Solving the Firewall/NAT Traversal Issue of SIP:

Who Should Control Your Security Infrastructure?

Ingate® Systems
www.Ingate.com
1 Executive Summary

Session Initiation Protocol (SIP) represents the third wave of Internet usage after SMTP (email) and HTTP (Web). Developed by the Internet Engineering Task Force (IETF), SIP has today become the signaling protocol of choice for establishing realtime communications, including Voice over IP (VoIP) calls. Research suggests that SIP is the VoIP protocol that has replaced H.323 and MGCP and that, for the foreseeable future, no replacement is expected (Business Communications Review, August 2005).

However, SIP-based communication does not reach users on the local area network (LAN) behind firewalls and Network Address Translation (NAT) routers automatically. Firewalls are designed to prevent inbound unknown communications and NAT stops users on a LAN from being addressed. Firewalls are almost always combined with NAT and typically still do not support the SIP protocol properly.

This issue of SIP traffic not traversing the enterprise firewall or NAT is critical to any SIP implementation, including VoIP. Eventually, all firewalls will need to be SIP capable in order to support the wide-scale deployment of enterprise person-to-person communications. In the interim, several solutions have been proposed to work around the firewall/NAT traversal problem. Several of these solutions have serious security implications while there are also solutions that allow you to remain in control. It is important to consider to what level you are prepared to surrender the control of your corporate infrastructure when choosing a NAT/firewall traversal solution.

The choice of method for traversing firewalls/NATs is, to a large extent, dependent on the answer to the questions: “Who should be in control of your security infrastructure: the firewall administrator, the user or a service provider?” and, “Do we want a solution that is predictable and functions reliably with SIP standard compliant equipment or is it sufficient with a best effort solution that works in certain scenarios and maybe only with a specific operator?”

Universal Plug-and-Play (UPnP) – The SIP client or Windows® is in control
Universal Plug-and-Play (UPnP) for NAT control allows Microsoft Windows® or a UPnP-capable SIP client to take control of the firewall. Both the client and firewall must support UPnP. This is a viable alternative only for those that can be sure there will never be anything malevolent on the LAN. UPnP is only supported by few firewalls and SIP clients. Due to the inherent high security risk in allowing a third party software to take control of the firewall this method is rarely used and in practice only for home users.

STUN, TURN, ICE – The SIP client is in control
These are all protocols proposed by the IETF for solving the firewall/NAT traversal issue with intelligence in the clients together with external servers. With these methods, pinholes are created in the NAT/firewall for SIP signaling and media to pass through. It is also the responsibility of the SIP client to emulate what the protocol should have looked like outside the firewall. These methods assume certain behavior from the NAT/firewall and cannot work in all scenarios. In addition, they remove control from the firewall, which must be sufficiently open to allow the users to create the necessary pinholes.

Session Border Controllers at Service Provider – The service provider is in control
Most service providers use some sort of session border controller (SBC) in their core network to perform a number of tasks related to their SIP services. One of these tasks is to make sure that the SIP services can be delivered to their customers. They may use STUN, TURN, ICE for this by acting as a server component for these protocols. However, not all clients support these protocols so the SBC may also use far-end-NAT traversal (FENT) technology for NAT traversal. The FENT function will aid remote SIP clients by transforming any SIP message by rewriting all relevant information and relay media, as well as keeping the
client on the NATed network reachable. This solution only works with firewalls that are open from the inside, and may not work with all equipment and in all call scenarios. FENT is best suited for road warriors working at a hotel or at a conference, rather than at fixed location where there are more reliable and secure solutions. FENT also removes control from the firewall, which must be sufficiently open to allow FENT from the service provider SBC to work.

SIP-capable Firewalls or enterprise SBC – The firewall administrator is in control
This is a long-term solution where the problem is solved where it occurs, at the firewall or in tandem with an existing firewall using an enterprise session border controller. When deployed at the enterprise edge the SBC offers the same security and control as it does for the service provider's core network. The enterprise SBC typically has a built-in SIP proxy and/or back-to-back user agent (B2BUA) functionality to give unparalleled flexibility in real-life enterprise deployments.

Most vendors of SBCs for service providers have products that can be deployed at the enterprise and then there are companies like Ingate Systems that have developed products for the enterprise market from the very beginning. Ingate's SIParator® is such an enterprise SBC.

