WHITE PAPER

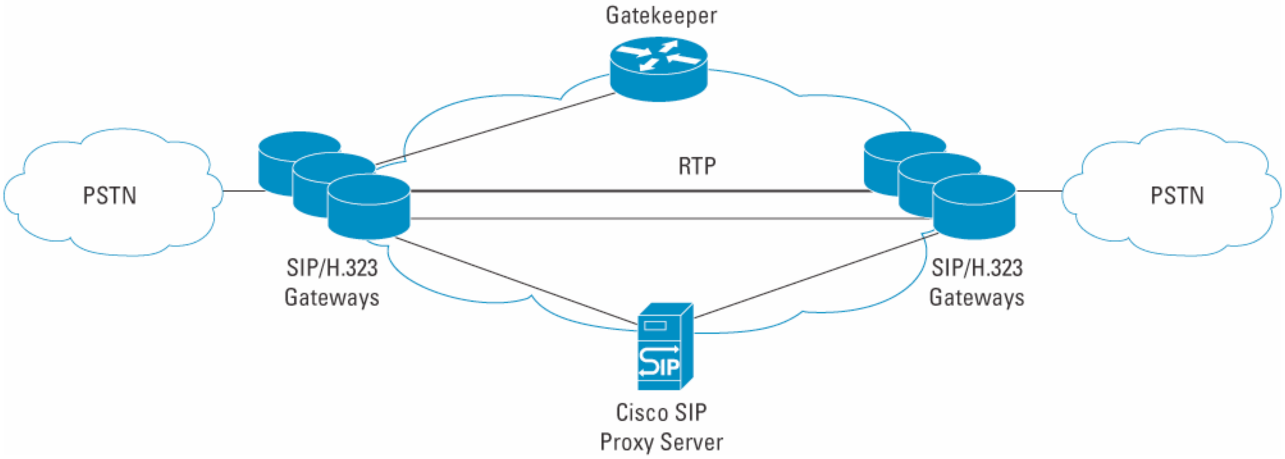
MITIGATING ATTACKS IN VOIP ENVIRONMENTS

Wide-scale voice over IP (VoIP) implementations based on Session Initiation Protocol (SIP) and H.323 are gaining traction and are starting to be heavily deployed in large enterprises (Figure 1). VoIP is becoming an equivalently pervasive solution in the small office/home office (SOHO) market.

Service providers are deploying packet telephony technologies as an alternative to traditional circuit-switched telephone networks in order to shift traditional voice services to packet-based networks and to create new services that combine data, voice, and video information. The lower cost associated with converged data and voice networks is a prime rationale for deployment of packet telephony.

With the widespread deployments of VoIP helping businesses and individuals manage and accomplish critical day-to-day tasks, the infrastructure supporting VoIP implementation must remain highly secure and available. A small amount of downtime can result in millions of dollars of lost revenue and greater customer support costs. The building blocks of a VoIP network—the endpoints, gateways, and the gatekeepers—must have hardened protocol implementations and be protected against possible denial of service (DoS) attacks, takeovers, or misuses.

Figure 1. SIP and H.323 Wholesale Call Transport



VOIP ARCHITECTURE

The VoIP infrastructure helps connect multiple PSTN and IP-based telephony endpoints. It consists of SIP/H.323 gateways that help the endpoints establish connections by providing them with conversion of the data format, directory services, and call signaling services. While SIP is growing in popularity due to its use of (and similarity to) Internet technologies such as HTTP, H.323 is more widely deployed and the standard is more mature.

