ALG for VoIP Overview

In general an ALG may offer the following functions:

- Allows endpoint applications to use known and ephemeral TCP/ UDP ports to communicate with the known ports used by the server applications and dynamic ports used by media/audio streams though firewalls and/or router NAT configuration. Without an ALG, ports would be blocked and/or packets dropped unless an explicitly defined large number of ports in the firewall {or static mapping within routers} were configured — rendering the network vulnerable to attacks on those ports and failing FIPS/PCI compliance etc.
- Enables conversion of the network layer address information found inside an application payload between the addresses acceptable on either side of the firewall/NAT. Often referred to as the 'gateway' function for an ALG
- Allows application-specific commands to be identified and offering granular security controls {deep packet inspection}
- Enables synchronizing between multiple streams/sessions of data between two hosts exchanging data for large file transfer etc. The main file transfer control connection will often remain idle during a transfer - ALG can prevent the control connection getting timed out by network devices while waiting during lengthy file transfer

ALG uses deep packet inspection of all the packets over a given network to make this functionality possible. Often used within VoIP enterprise to control address translation through NAT and firewalls for H.323 and SIP.

For instance, for Session Initiation Protocol (SIP) Back-to-Back User agent (B2BUA), an ALG can allow firewall traversal with SIP. If the firewall has its SIP traffic terminated on an ALG then the responsibility for permitting SIP sessions passes to the ALG instead of the firewall. An ALG can solve another major SIP headache: NAT traversal. Basically a NAT with a built-in ALG can rewrite information within the SIP messages and can hold address bindings until the session terminates. A SIP ALG will also handle SDP in the body of SIP messages (which is used ubiquitously in VoIP to set up media endpoints), since SDP also contains literal IP addresses and ports that must be translated.

An ALG is very similar to a proxy server, as it sits between the client and real server, facilitating the exchange. There seems to be an industry convention that an ALG does its job without the application being configured to use it, by intercepting the messages. A proxy, on the other hand, usually needs to be configured in the client application. The client is then explicitly aware of the proxy and connects to it, rather than the real server.

ALG is often wrongly reported as the cause of issues with VoIP implementation and blamed for one-way-audio etc. when it is really just a lack of knowledge by the design and implementation teams of SIP SDP and NAT. This document attempts to explain the functionality of ALG for VoIP.

This document assumes knowledge of VoIP protocols and uses SIP (Session Initiation Protocol) as the example to describe how ALG works.

IF YOU REQUIRE A QUICK EXPLANATION OF SIP PROTOCOL AND HOW IT WORKS THROUGH NAT AND FIREWALLS, SEE {LINK} - any questions → training@kccvoip.com

ALG for VoIP Overview

It is true that some home routers and SOHO routers/firewalls have not implemented ALG correctly and have caused issues. For very small installations where there are only a few telephones and less than two voice servers, it is perfectly OK to disable ALG for SIP and either use dedicated port-mapping and/or symmetrical RTP to design and configure a working solution along with STUN/TURN etc. However, if the enterprise calls for multiple voice servers and/or the ability to transfer calls (SIP re-invite), use direct media, asymmetrical RTP/SRTP, CODEC translation, multiple providers, distributed DSP farms etc…. AND have this all behind NAT, then ALG {such as Cisco ALG or CUBE, Juniper ALG etc.} is absolutely ESENTIAL. Alternatively, larger implementations tend to use public addressed SBC {session border controllers} to remove most NAT and security issues {Acme Packet, Sangoma, Kamailio etc.}

Using Session Initiation Protocol (SIP) as our example VoIP protocol for this document - SIP is an Internet Engineering Task Force (IETF)-standard protocol for initiating, modifying, and terminating multimedia sessions over the Internet. Such sessions might include audio telephony, multimedia, conferencing, instant messaging and application-level mobility in network environments.

