Voice Over IP and Firewalls

By Mark Collier
Chief Technology Officer
SecureLogix Corporation
mark.collier@securelogix.com

Introduction

Use of Voice Over IP (VoIP) in enterprises is becoming more and more common. While most enterprise VoIP deployments are campus-based and internal, the move to Hosted VoIP/IP Centrex (e.g., for branch sites) and external VoIP trunking is just around the corner. These coming deployments will create problems for traditional data firewalls, which must be upgraded or augmented with VoIP firewalls.

Firewall Basics

A firewall is an-line device or software component that monitors traffic between a trusted and untrusted network. By being inline, a firewall can use a rule–based policy to determine which packets are allowed to pass through and which are not. A firewall may be a dedicated device, may be running on a device such as a router, or may be software running on a PC, such as a personal firewall. Firewalls can be deployed at various locations, but are almost always present on the enterprise connection to the Internet. Firewalls are used as the first line of defense against attacks originating from the Internet. In addition, firewalls may be placed on a dedicated Wide Area Network (WAN) link and at strategic locations inside the network to protect a critical server, such as an IP PBX.

By default, most firewalls deny all access. Rules are added to explicitly allow certain types of access. Rules are normally added to allow certain types of TCP/IP traffic, while virtually all UDP traffic is disallowed. TCP/IP traffic is generally viewed as more secure, because it is connection-oriented and more difficult to spoof. Inbound traffic is restricted to known types, such as email, web access, and name service, and only allowed to go to specific IP addresses. Outbound traffic is less restricted, but is normally confined to outbound web access to Internet web servers. The following figure illustrates the basic data flow for a firewall.

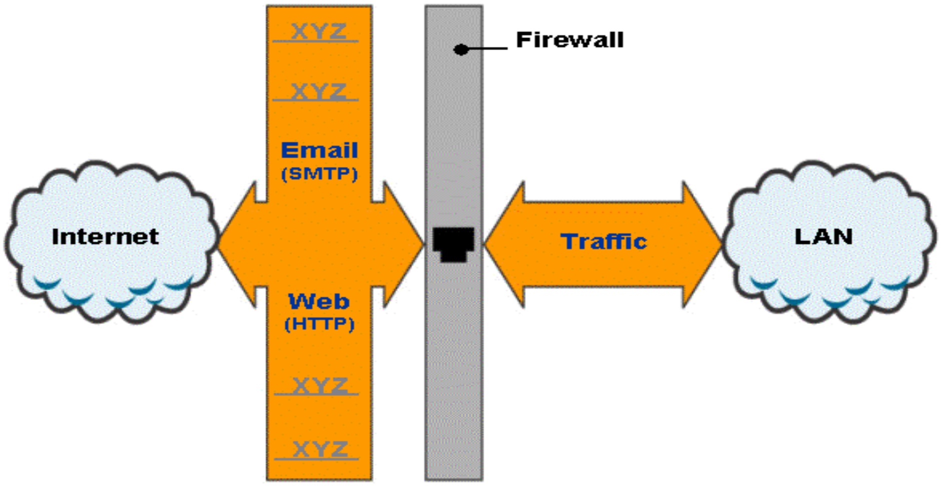

Figure 1 – Basic Firewall Data Flow

A firewall may include other applications, such as a Virtual Private Network (VPN), Intrusion Detection/Prevention System (IDS/IPS), content filter, anti-virus, etc. Firewalls used at small sites often combine these applications onto a single platform. At larger sites, these applications are generally present on separate platforms.

Firewalls commonly perform Network Address Translation (NAT). NAT was designed to preserve the limited IP space available with IP Version 4 (IPv4). NAT enables use of internal/private IP addresses, which share a limited number of public IP addresses and ports. There are various types of NATs, with symmetric NATs most common in enterprises. For those connections allowed through the firewall, the NAT converts internal address/port pairs to a separate external address/port pairs. This allows for example, a small organization with a single IP address (and many ports), to support many connections to the Internet. See the following figure for an example of NAT.

