropriate protection mechanisms, and properly configure them.
10. Consider segmenting voice and data traffic by using a virtual LAN. This will limit the threat posed by packet-sniffing tools and minimize disruption in the event of an attack.
11. If possible, run the VoIP network on a separate physical network. Converge your network

management tools.

12. Employ different subnets with separate RFC 1918 address blocks for voice and data traffic, with separate DHCP servers for each.
13. Utilize proxy servers or Application Layer Gateways (ALGs) in front of corporate firewalls to process incoming and outgoing voice data.
14. Utilize Session Border Controllers (SBCs) at the VoIP network edge if H.323 encryption is used.
15. Use strong authentication and access control on the voice gateway systems.
16. Harden server-based IP PBX operating systems.
17. Test backup power solutions.
18. Forge strong relationships with your ISP's in order to defend against external DoS attacks.
19. Use IPSec or Secure Shell (SSH) for all remote management and auditing access.
20. If practical, avoid using remote management at all and access IP PBXs from a physically secure system.

The author would like to thank Anton Rager, Brian Boyter, and Rick Robinson for their helpful discussions in preparing this article.

About the author

[Thomas Porter](#), Ph.D. is a Senior Security Consultant with Avaya's Enterprise Security division. He has spent over eleven years in the networking and security industry as a consultant, speaker and developer of security tools; he also holds numerous security certifications. Tom lives in Chapel Hill, North Carolina with his wife.

[Privacy Statement](#)

Copyright 2006, SecurityFocus