SIP control packets provide the function to distribute session-description information (SDP = Session Description Protocol), and during any session setup and/or session transfer/change, to negotiate and modify the parameters of the session. SIP is also used to terminate the session, signal a call establishment, provide failure indication, status and methods for endpoint registration.

SDP information is included in SIP INVITE, Re-INVITE, 200-OK and ACK messages etc. and indicates the required multimedia parameters of the session;

Although SIP can use different description protocols to request the session, most SIP Application Layer Gateway (ALG) implementations support only the Session Description Protocol (SDP).

SDP provides essential information that a system can use to join a multimedia session or setup a telephone call etc. - each endpoint will send SDP content to show which IP address, port number, CODEC, multimedia types that endpoint wants to receive. These SDP packets are formed by the endpoint software and use the local addressing and port information etc. So it is not unusual to have a telephone sending SIP invite packets containing SDP with private IP addressing. IF YOU DO NOT KNOW HOW SIP/SDP PACKETS ARE FORMED USING THE LOCAL IP ADDRESSING, PORT AND CODEC INFORMATION - For a basic SIP/NAT explanation see LINK

ALG for VoIP Overview

SIP messages consist of requests from client to server and responses to the requests from server to client while establishing or modifying a session. A SIP user agent (UA) is an application that runs at the endpoints and consists of two parts:

- user agent client (UAC) - sends SIP requests on behalf of the user
- user agent server (UAS) - listens to responses and notifies the user

UA are SIP proxy servers and phones etc. UAS could be voice servers, PBX, providers etc.

ALG SIP Operation Example

SIP traffic consists of two main traffic types, the signalling and the media/audio stream.

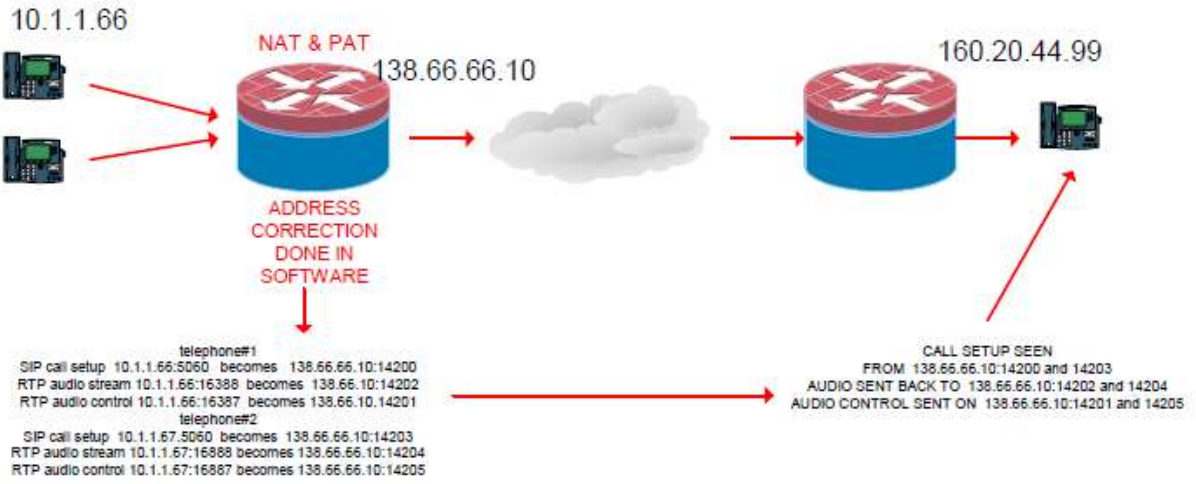
SIP signalling request and response messages between client and server use transport protocols such as UDP or TCP. The media/audio stream carries the data using transport protocols usually UDP/RTP, UDP/SRTP or TCP/SRTP.