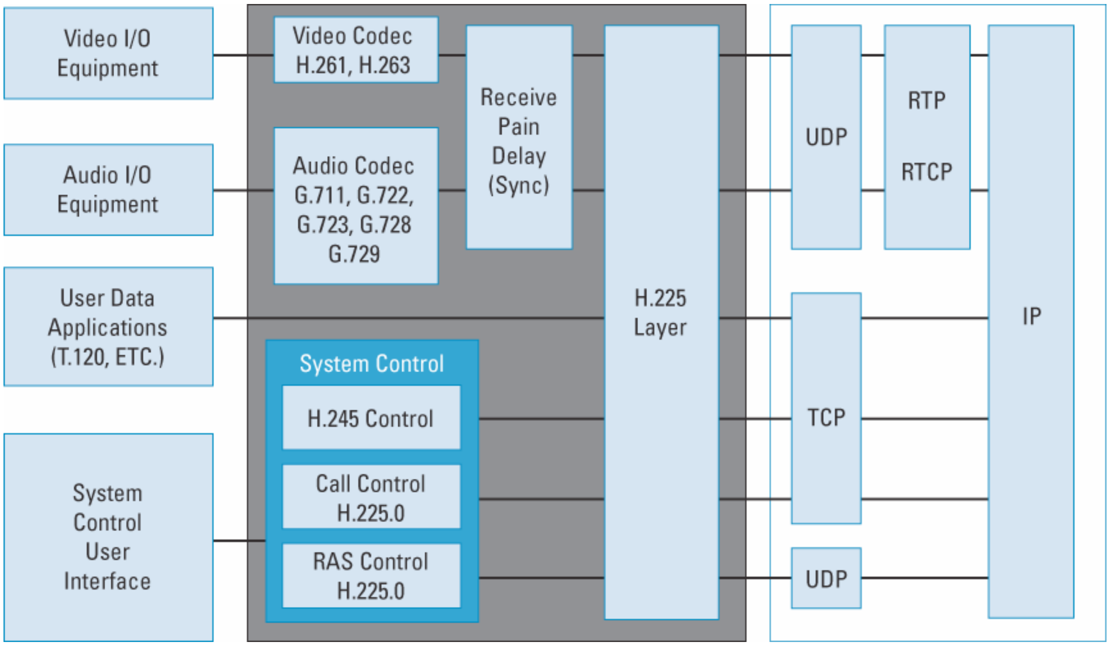
One of the main advantages of H.323-based networks is the ability to manage available resources for call routing via H.323 gatekeepers. The gatekeeper is the logical “switch” of the H.323 network, providing several basic services to all endpoints in its zone.

Gatekeeper services include address translation (alias name/number-to-network address), endpoint admission control (based on bandwidth availability, concurrent call limitations, or registration privileges), bandwidth management, and zone management (the routing of calls originating or terminating in the gatekeeper zone, including multiple path reroute). Gateways coordinate calls by communicating with gatekeepers using the Registration, Admission, and Status (RAS) Protocol.

H.323 OVERVIEW

H.323 is a collection of protocols and other standards that together enable conferencing over packet-based networks (Figure 2). H.323 is embraced by major vendors, and implementations range from desktop applications to industrial-grade gateways.

Figure 2. H.323 Protocol Suite Overview



H.225 OVERVIEW

The H.225 call signaling protocol consists of many subprotocols and, as seen earlier, is part of the H.323 suite. H.225 is used for connection establishment and termination between endpoints. The H.225 call signaling protocol also supports status inquiry, ad hoc multipoint call expansion, and limited call forwarding and transfer. H.225 call signaling messages are exchanged over Q.931. The Q.931 messages are exchanged over a TCP stream demarcated by Transport Protocol Data Unit Packet (TPKT) encapsulations. The H.225 call signaling message is transported as part of the user information element of the Q.931 protocol. The ASN.1 representation of the H.225 message is encoded using the Packed Encoding Rules.