For an enterprise, there are special security and functional requirements that make the SIP-capable firewall or enterprise SBC the solution of choice. First, it is the only solution that allows the firewall to maintain control of what is traversed between the LAN and the outside world. In addition, it is becoming more and more common to have a SIP server on the LAN. In fact, all SIP-based IP PBXs are SIP servers. In order for these SIP servers to communicate over IP with the outside world, the firewall simply must be SIP-enabled. As many IP-PBXs have done their own SIP-extensions outside the SIP standard it is very important that the firewall or enterprise SBC be adapted to support these extensions.

Due to the complexity of real-life installations, it is highly recommended to deploy specialized SIP-enabling devices such as SIP-proxy firewalls even in the SOHO and SMB markets, even if it might not be motivated from a security policy perspective.

2 SIP, NATs and Enterprise Firewalls
The market growth of live SIP-based person-to-person communications is expected to be the next big wave of Internet usage after email and the Web. SIP has quickly become the standard signaling protocol for these “realtime” IP communications, including VoIP. Research suggests that SIP is the VoIP protocol that has replaced H.323 and MGCP and that, for the foreseeable future, no replacement is expected (Business Communications Review, August 2005).

This is in large part due to the design of the SIP protocol. The protocol was engineered to:

- Be used for Internet-based applications
- Be compatible with all pre-existing Internet protocols
- Reduce -- even eliminate -- interoperability issues
- Offer tremendous flexibility

As a result, SIP now enjoys the greatest level of adoption by vendors. Today’s SIP implementations are both robust and feature rich.

However, SIP-based communications cannot reach LAN users behind firewalls and NATs automatically because firewalls are designed to prevent inbound unknown communications. This is the reason we place firewalls at the edge of our networks and then poll email and Web servers for downloads. NAT hides the private IP addresses on the LAN, stopping users on the LAN from being addressed from the outside. Very few, if any, communications are received directly from outside our local area networks. This provides us with comfort in knowing that only authorized users can gain access to our networks and the valuable information stored on our local servers and computers.

The Network Address Translation (NAT) that is created on the firewall or by routers is also a part of the security fabric. NATs are necessary primarily because the Internet IPv4 standard does not support enough unique IP addresses to allow all of the devices connected to the Internet to have their own identity i.e. unique IP address. With Network Address Translation, only the firewall or router is given a publicly routable IP address. Each device is then assigned a private IP address that is only known inside the firewall-protected space. While this works fine for the types of traffic that are typically supported on the LAN, it prevents inbound communications from reaching the intended recipient behind the firewall because the IP address of the client device is unknown and not routable.

Finally, most firewalls do not support the SIP protocol. Just as with all other protocol types, the firewall must recognize the format of the signaling in order to admit it to the network. Since many firewalls installed today do not support SIP, the inbound traffic will be stopped for this reason alone.

But why is this important?
There are a number of available methods for firewall traversal. Each has its own benefits; many have significant drawbacks. These drawbacks impact security. The choice of method for traversing firewalls/NATs determines the amount of control and security you maintain of your network. Is your security best left to your firewall administrator? The individual user? Your service provider? Also, does the solution need to work with all operators, or only one specifically? How SIP-compliant does it really need to be? Will SIP interoperability issues affect security? The answers can help determine which method of firewall/NAT traversal is right for your network.

The choice of traversal method also impacts how future-proofed your network will be. When the IETF first introduced the protocol they did so with a vision of SIP-based realtime communications becoming a universal protocol supported on all devices - from computers to phones - so that we will be able to reach anyone, anytime, wherever they may be located at that moment. This is the vision of true Global Connectivity over the Internet. And it’s not just about voice. When SIP is widely deployed, interaction will become more collaborative, with partners, vendors, employees and even customers using the most effective tool for every occasion, whether that be Instant Messaging (IM), presence, voice, video, application sharing, whiteboarding or file sharing.

The SIP servers providing these functions are often placed on the LAN. However, in order for them to communicate over IP with the outside world, SIP traffic must be able to traverse the firewall.