The ALG configuration is often referred to as packet inspection or ALG service policy. ALG default ports for SIP signalling will usually be UDP 5060 and TCP 5061, however these can be changed to match the requirements. The media/audio stream uses ephemeral port numbers that can change several times during the course of a call. These dynamic ports for the media/audio stream are usually the cause of issues due to lack of knowledge on how they are selected and how they are mapped through any firewalls and NAT.

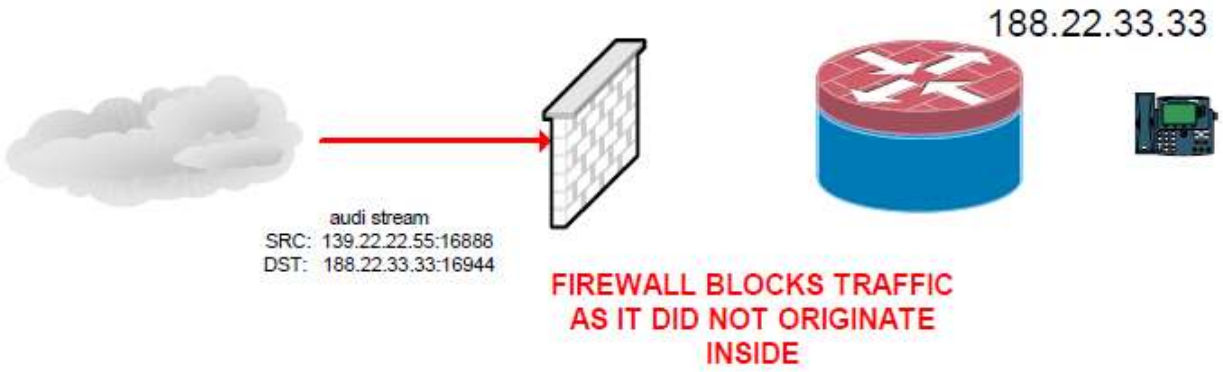
Often the range of ports used for multimedia/audio have a vast range such as 32182 – 64000 with different ISP providers using different but overlapping port ranges trying to connect to multiple servers within the enterprise - making it almost impossible to implement multiple ISP trunks to multiple providers using static mapping even with symmetrical RTP + it is also considered very insecure to allow vast port ranges through the enterprise firewalls and impossible to maintain FIPS or PCI compliance.

The ports used for the media/audio sessions are not known until call setup as they are dynamically allocated at each end of the call path, but the ports used for the SIP signalling and negotiation are well-known (or predefined). The ALG therefore takes interest in the signalling packets from the control session and inspects the negotiation for the transport information used for the media/audio session to extract the information for SDP.

ALG for VoIP Overview



ALG corrects the SIP/SDP addressing and creates pinholes using the values found in the SDP packets. The pinholes allow the traffic through the firewall and creates the necessary NAT mapping to ensure the media/audio streams can be routed in BOTH directions. NOTE – when SDP is encrypted, additional configuration will be required as the ALG will not find the IP address and port information within SDP.



Without ALG, one-way media/audio will often be experienced due to one of the RTP streams blocked by firewall and/or lack of NAT entry to allow the traffic through to the endpoint. In this example the audio stream from 139.22.22.55 is blocked by the firewall because there are no pinholes or static mapping to allow the routing.

ALG for VoIP Overview

SDP Session Descriptions

A session is described by a series of attribute/value pairs, one per line. The attribute names are single characters, followed by =, and a value. Optional values are specified with =*. Values are either an ASCII string, or a sequence of specific types separated by spaces. Attribute names are only unique within the associated syntactic construct, such as within the session, time, or media only.

The main two fields used by ALG contain Transport Layer information.