Figure 3. Basic Constituents of a H.323 Call

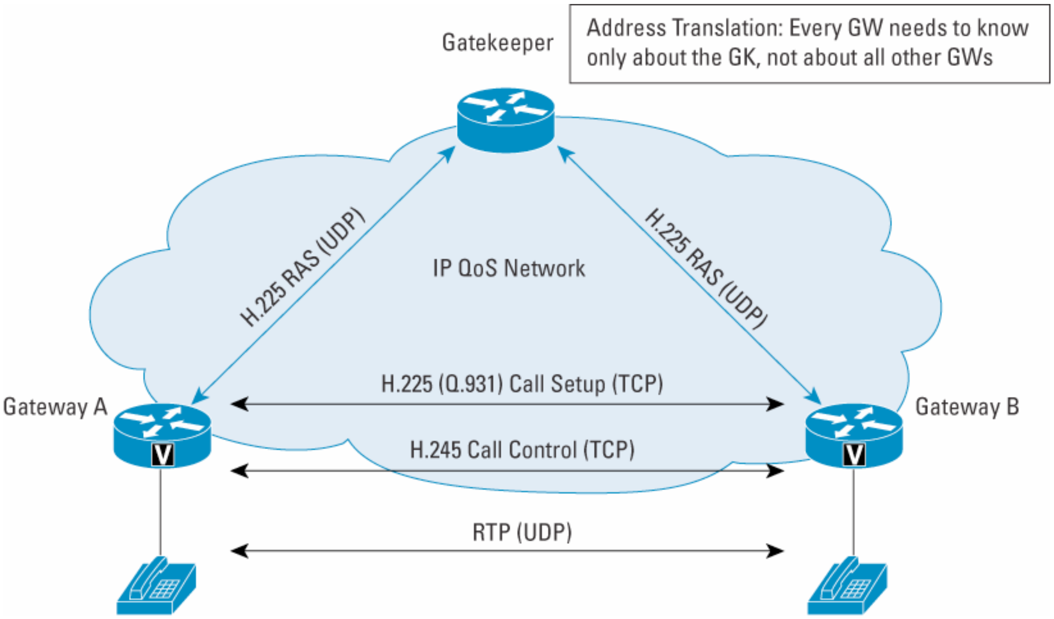
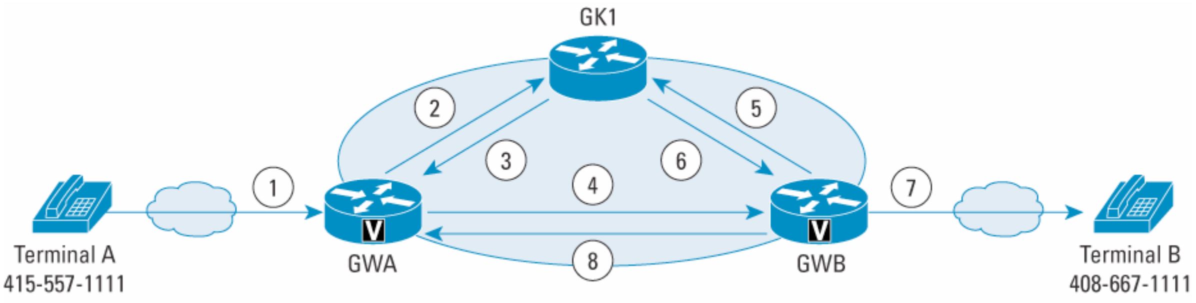

Gatekeeper
Address Translation: Every GW needs to know only about the GK, not about all other GWs
H.225 RAS (UDP)
IP QoS Network
H.225 RAS (UDP)
Gateway A
H.225 (Q.931) Call Setup (TCP)
Gateway B
H.245 Call Control (TCP)
RTP (UDP)

Figure 3 represents a basic message sequence and the constituents of an H.323 call. The general approach to starting a call is to send a mandatory admission request on the RAS channel, followed by an initial setup message on a reliable channel transport address (this address may have been returned in the admission confirmation message, or may have been known to the calling terminal). As a result of this initial message, a call setup sequence commences based on H.225 call signaling operations. The sequence is complete when in the "Connect" message, the terminal receives a reliable transport address on which to send H.245 control messages.

Note: A reliable transport address is used for call setup for the terminal-to-terminal case, and also for the gatekeeper-mediated case. The reliable call signaling connection is kept active until a "Release Complete" message is received for all active calls signaled over the call-signaling channel.