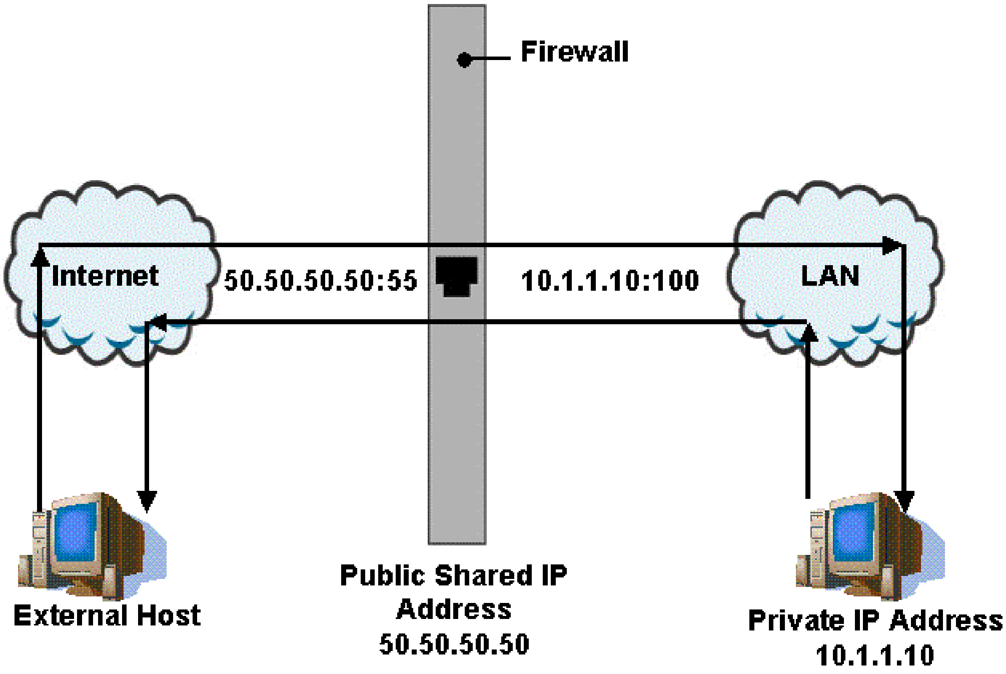
Firewall
Internet
50.50.50.50:55
10.1.1.10:100
LAN
External Host
Public Shared IP Address 50.50.50.50
Private IP Address 10.1.1.10

Figure 2 - Network Address Translation (NAT)

In this example, an internal host initiates a connection to a host on the Internet. The internal host selects a source IP/port of 10.1.1.10:100. As this address flows through the NAT, it is converted to a shared IP address and port, in this case, 50.50.50.50:55. As packets flow back, the opposite operation is performed.

Firewall Issues

This section describes issues you are likely to encounter when trying to transport VoIP through traditional data firewalls and NATs.

Separate Signaling Media and NAT

Most VoIP protocols use separate streams for signaling and media/audio. The signaling is normally exchanged on well-known ports (they must be well known to be found), while the media is carried over multiple dynamically assigned UDP ports. Although your firewall can be easily set up to allow signaling to and from an IP PBX with a known address and port, it can't handle the dynamic UDP ports. The firewall "could" open a wide range of UDP ports, but this is highly non-secure (some consumer VoIP products actually recommend this!). Another issue is that the dynamic UDP IP address/port numbers are embedded within the payload of VoIP signaling packets where a standard NAT can't find them. This causes the untranslated address to be passed to the external host, which prevents the media from reaching the correct host. The following figure illustrates these issues.