This means that the networks which support voice and video must be converged with the data networks. With the features available today with SIP, this vision is a reality. But the firewall and NAT traversal issues must be solved at the enterprise edge if the vision is to truly become a reality in the workplace.

3 Methods for Solving NAT/Firewall Traversal of SIP

Eventually, all firewalls will need to be SIP capable in order to support the wide-scale deployment of realtime communications. In the interim, several solutions have been proposed to work around the firewall/NAT traversal issues that limit SIP-based communication.

3.1 SIP-capable firewalls

This is a long-term solution where the problem is solved where it occurs, at the enterprise firewall or in tandem with the firewall using an enterprise session border controller.

![Figure 2 - SIP-capable firewalls and enterprise SBCs](image)

3.1.1 SIP ALG-based SIP-capable firewalls

The majority of all SIP-capable firewalls today use the SIP Application Level Gateway (ALG) architecture. This works for basic call scenarios but has limited functionality for real deployments of enterprise SIP-based realtime communications. The SIP ALG architecture tries to solve the firewall traversal problem of the SIP traffic by “taking care of the SIP packets on the fly,” making sure that they reach the right destination on the LAN. This architecture does not provide the enterprise with the full protection and flexible functionality of a SIP proxy-based firewall solution.

3.1.2 SIP proxy-based SIP-capable firewalls

The SIP proxy architecture is a complete solution to the firewall and NAT traversal issues presented by the enterprise firewall. A proxy is designed to briefly stop the packets so that each signaling packet can be inspected before the header information is rewritten and the packets are delivered to the appropriate endpoints. This provides the enterprise with a flexible, controlled implementation of SIP-based communications.
In addition, the SIP proxy can offer benefits\(^1\) not available with the ALG architecture:

- Far-end NAT traversal to support remote workers such as road warriors and home users
- Encrypted SIP signaling (TLS) and media (SRTP)
- Authentication
- Advanced filtering
- Advanced routing and control features
- Intelligence to enable the firewall to act as a backup for a hosted or centralized IP-PBX

To gain unparalleled flexibility, some SIP proxy solutions – such as Ingate’s – also encompass a so-called back-to-back-user-agent (B2BUA) functionality. The B2BUA allows the firewall to have two different call legs in the same session, one on each side of the firewall. This can help if the service provider on the Internet side does not support call transfers with the SIP method REFER. The firewall can then utilize “local call transfer” by just changing the call leg on the LAN side from one client to the other.

![Figure 3 - B2BUA functionality](image1)

### 3.2 Enterprise session border controllers

Many enterprise customers are reluctant to replace their existing firewalls with new SIP-capable firewalls because they have spent a great deal of effort setting up security policies. Also, they trust the equipment they have. Yet enterprises must overcome the limitations of their existing firewalls, whether they have firewalls with no SIP functionality or SIP ALG firewalls with limited SIP functionality. This need has triggered the development of a new type of product which some people call the “enterprise session border controller.” The Ingate SIParator\(^2\) is an example of such a device designed to work in networks where a corporate firewall is already in place. The SIParator can be considered a firewall just for SIP traffic which can be installed either in a standalone configuration, or as part of the DMZ of the existing firewall. Essentially the SIParator assumes control of SIP traffic without involving the existing firewall in the process.\(^2\)

### 3.3 Session border controllers at the service provider edge

Most service providers use some sort of SBC in their core network to perform a number of tasks related to their SIP services. One of these tasks is to make sure that the SIP services can be delivered to their customers.

![Figure 4 - Session Border Controller at the Service Provider](image2)

They may use STUN, TURN, ICE (described in next section) for this by acting as a server component for these protocols. However, not all clients support these protocols so the SBC may also use a far-end NAT traversal (FENT) technology for NAT traversal. The FENT function will aid remote SIP clients by transforming any SIP message by rewriting all relevant information and relay media, as well as keeping the client on the NATed network reachable. Typically, this far-end NAT traversal solution is

---

1. The suggested benefits noted here are derived from the capabilities of the Ingate Firewall\(^\text{®}\) and the Ingate SIParator\(^\text{®}\) products
2. The existing firewall can still log all of the traffic entering the network, even SIP traffic, if the SIParator is configured in the DMZ mode.
This solution only works with firewalls that are open from the inside, and may not work with all equipment and in all call scenarios. FENT is best suited for road warriors working at a hotel or at a conference, rather than at fixed location where there are more reliable and secure solutions. FENT also removes control from the firewall, which must be sufficiently open to allow FENT from the service provider SBC to work.