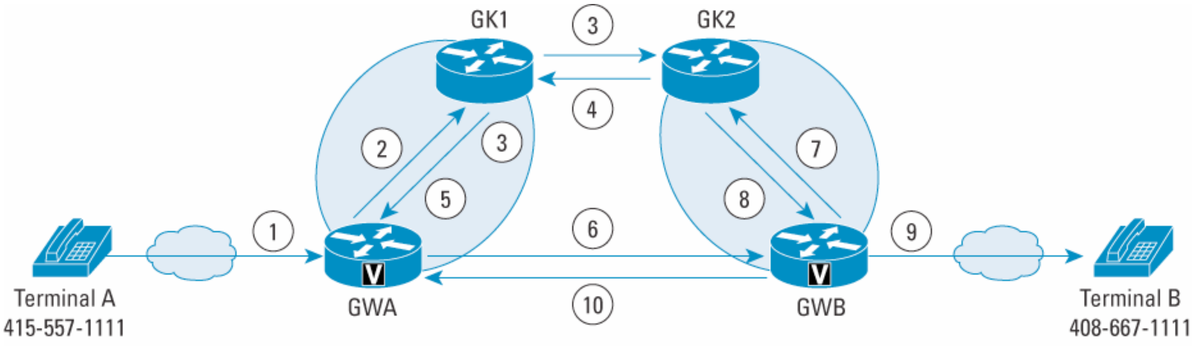
Figures 4 and 5 depict the message sequence of a normal inter- and intra-zone call-message sequence.

Figure 4. Intra-Zone Call Message Sequence



1) Terminal A dials the phone number 408-667-1111 for Terminal B
2) GWA sends GK1 an ARQ, asking permission to call Terminal B
3) GK1 does a lookup and finds Terminal B registered; returns an ACF with the IP address of GWB
4) GWA sends a Q.931 Call-Setup to GWB with Terminal B's phone number
5) GWB sends GK1 an ARQ, asking permission to answer GWA's call
6) GK1 returns an ACF with the IP address of GWA
7) GWB sets up a POTS call to Terminal B at 408-667-1111
8) When Terminal B answers, GWB sends Q.931 Connect to GWA
9) GWs sends IRR to GK after call is set up

Figure 5. Inter-Zone Call Message Sequence



1) Terminal A dials the phone number 408-667-1111 for Terminal B
2) GWA sends GK1 an ARQ, asking permission to call Terminal B
3) GK1 does a lookup and does not find Terminal B registered; GK1 does a prefix lookup and finds a match with GK2; GK1 sends an LRQ GK2, and RIP (Request In Progress) to GWA
4) GK2 does a lookup and finds Terminal B registered; returns an LCF with the IP address of GWB
5) GK1 returns an ACF with the IP address of GWB
6) GWA sends a Q.931 Call-Setup to GWB with Terminal B's phone number
7) GWB sends GK2 an ARQ, asking permission to answer GWA's call
8) GK2 returns an ACF with the IP address of GWA
9) GWB sets up a POTS call to Terminal B at 408-667-1111
10) When Terminal B answers, GWB sends Q.931 Connect to GWA

H.225 SECURITY CONSIDERATIONS

As shown in earlier figures, H.225 call signaling and status messages form an inherent part of the H.323 call setup. Various H.323 entities in the network like the gatekeeper, gateways, and endpoint terminals run implementations of the H.225 protocol stack. In scenarios like this, it becomes increasingly important to have robust implementations of these protocols and to have proper security checks to avoid protocol misuse and allow attackers to use bugs in these implementations as attack vectors. Attackers can try and compromise the H.225 protocol implementations; it is possible to adversely affect the VoIP network, hijack calls, or lead to misuse of the VoIP network.

Buffer Overflow Attacks

Since H.225 messages are PER encoded, the attacker can misencode the PER encoding lengths and try and cause buffer overflow at the receiving endpoint. The ASN.1 representation of the H.225 protocol lays down some specific bounds on the lengths of the fields, and protocol modules may be susceptible to attacks based on these fields.