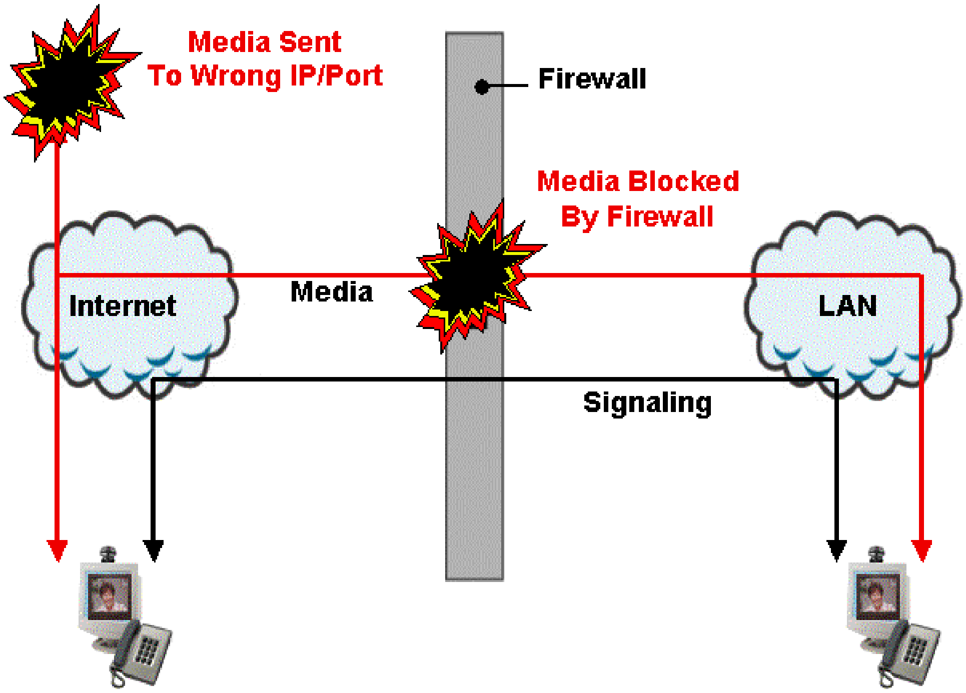

Figure 3 – Issues with Separate Signaling and Media Streams

As shown in the illustration above, the media UDP IP address/port sent to the external host is not translated, and therefore cannot reach the internal host. Also, because the firewall will normally block UDP traffic on dynamic ports, media that does reach the firewall will be blocked.

Signaling Encryption

VoIP equipment vendors are starting to offer encryption for signaling to provide strong authentication and privacy. Encryption of signaling greatly enhances security, but creates issues for NATs. If the VoIP signaling is encrypted, then the NAT cannot inspect the payload and convert addresses. A "solution" is to not use encryption through the NAT, but this is not desirable.

Performance and Reliability

Traditional data firewalls can create performance and reliability issues. Firewalls must process large numbers of packets from many sources. This traffic can create congestion at the firewall, resulting in delay, jitter (variable latency), and possibly dropped packets. This occurs because the firewall does not give priority to VoIP packets. This delay can degrade VoIP media audio quality (and in some cases signaling as well). The firewall may also create issues if it does not preserve Quality of Service (QoS) markings on packets. QoS markings are used to indicate high priority packets, and must be preserved so that other devices in the network know to prioritize processing of these packets.

Like any network device, firewalls will fail on occasion. Normally, when a firewall fails, all active sessions, including VoIP calls, will be torn down.

No Application Awareness

Most firewalls do not inspect the packets at the application layer and detect various types of application-level attacks, including registration hijacking, session tear down, Denial of Service (DoS), etc. Some firewalls do perform some generic application inspection, but other products, such as web server firewalls or email content filtering applications, provide comprehensive application-specific security. This is expected to be true as well for VoIP, due to the complexity of the application and wide number of protocols in use. The following figure illustrates this:

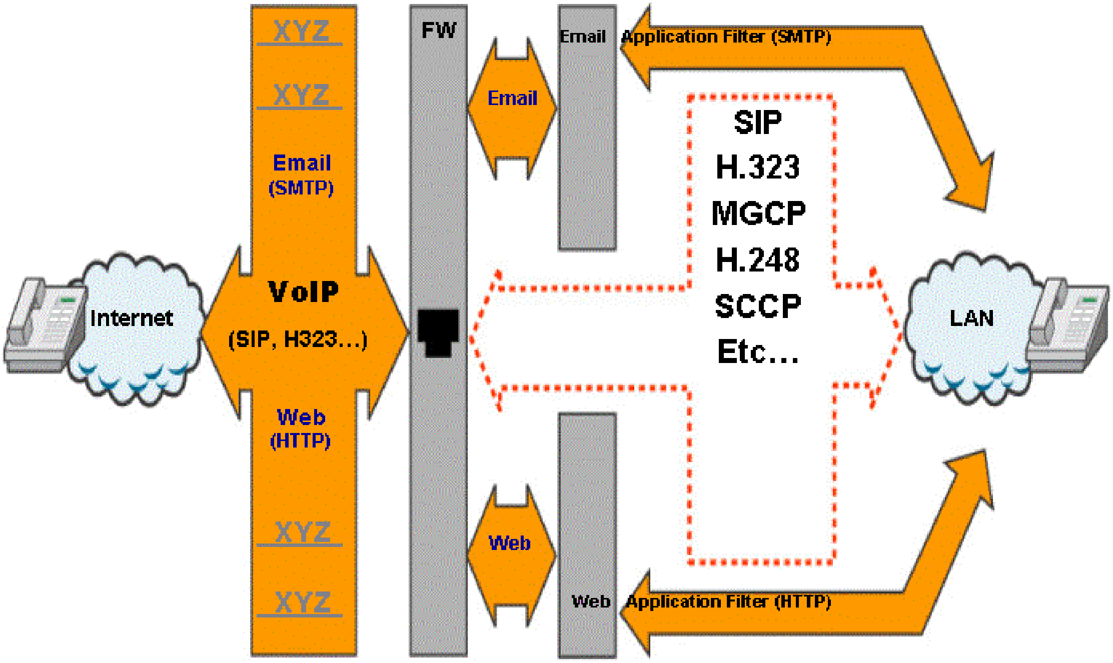

Figure 4 – Need for a Separate VoIP Firewall

Firewalls will also not detect application-service type attacks, such as toll fraud, harassing calls, and various types of fax and voice SPAM, often referred to as SPAM over Internet Telephony (SPIT).

No Protection for Circuit-Switched Networks

VoIP can be used as a means of attack against legacy circuit-switched voice networks. Traditional data firewalls do not provide any protection for circuit-switched networks. As consumers, service providers, and enterprises adopt more VoIP, this increases the threat of certain types of attacks, even to those enterprises still using circuit-switched access to the public network. More VoIP means more potential sources of attacks, including call floods and SPIT. These attacks are easier to originate with VoIP, and can affect a target circuit-switched network as easily as a VoIP network.

Recommendations

You can address the issues described above in one of two ways. Some traditional data firewalls will provide upgrades, which address some of the issues. An alternative is to use a VoIP-specific firewall to augment the data firewall and focus on VoIP-specific security. In either case, the following features must be provided:

- Monitor the signaling and perform NAT for the IP address/port pairs in the payload. For example, modifying values in the Session Description Protocol (SDP) in Session Initiation Protocol (SIP) signaling.
- As calls are started and ended, dynamically manage (open and close) UDP ports for the media sessions. These ports must be opened quickly, to prevent loss of initial media packets.
- Support the various protocols that traverse the firewall/NAT, including SIP, H.323, the Media Gateway Control Protocol (MGCP), and potentially, proprietary IP PBX vendor protocols.
- Enable use of encryption through the firewall/NAT. The firewall must either have access to the keys or act as a trusted proxy/Back to Back User Agent (B2BUA) and terminate and regenerate the encrypted signaling streams.
- Give priority to media packets. The firewall must maintain separate queues, to ensure that media packets are not delayed by other, non-real-time traffic being processed by the firewall. The firewall must also preserve any QoS markings on the media packets. This same treatment is necessary for certain signaling protocols, which involve man low-level messages to IP phones.
- Provide a design that is either fail-safe or redundant. A fail-safe design allows calls to be unaffected when the firewall fails (this can only be used when the firewall is not performing NAT or encryption termination). A redundant design involves use of two firewalls, which exchange state information and fail-over without dropping calls.
- Monitor signaling and media for packet-level and application-level attacks, such as registration hijacking, illegal teardowns, register floods, call floods, malformed packets, and other forms of DoS. The firewall must detect these attacks and discard illegal packets.
- Monitor for application-service level attacks, including toll fraud, harassing calls, and SPIT. Here the firewall is monitoring for abusive calls or call patterns, as opposed to packet-level attacks.
- Support appropriate functions for circuit-switched sites, which need protection from certain VoIP-originated attacks.