### 3.4 STUN, TURN and ICE

These are all methods proposed by IETF in an attempt to solve the firewall/NAT traversal issue with intelligence in the clients together with an external server.

#### 3.4.1 STUN

STUN (Simple Traversal of UDP through NATs) requires a STUN client on the phone or other end point device, which sends packets to a STUN server on the Internet. The STUN server replies with information about the IP address and ports from which the packets were received and detects the type of NAT device through which the packets were sent. The STUN client in the end point uses this information in constructing its headers so that external contacts can reach them without the need for any other device or technique.

![STUN Protocol Diagram](chart)

Figure 5 - The STUN protocol

STUN requires that the NAT device allow all traffic that is directed to a particular port, and that the traffic is forwarded to the client on the inside. This means that STUN only works with less-secure NATs, so-called “full-cone” NATs, and that the internal client will be exposed to an attack from anyone who can capture the STUN traffic. STUN may be useful for some, but is generally not considered a viable solution for enterprises. In addition, STUN cannot be used with symmetric NATs. This may be a drawback in many situations as most enterprise-class firewalls are symmetric.

#### 3.4.2 TURN

TURN (Traversal Using Relay NAT) allows an end point behind a firewall to receive SIP traffic on either TCP or UDP ports. This solves the problems of clients behind symmetric NATs which cannot rely on STUN to solve the NAT traversal issue. TURN connects clients behind a NAT to a single peer. Its purpose is to provide the same protection as that created by symmetric NATs and firewalls. The TURN server acts as a relay; any data received is forwarded. The client on the inside can then be on the receiving end, rather than the sending end, of a connection that is requested by the client on the inside.

![TURN Protocol Diagram](chart)

Figure 6 - The TURN protocol

This method is appropriate in some situations, but since it essentially allows inbound traffic through a firewall, with only the client in control, it has limited applicability for enterprise environments. It also scales poorly since the media must go through the TURN server.
3.4.3 ICE
ICE (Interactive Connectivity Establishment) uses STUN, TURN and other methods to solve the NAT traversal issue. ICE allows endpoints to discover other peers and then establish a connection. ICE essentially incorporates all of the methods proposed for NAT traversal of SIP that do not rely on the firewall or NAT device. ICE is a complex solution to the problem of NAT traversal, but since it encompasses multiple solutions it is regarded as one that will always enable the connection, regardless of the number of NATs involved. However, ICE still relies on client-server based approaches, and removes control from the enterprise. Due to its complexity there is very limited client support for ICE today.

STUN, TURN and ICE are methods that assume certain behavior from the NAT/firewall and do not work in all scenarios. The control is removed from the firewall which has to be sufficiently opened to allow users to create the pinholes needed to let the communication through.

3.5 Universal plug-and-play (UPnP)
Universal Plug-and-Play (UPnP) for NAT control allows Microsoft Windows® or a UPnP-capable SIP client to take control of the firewall. Both the client and firewall must support UPnP. This is a viable alternative only for those that can be sure there will never be anything malevolent on the LAN. UPnP is only supported by few firewalls and SIP clients. Due to the inherent high security risk in allowing a third party software to take control of the firewall this method is rarely used and in practice only for home users.

![Figure 7 - Universal Plug and Play (UPnP)](image-url)

Service Provider

Site A  Site B

SW  SW
4 SIP Proxy-Based Firewalls and Enterprise SBCs: Security Advantages of the SIP Proxy

4.1 Controlling media

SIP proxy technology is an excellent way to add a level of control to the flow of SIP media. This control offers tremendous advantages with regard to security.

The main purpose of SIP is to set up a media session between clients. Media is handled by other protocols (often RTP). For media to traverse the enterprise edge, the SIP proxy must dynamically open the media ports for media to flow during the duration of the call. As soon as the call is completed the media ports are closed. This behavior is much more secure than solutions with non-SIP-aware firewalls/border elements where a media port range constantly needs to be open. In general the SIP proxy approach is more secure than the IETF specified STUN/TURN/ICE methods, which requires that ports are left open from the inside of the firewall to allow media port negotiation to succeed.

