INTEROP® LABS

SIP and Firewalls

A firewall is a device that stands astride a user's pathway to the Internet that denies entry and exit to all but acceptable forms of traffic.

SIP based Voice over IP (VoIP) is not naturally friendly to firewalls or Network Address Translation (NAT).

Firewalls open and close pathways depending on the real-time needs of user applications. Very few applications communicate their intentions directly to firewalls. Instead, firewalls must observe application traffic and indirectly derive the application's intention and open or close the appropriate holes through the firewall.

For application protocols that use well-known ports, such as HTTP, this is often a relatively easy task that can be handled by static rule sets that indicate what layer 3 (IP) protocols, ports, and addresses may pass and which may not.

But for protocols that assign ports dynamically, particularly protocols based on UDP, the firewall has to inspect application-level protocol information. In effect the firewall has to become aware of the nature and operation of application layer protocols. This adds considerable complexity to the firewall software and can consume considerable amounts of firewall memory and CPU.

Firewall rules typically presuppose that the only services available to "the outside" are those provided by a short, and static, list of well known servers (such as web and email servers) on well known IP addresses and ports.

SIP, however, turns every internal phone into a device that may potentially be called from outside phones. This means that SIP aware firewalls may have to evaluate dynamic rules whenever an incoming call arrives. Or the SIP aware firewall may delegate that responsibility to a specialized SIP device (which may be physically embedded into the firewall box) that may, in turn, impose a VoIP-specific security policy.

Firewalls are a technology separate and distinct from Network Address Translation. NATs are really about translating IP address and TCP or UDP port information from one IP address space to another. Firewalls and NATs are often integrated into a single device.

Although firewalls and NATs often have to do similar kinds of application protocol tracking the way they respond to that information is significantly different. Firewalls generally do not rewrite packet contents, NATs do. NATs are not really security devices, but they are perceived as such. This note deals with the interaction of SIP and firewalls. This does not mean that SIP does not have issues with NATs – quite the contrary – but this is a topic for another note.

SIP + RTP/RTCP

Let us back up for a moment and recognize that when we talk about SIP based VoIP we are really talking about a bundle of separate protocols. SIP itself is but one protocol in this bundle. SIP is a call-setup (signaling) protocol. SIP does not itself carry the voice or video media. SIP depends on the RTP/RTCP protocol (and variations for security, such as SRTP) to move the actual media content.

SIP uses the well-known UDP and TCP port (5060) and is thus easy for firewalls to intercept. However, SIP carries several types of IP address and port information that the firewall must examine. And because of the multiplicity of ways that SIP can format this data this examination can be difficult.

The RTP/RTCP protocol is not anchored to any well-known fixed port. The UDP port used for any particular RTP/RTCP stream is dynamically assigned by SIP. Until the SIP transactions are parsed the firewall does not know which packets contain the RTP/RTCP streams carrying a call's media. In addition, the firewall must continue to monitor the SIP activity in order detect SIP RE-INVITES and know when the call has terminated.

Tracking SIP To Discover UDP Ports To Be Used By RTP/RTCP

A SIP-capable firewall must follow and comprehend each SIP exchange and extract from each exchange the RTP/RTCP port information that is being negotiated. The nature of SIP requires that the firewall continue to monitor the SIP exchanges through the lifetime of the call – The firewall must continue to monitor the SIP conversation in case the call is modified by subsequent SIP activity (as might occur with a call that is forwarded to voice-mail or processed by an interactive-voice-response [IVR] menu system.)

Unfortunately SIP does not carry this information in a nice easy-to-digest format. Instead the media stream and port information is spread over multiple SIP packets that occur near the start of the call and may be revised in subsequent SIP packets that may occur during the life of the call.

Moreover, the way that SIP encodes data further complicates the life of a firewall. SIP is a text protocol that has blended techniques used in e-mail (SMTP) and HTTP and then added enough redundant encoding options and variations to choke a camel.

As if this were not enough, some SIP devices have adopted mechanisms, such as STUN, to detect the presence of NATs and pre-modify SIP packet contents in anticipation of NAT processing.

SIP is flexible – it can run over the reliable, sequenced, connection-oriented TCP or it can run over the unreliable, non-sequenced, connectionless UDP. Most SIP devices of today use UDP. This means that a firewall must anticipate that the SIP packets for a call might be lost, duplicated, reordered, or delayed. But even when SIP is run over TCP, the firewall must partially replicate TCP processing in order to reform the packets into a sequenced stream.

SIP is overgrown with complexity; the IETF has not yet pruned SIP to be a thing of streamlined simplicity. Firewalls, because they have to take whatever is thrown at them, must be very robust and engineered to accept and handle all of SIP's complexity.

All in all, a SIP aware firewall must be a creature rather well endowed with intelligence, capacity, agility, and reliability.

The Effect of Encryption

Several security protocols and secure tunnels use encryption to protect the data carried by IP packets from being observed or altered as it flows from source to destination. Firewalls, because they need to inspect that data, become blind and unable to function when the SIP packets are encrypted.

Firewall-Like Capabilities In SIP and RTP/RTCP Specific Devices

If we move above the level of the classical firewall we find a number of specialized SIP devices that can be adapted to enforce an enterprise's security policies.

SIP can be used as a peer-to-peer protocol or it can be driven through one or more intermediary devices.

SIP based phones can be coerced so that they operate through intermediary devices that may enforce a security policy. These intermediary devices are generally known as a SIP Registration Server and a SIP Proxy server. These are often found in the same physical device and they often coordinate their activities.

Phones can be configured to make outgoing calls via a SIP proxy. The proxy may block or allow calls and may force the phone to register itself with a Registration Server.

The Registration Server, also known as a Presence Server, is used to direct incoming calls to the proper phone. The Registration Server may evaluate the phone's credentials and chose whether to accept the registration request.

As mentioned previously, SIP is a signaling protocol; the real work of carrying the media is done by RTP/RTCP. There is no requirement that the RTP/RTCP packets flow along the same paths as the SIP packets. In fact the SIP and RTP/RTCP packets for a call often follow quite separate network pathways.

The media flows may be directed by SIP to rendezvous at specialized media servers – such as conference call bridges – that are capable of exercising security controls.

Summary

SIP and RTP/RTCP form a very flexible system. Classical firewalls are only one tool among many for imposing security policy on SIP based VoIP. These tools may be bundled by a vendor into a single physical box.

The enforcement tools for SIP and RTP/RTCP are complex and varied. But tools are just that: tools. They must enforce a policy. If experience with telephone dial plans is any guide, we can anticipate that an enterprise's VoIP security policy will be forever evolving and often very intricate.