- **c= for connection information**
  This field can appear at the session or media level. It appears in this format:
  c=<network-type><address-type><connection-address>
  IPv4 and IPv6 address types are usually supported.
  If the destination IP address is a unicast IP address, the SIP ALG creates pinholes using the IP address and port numbers specified in the media description field m=
- **m= for media announcement**
  This field appears at the media level and contains the description of the media. It appears in this format: **m=<media><port><transport><fmt list>**
  RTP is often the required Application Layer transport protocol. The port number indicates the destination port of the media stream (each UA selects the port number it will listen for the media/audio streams). The format list (fmt list) provides information on the Application Layer protocol that the media/audio uses (each UA will be configured to support a list of CODECs).
  ALG opens ports for RTP and Real-Time Control Protocol (RTCP). Every RTP session has a corresponding RTCP session - the ALG will create pinholes for both RTP and RTCP traffic. By default, the port number for RTCP is RTP port number + 1.

Pinhole Creation

Two pinholes are created (one for RTP and for RTCP) with the same destination IP address. The IP address is derived from the c= field in the SDP session description either from the session-level or the media-level portion of the SDP based on the following SDP convention:

- First - 'c=' field containing an IP address in the media level. If there is such a field, the ALG parser extracts that IP address, and the SIP ALG uses that address to create a pinhole for the media/audio RTP/RTCP.
- If there is no 'c=' field in media level, ALG parser extracts the IP address from the c= field in the session level, and uses that IP address to create a pinhole for the media/audio RTP/RTCP. If the session description does not contain a 'c=' field in either level, this indicates an error and usually causes the packet to be dropped and logged.

The SIP ALG also opens pinholes for signalling. Signal pinholes are used after session timeout, and are also used for the signalling to be sent to a third-party addresses etc. The SIP ALG signal pinholes never age out, unlike RTP or RTCP pinholes, where only the destination IP and destination port are specified.

The SIP ALG opens signal pinholes as needed for following headers:

- CONTACT
- ROUTE
- RECORD-ROUTE
- VIA

ALG needs the following information to create the pinholes. This information either comes from the SDP session description or from the SIP headers (as listed above).