DoS Attacks

Attackers can try and send huge messages by specifying out-of-bound and large messages or fields. This leads to excessive memory usage at the endpoints and gateways and can lead to a DoS attack. The attackers can try to use PER encoding coupled with the ASN.1 representation to encode excessive recursive fields and lead to huge processing and memory overhead at the endpoint.

Invalid Protocol Fields/Misuse

Attackers may use a vulnerability in the endpoint implementation by sending invalid protocol fields, or may misuse the misinterpretation of endpoint software. This can lead to inadvertent leakage of sensitive network topology information, call hijacking, or a DoS attack.

Attacks Using Bad Patterns in String Fields

Attackers may use certain string fields in the Q.931 and H.225 protocols to insert specific patterns and compromise the endpoint implementation to run specific attack code, like opening a back door for further attacks.

THE CISCO IPS SENSOR SOFTWARE VERSION 5.0 H.225 ENGINE

To protect different H.225 implementations and provide for a single point of misuse or attack detection for H.225 implementations in the network, an IPS is an ideal solution.

Because the call-signaling messages are exchanged over TCP PDUs, they need deep-packet inspection. An IPS can be ideal to detect such attacks on multiple H.323 gatekeepers, VoIP gateways, and endpoint terminals by sitting at the edge of the network. An IPS can also act as a central point for various policy enforcements on H.225 messages coming into the network.

In order to use the IPS as a detection/protection point, to protect against attacks by specially crafted invalid H.225 messages, and to protect against the misuse and overflow attacks on various protocol fields in these messages, we need to analyze the H.225 protocol. By analyzing the messages and fields and applying static and user-tunable signature checks for the protocol implementation, an IPS can provide customers with a solution to protect their H.323 implementations against such attack vectors. Especially important is the SETUP call-signaling message; this is the first message exchanged between H.323 entities as part of the call setup. The SETUP message uses many of the commonly found fields in the call-signaling messages, and implementations that are exposed to probable attacks will also likely fail the security checks for the SETUP messages. As a result, it is highly important to check the H225 SETUP message for validity and to enforce checks on the perimeter of the network.

Mitigation of VoIP Threats with the H323/H225 Engine Delivered With Cisco IPS v5.0 Sensor Software

Cisco IPS v5.0 Sensor Software utilizes advanced algorithms to accurately identify and stop threats in a VoIP environment using the following techniques:

TPKT Validation and Length Checks

For TCP streams, checks on the format of the TPKT (RFC 1006), version number, and maximum length are performed. This helps protect the gateway from very large TPKTs or bad TPKT length attacks, which in turn helps to ensure the sanity of the TPKT fields and that the TPKT length is within the bounds defined by the policy.

Field/Message Validation (Empty or Absent Fields)

Validation in terms of empty or absent information elements is performed. The actual H.225 message is also checked for adherence to the protocol specification in terms of field presence, length, and other criteria.

Information Elements Validation and Length Checking on Information Elements

Checks lengths and validations like presence/absence of information elements; checks for reserved values on Q.931 information elements for SETUP messages. This helps ensure the sanity of the SETUP message and that the SETUP message is not being misused.

Setup Message Validation

SETUP message checks, including presence/absence of actual H.225 SETUP message payload, making sure that the user information element is the last one in the SETUP message, and making sure that the total length of the actual H.225 SETUP message is valid, are performed to minimize the possibility of invalid or too-large SETUP messages reaching the endpoints/gateways.

Regexp Pattern Searches

Regexp pattern search capabilities on various SETUP message string components (URLs and e-mail IDs) and Q.931 string components (display information) are implemented to allow the system administrator to selectively apply policies and signatures on fields that are subject to attack based on the endpoint implementations.

Address Checks

Using the pattern search capabilities and length check signatures, address fields like e-mail ID, URL ID, and H.323 ID can have policies applied. These policies are highly detailed; they allow applying policies based on the above fields occurring in specific places, and not all occurrences of the address fields across the messages.