A firewall processing VoIP can be deployed at different points in the network:

- "In front" of an IP PBX–to protect it from internal threats from the Local Area Network (LAN).
- On the Internet connection–to monitor VoIP traffic coming from remote workers. Even users with a VPN can contract viruses, worms, etc., which could potentially tunnel through the VPN and affect the IP PBX. Remote users should also run personal firewalls to protect themselves from the Internet.
- On an IP Centrex/Hosted IP or external VoIP trunk–to detect attacks originating from the service provider network. When VoIP is exchanged with a service provider network, even if it is not the Internet, a firewall is still required to detect attacks originating from that network.

If a separate VoIP firewall is used to augment the traditional data firewall, there are multiple configurations available to allow the two to work together. The firewalls can run in series, in parallel, or in a Demilitarized Zone (DMZ) host configuration. The following figure illustrates these configurations:

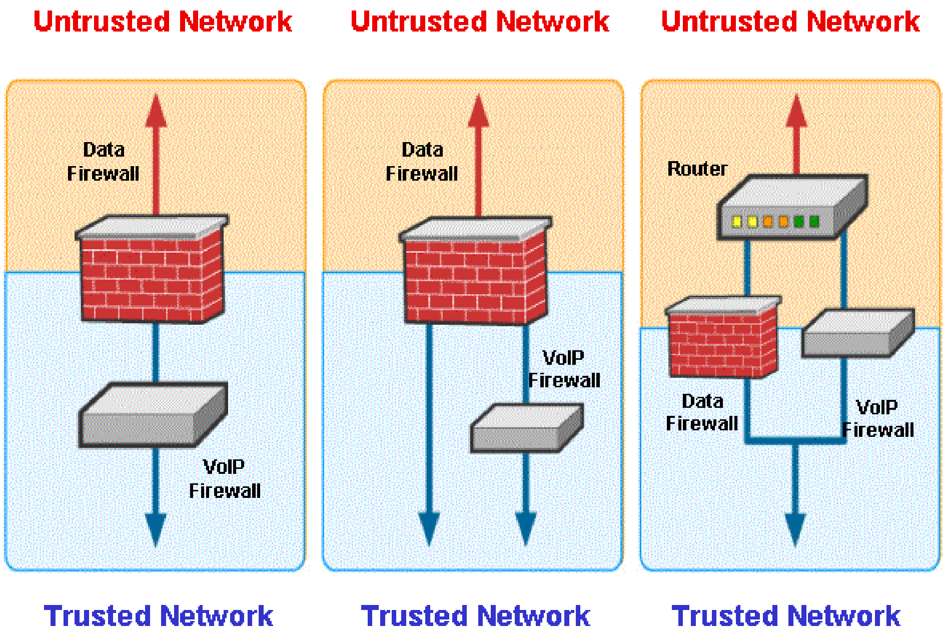

Figure 5 – VoIP Firewall Deployment Options

Conclusions

VoIP creates issues for traditional data firewalls. These issues are caused by the unique nature of VoIP, and include separate signaling/media, embedding IP addresses in signaling payload, use of encryption, performance/reliability requirements, lack of application awareness, and no protection for circuit-switched networks. As VoIP becomes commonly deployed, some traditional data firewalls will add support for VoIP. Alternatively, you have the option of using VoIP firewalls to address VoIP's unique issues.