In addition to the dynamic opening and closing of media ports, the edge device should only accept incoming media from the endpoint that receives media from the edge device. This protects against hackers trying to inject media from other endpoints or devices.

To protect media from being overheard by unauthorized persons, media encryption comes into play. The industry seems to have chosen SRTP using sdescriptions for key exchange as the de facto standard for media encryption. Using SRTP to encrypt media traversing the Internet effectively stops eavesdropping. The integrity of the call is much stronger than ever possible on PSTN.

4.2 SIP signaling

Firewalls with a SIP server and full SIP proxy play a critical role in maintaining enterprise security, and securing VoIP. They can rewrite SIP signaling and process in a very flexible way, ensuring correct routing and interoperability with other systems built to RFC 3261 and related standards.

One important part of the SIP proxy is the SIP parser. The SIP parser verifies that the SIP message is valid and that it may be forwarded to the local LAN. Malformed SIP messages are discarded. The SIP parser must be robust enough to withstand any types of malformed SIP messages without crashing. Also, to mitigate DoS attacks, the parser should be able to process a very large number of packets.

The SIP proxy should include support for the optional loop detection mechanism defined in the SIP specification. This mechanism discerns whether a SIP message is looping (sending the SIP message to itself) and, if so, aborts this behavior. This detection mechanism also protects against DoS attacks where a SIP message is constructed to create loops and thus keep the SIP proxy too busy to engage in useful processing.

In order to protect resources, e.g. a PSTN gateway, authentication of SIP users should be supported. The standard means of authentication of SIP users is via the Digest protocol. SIP users’ credentials should be stored in a centralized database e.g. on a RADIUS server. This is more secure and likely easier to maintain.

SIP signaling consists of messages in ASCII text (plain text), and are therefore easy to read and manipulate. It is strongly recommended to encrypt and authenticate SIP signaling. This is normally achieved by supporting TLS or MTLS. MTLS is the most secure method as both server and client mutually authenticate each other using CA-signed certificates or certificate chains.

In order to provide greater and more flexible protection mechanisms, filters are useful features. A typical filter would include the following:

• SIP methods can be allowed or prohibited per network.
• Authentication can be enabled or disabled per network and SIP method.
• SIP messages can be filtered on content type.
• Incoming callers can be restricted to a white list; this list can be individually enabled/disabled per user.
• A filter based on from/to header may be used to allow or disallow processing.
5 Which NAT/Firewall traversal solution is right for you?

The choice of method for traversing firewalls/NATs is, to a large extent, dependent on the answer to the questions: “Who should be in control of your security infrastructure: the firewall administrator, the user or a service provider?” and, “Do we want a solution that is predictable and functions reliably with SIP standard compliant equipment or is it sufficient with a best effort solution that works in certain scenarios and maybe only with a specific operator?”

Figure 8 - Positioning of NAT traversal solutions

The choice of NAT/firewall traversal solution must be selected with two very important things in mind:

1) The level of security policy and control you need
2) The level of flexibility in terms of configuration you need

The level of security policy and control you need

If you run a business and want to maintain the control of your own security infrastructure - with a high security policy e.g all ports in your firewall closed from the inside and deep packet inspection of the SIP traffic -- then there is really only one choice: a SIP proxy-based SIP-capable firewall or a SIP proxy-based enterprise session border controller.

If you have no or a very low security policy and do not mind that a service provider asks you to open up certain ports in your firewall, then you may consider the far-end NAT traversal solution offered by the SBC in the service provider’s core network. With this solution control of your security infrastructure is given to the service provider.

If you have no or a very low security policy and have a simple NAT box or a residential firewall with all ports open from the inside, then the standard STUN, TURN, ICE solutions may be your choice. You must however remember that these protocols were built by the IETF with a vision that all SIP traffic should be point-to-point over the Internet, with very little consideration of the security policies that need to be maintained in the enterprise. With these solutions you will put the control over to the clients and the firewall will have no control over what is admitted to the network.

With the proliferation of viruses and spyware, we do not believe that anyone including the home user should use the Universal Plug and Play method (UPnP) for admitting SIP traffic through the firewall. It is highly dangerous to let external software take control over the firewall.