Signature to Track the Maximum Number of SETUP Messages Seen on a Single Call-Signaling Connection

This specific check is incorporated to support limiting the total number of SETUP messages that are seen in a call-signaling connection. This can help to alleviate or stop DoS attacks based on SETUP messages.

ASN.1 PER validations

The H.225 engine facilitates catching PER encoding errors like illegal or unknown choice/sequence of a field, ASN constraint violations and the message being too short for complete protocol decode.

Other Features Supported by the H323/H225 Engine in Cisco IPS v5.0 Sensor Software

- Built-in support for the inspection of fragmented H.225 messages across TPKTs, and the presence of multiple H.225 messages in the same TPKT.

All in all, the IPS H.225 engine helps the network administrator ensure that the SETUP message coming into the VoIP network is valid and within the bounds of the installed policies. It also helps ensure that the addresses and Q.931 string fields like URL IDs, e-mail IDs, and display information adhere to specific lengths and do not have possible attack patterns in them.

With the inbuilt signatures for TPKT validation, Q.931 protocol validation, and ASN.1PER validations for the H.225 SETUP message, the IPS H.225 engine is ready to be used right out of the box. The Q.931 and TPKT length signatures are tunable by the user to easily accommodate the needs of the protocol implementation. Also, it is flexible enough to be able to add and apply detailed signatures on specific H.225 protocol fields and to apply multiple pattern search signatures of a single field in the Q.931 or H.225 protocol.

In summary, with the increased dependence on VoIP applications in today's networks, businesses are evermore exposed to threats that specifically target voice protocols and VoIP infrastructure. Attacks are expected to become increasingly sophisticated and attack tools more widely accessible. The effects of such attacks generally lead to disruption of VoIP services that invariably result in loss of revenue. In order to maintain business continuity and maximize up time of critical VoIP applications, Cisco IPS v 5.0 Sensor Software delivers a comprehensive implementation of a H323/H225 Engine that accurately classifies VoIP threats and stops such attacks through a variety of automated inline IPS response actions.

Corporate Headquarters
Cisco Systems, Inc.
170 West Tasman Drive
San Jose, CA 95134-1706
USA
www.cisco.com
Tel: 408 526-4000
800 553-NETS (6387)
Fax: 408 526-4100

European Headquarters
Cisco Systems International BV
Haarlerbergpark
Haarlerbergweg 13-19
1101 CH Amsterdam
The Netherlands
www-europe.cisco.com
Tel: 31 0 20 357 1000
Fax: 31 0 20 357 1100

Americas Headquarters
Cisco Systems, Inc.
170 West Tasman Drive
San Jose, CA 95134-1706
USA
www.cisco.com
Tel: 408 526-7660
Fax: 408 527-0883

Asia Pacific Headquarters
Cisco Systems, Inc.
168 Robinson Road
#28-01 Capital Tower
Singapore 068912
www.cisco.com
Tel: +65 6317 7777
Fax: +65 6317 7799

Cisco Systems has more than 200 offices in the following countries and regions. Addresses, phone numbers, and fax numbers are listed on **the Cisco Web site at www.cisco.com/go/offices.**

Argentina • Australia • Austria • Belgium • Brazil • Bulgaria • Canada • Chile • China PRC • Colombia • Costa Rica • Croatia • Cyprus Czech Republic • Denmark • Dubai, UAE • Finland • France • Germany • Greece • Hong Kong SAR • Hungary • India • Indonesia • Ireland Israel • Italy • Japan • Korea • Luxembourg • Malaysia • Mexico • The Netherlands • New Zealand • Norway • Peru • Philippines • Poland Portugal • Puerto Rico • Romania • Russia • Saudi Arabia • Scotland • Singapore • Slovakia • Slovenia • South Africa • Spain • Sweden Switzerland • Taiwan • Thailand • Turkey • Ukraine • United Kingdom • United States • Venezuela • Vietnam • Zimbabwe

204107_ETMG_MH_11.04

Printed in the USA