SIP and Security

Enterprises using IP-based telephony will expect a level of reliability, privacy, authenticity, and accountability on par with traditional analog networks. With this entrance into the production environment, all of these security considerations, so often forgotten in the lab, take on renewed importance. Properly securing a VoIP network based on SIP requires the consideration of many factors. Because SIP, itself a signaling protocol, is separate from RTP, the protocol used to carry audio and video media streams, each of these protocols must be secured separately. Both SIP and RTP have protocol elements that can be used to provide access control, data integrity, and confidentiality. However, these features must be separately enabled in a secure deployment. Your choice of what elements and features to use will depend on your security requirements and network architecture.

In addition to internal protocol features, most security architects are very concerned with border security, both within the enterprise and as IP-based telephony extends across the Internet. Two strategies for border security are commonly used. One is to make use of SIP-aware firewalls, NAT devices, and a new breed of devices called Session Border Controllers to control and manage VoIP traffic as it passes network boundaries. An alternative approach is to avoid the whole problem by keeping all SIP devices, both end stations and proxies, within your administrative control, using VPN technologies to extend that control to remote locations.

SIP-Aware Security Devices

Firewalls tend to be the first device that comes to mind when the topic of security comes up. Traditionally, firewalls have simple rule sets that allow communication to specific hosts at defined ports. With the dynamic nature of the RTP media streams, this approach is insufficient for SIP. Like protocol such as FTP, SIP requires an application layer gateway (ALG) in firewalls for proper operation. The ALG watches the protocol and dynamically opens and closes ports through the firewall as needed for both SIP and RTP messages to pass. Because SIP is much more complex than FTP, ALGs in firewalls for SIP have only recently been available.

NAT devices, while technically not security devices, are commonly perceived as such and are so ubiquitously deployed that they must be considered. NAT and SIP interact in several ways. First, the SIP protocol itself carries IP addresses of the end stations in the call setup messages. When the endpoint embeds its own address in SIP messages, this address will be useless once it passes through the NAT, and calls cannot be completed. Secondly, while any NAT device will be able to setup a translation for SIP control traffic, since it is initiated from within the private side, a NAT device without SIP features will not be able to create the necessary translations for the inbound RTP traffic. This leads to the scenario of being able to set up a call, but not having either no audio or one-way audio.

A SIP-aware NAT device solves these problems by changing embedded private addresses in call setup messages to be external (non-NATed) public address. The NAT also can also see the RTP session description, and create the proper translations to allow the RTP media stream through. Unfortunately, most consumer grade NAT devices aren't SIP aware, so many SIP endpoints have implemented a NAT-Traversal mode. In this mode, the endpoint is manually configured with the public address of the NAT, and it uses this address in the signaling. Also, once the endpoint behind the NAT has completed call setup, it sends a RTP packet to "prime" the NAT for the incoming media stream. In cases where the public address of the NAT is unpredictable, other mechanisms have been created to workaround that problem. For example, some SIP endpoints support the STUN (Simple Traversal of UDP through NAT) protocol for this purpose. NAT interoperability is an area of considerable concern to SIP developers and service providers.

A Session Border Controller (SBC) is a device that resides at the border between an enterprise and the Internet, effectively a specialized firewall for SIP. It provides a single point of transit for SIP communications, acting as a relay for both for call control and media streams. Conceptually similar to an external email relay or DNS server, the SBC provides a single point for the application of policy, protocol validation, and detection of malicious SIP traffic. SBCs are valuable tools in environments where a large amount of SIP traffic will pass between the corporate network and the general Internet.

Using SIP and RTP protocol features for Security

The SIP protocol itself defines a few mechanisms to provide access control and call control security. Access control is done through the SIP REGISTER mechanism. When a SIP endpoint initially connects to a SIP proxy, it sends a REGISTER message with identifying information. While this REGISTER can be completely unauthenticated, limited access control is provided by a cleartext password, or for somewhat greater security, an MD5 digest.

SIP call control messages from a registered endpoint can be secured at two different levels. At the highest level, Secure MIME (S/MIME) can be used to provide end-to-end integrity and confidentiality. At the network transport layer, various hop-by-hop security mechanisms can be negotiated, such as TLS (SSL) or IPsec.

In addition to the SIP signaling, the actual media streams may need to be secured. RTP has an extension, Secure RTP (SRTP), which can add confidentiality, message authentication, and replay protection for the media streams. SRTP is specifically designed to be low overhead, both from a bandwidth and computational perspective. The use of SRTP is negotiated as part of the SIP INVITE message.

Combining these protocol features with SIP-aware security devices provides a great deal of flexibility in achieving both security goals and successful administrative application of policy. However, the biggest issue surrounding all of these mechanisms is the level of implementation and interoperability in SIP products. In our interoperability testing, we did not find that most products had implemented security on either the SIP call control messages or on the RTP media stream.

Avoiding the problem altogether

While the mechanisms mentioned above can be combined in various ways to secure just about any conceivable network topology, they add complexity and cost to the overall SIP deployment. Also, at this point in the development of SIP devices, full implementation and interoperability of all of these mechanisms is far from a given. Hence, many enterprise environments have opted to bypass the problem entirely. To do this, they've ensured that all SIP related traffic remain within the enterprise's secured network. When deploying to remote locations, be it remote offices or the homes of tele-workers, common VPN technology such as IPsec is used. Thus, SIP simply becomes trusted traffic within the border security of the enterprise.

This option is attractive and easy to implement, but also has inherent limitations. In the simplest form of this model, SIP traffic must stay within the enterprise, becoming a replacement for the traditional PBX, thus eliminating any potential benefits of using the Internet to carry voice traffic. This strategy should be considered a temporary avoidance, rather than a permanent solution, as other security issues, such as network availability, will eventually need to be addressed in any enterprise deployment.

How is SIP being secured in the iLabs?

In the iLabs, we have built a security testbed consisting of a number of different firewalls, a NAT device, and an array of phones and a SIP proxy behind them. This testbed allows us to demonstrate both control and media streams passing through one or more firewalls, as well as through a typical consumer-grade NAT device. Our tests range from placing a simple call to an end station beyond a single firewall to a very complex multi-way call to end stations behind different firewalls using a SIP proxy behind yet another firewall.

In general, most of our tests were successful, though there were definitely issues to resolve. For example, a phone behind the NAT device successfully registered with a SIP proxy, and could make outbound calls. However, inbound calls only worked if an outbound call had been made recently, reflecting the lack of SIP awareness in the NAT device. Our testing revealed the firewalls to be primitive in their SIP-specific debugging and logging capabilities as well. All in all, though, we found SIP-aware firewalls to have significantly matured in the last year and finally reached a level of reliability and interoperability required for enterprise deployment.

Pointers and References

Like most complex Internet protocols, SIP and RTP are defined and extended through a large set of RFC documents.. The security-related SIP RFCs in the following list can be found at `http://www.ietf.org.`

RFC3261: SIP: Session Initiation Protocol
RFC3853: S/MIME Advanced Encryption Standard (AES) Requirement for the Session Initiation Protocol (SIP)
RFC3329: Security Mechanism for the Session Initiation Protocol (SIP)
RFC3489: Simple Traversal of UDP through NAT (STUN)
RFC3550: RTP: A Transport Protocol for Real-Time Applications
RFC3711: The Secure Real-time Transport Protocol (SRTP)

In addition to the detailed specifications, anyone interested in SIP Security should read the excellent whitepaper on VoIP Security produced by the National Institute of Standards and Technology (NIST),
http://csrc.nist.gov/publications/nistpubs/800-58/SP800-58-final.pdf