The level of flexibility in terms of configuration you need

Even if you are not too concerned about security you want to make sure that the solution works without any issues. In our experience real life deployments of, for example, an IP-PBX on the LAN is so complex so that the basic NAT traversal offered by some solutions just is not good enough. This is why there is a good market for low-cost SIP proxy-based firewalls for the SOHO and SMB markets (such as the one from Ingate’s sister company, Intertex Data).
So, it really comes down to a two dimensional choice with security and flexibility as the core parameters.

![Figure 9 - Security and Flexibility](image)

### 6 Security considerations for SIP deployment in the enterprise

#### 6.1 Threats

Connecting a device to the Internet exposes the entire network to many types of threats. One example is a brute force attack where the intruder tries to log into a service using a user/password database trying a huge number of username and password combinations until the intruder finally succeeds in finding the right one. Once access has been granted the intruder may be able to launch other types of attacks based on known vulnerabilities to the service in question and in this way get access to other services or data.

Another example of a threat would be Denial of Service (DoS) attack where the attacker uses many different hosts or “zombies” to send a large number of packets to make the host drown or crash due to the vast amount of traffic.

The above are two examples of traditional data communication attacks. These and many others can easily be transformed into attacks on VoIP equipment. The VoIP Security Alliance or VOIP-SA has categorized possible attacks and threats on a VoIP system and made this information publicly available. This information is a resource for understanding what threats needs to be taken into account when it comes to securing VoIP.

#### 6.2 Importance of a stable platform

Firewall vendors have developed significant expertise in securing data communication. They know how to design stable systems that are locked down to only admit services that have been configured to pass. Firewalls inspect and log traffic and, if intelligent enough, they can even block suspected attacks including traffic from known malevolents.

Firewalls alone cannot prevent DoS attacks, but they can be built to withstand attacks, making them harder to occur. Firewalls can also lay the foundation for a swift recovery. More importantly, they can be built to protect the enterprise LAN from being reached by the DoS attack. A good enterprise SBC should have the same stable platform so it could be considered a “firewall specialized in VoIP.”

### 7 Conclusion

The vision of global connectivity is now upon us, now that the SIP protocol has gained acceptance by vendors and service providers. However, the issue of NAT traversal is still an impediment to widespread adoption of SIP and the reality of converged communications. Many companies today are looking for the right solution to bring the benefits of such collaboration into their network – and to do so while maintaining security and control.
Several solutions exist today, all of which have a place in the convergence toolbox. Some of the options can be characterized as being useful in situations where security demands are light and where there are no SIP servers on the LAN.

In those cases where security concerns are greater, and in situations where tight corporate firewall environments prevent the use of some of these simpler solutions, a more robust technique should be employed. SIP-capable firewalls and SIP-enabling edge devices provide the means to solve the traversal issue even in scenarios where there are tight security policies and SIP servers on the LAN that should be used for communication with the outside world.

The choice really comes down to the level of control that an organization is willing to cede to third parties or to the clients and users themselves. SIP-capable firewalls and enterprise SBCs offer the greatest amount of control with flexibility for the company to install the collaboration tools that it needs without sacrificing the security and integrity of the network.

Whichever solution is chosen, the resulting advantages that come from adopting SIP-based real-time, converged global communications offer productivity benefits to every company.

**About Ingate® Systems**

Ingate® Systems develops firewall technology and products that enable SIP-based live communication for the enterprise while maintaining control and security at the network edge. Ingate has a long history of developing next-generation firewall technology that solves the NAT/firewall traversal issue with SIP communications. In addition to an extensive line of Ingate Firewalls®, the company also produces the award-winning Ingate SIParator®, a device that connects to an existing network firewall to seamlessly enable SIP communications. Ingate products currently protect the networks of retail companies, financial institutions, industrial firms, government agencies and small-to-large enterprises throughout Europe, Asia and North America. Ingate Systems AB is headquartered in Sweden with offices in Stockholm and Linköping. Its wholly-owned subsidiary, Ingate Systems Inc., is located in Hollis, New Hampshire, with a U.S. technology center in Frisco, Texas. For more information on Ingate Systems, visit [www.ingate.com](http://www.ingate.com).