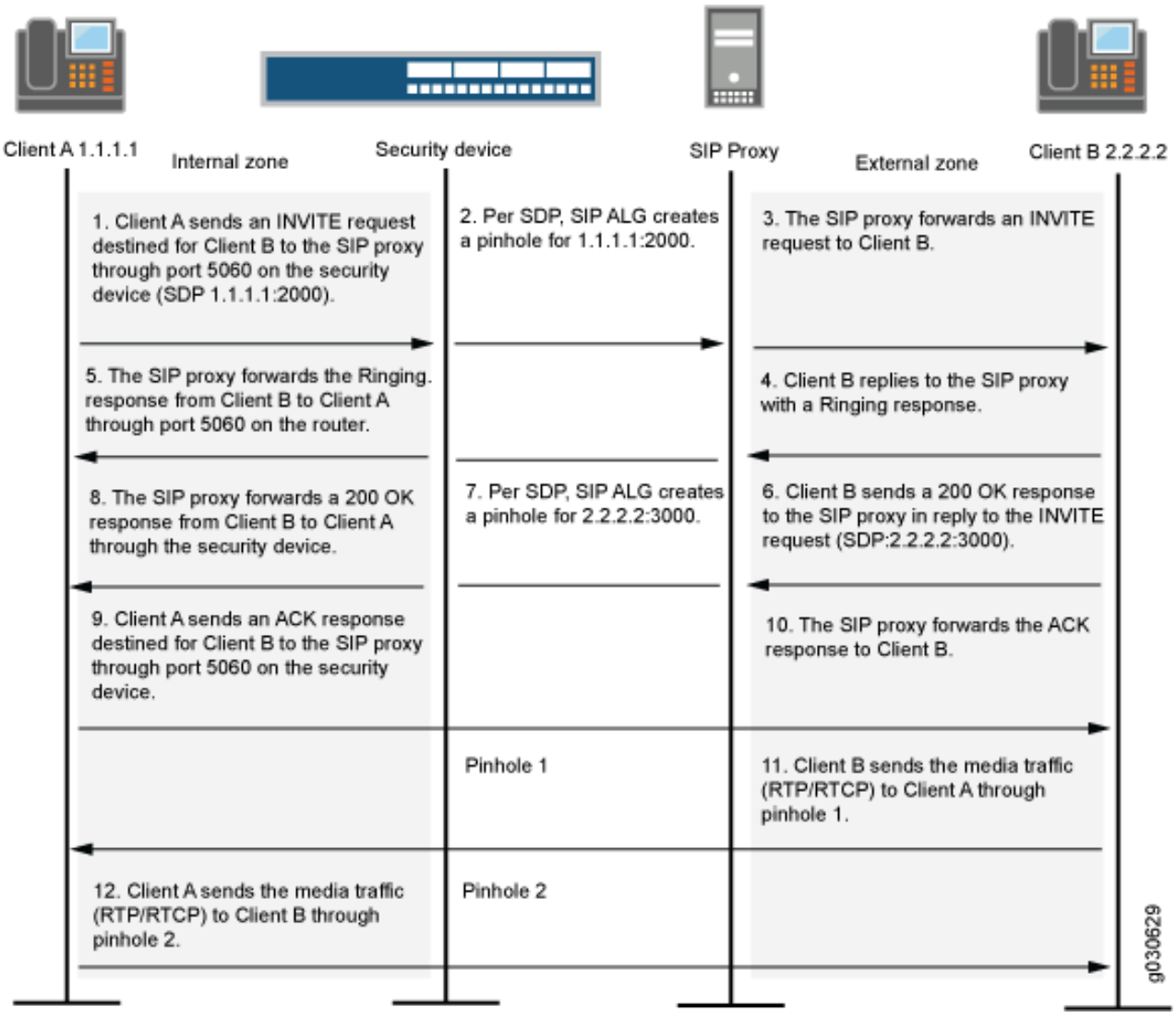
- Protocol—UDP or TCP
- Source IP
- Source port
- Destination IP—destination IP address from the c= field at the media or session level
- Destination port—destination port number for RTP from the m= field in the media level and the destination port number for RTCP {RTP port +1}
- Lifetime—This value indicates the length of time (in seconds) during which a pinhole is open to allow a packet through.

{discuss examples}

ALG for VoIP Overview

Example call setup between two SIP clients and how the SIP ALG creates pinholes to allow RTP and RTCP traffic.

SIP ALG Call Setup Example



{DISCUSS PORTS AND PROTOCOLS SEEN ABOVE}

ALG for VoIP Overview

Understanding IPv6 Support for SIP ALG

IPv6 is supported within most enterprise ALG along with NAT-PT mode and NAT64 address translation.

The SIP ALG processes the IPv6 address in the same way it processes the IPv4 address for updating the payload if NAT is configured and opening pinholes for future traffic.

Special processing occurs for the following formats:

- IPv6 in SIP URI = much the same as a URI with IPv4 addresses, an IPv6 address is enclosed in square brackets with address blocks separated by colons.
- IPv6 in SDP = use of an IP6 marker
- The SIP ALG with IPv6 limitation:
    - If NAT64 with persistent NAT is implemented, ALG adds the NAT translation to the persistent NAT binding table if NAT is configured on the Address of Record (AOR). NAT cannot duplicate the address configured, coexistence of NAT66 and NAT64 configured on the same address is not supported.
      Hence, only one binding is created for the same source IP address.

Understanding SIP ALG Request Methods

The SIP request and response messages contain *method* fields that denotes the purpose of the message - these method types and response are usually supported:

- REGISTER—A user sends a REGISTER request to a SIP registrar server to inform it of the current location of the user. A SIP registrar server records all the information it receives in REGISTER requests and makes this information available to any SIP server attempting to locate a user.
- INVITE—A user sends an INVITE request to invite another user to participate in a session. The body of an INVITE request can contain the description of the session.
- ACK—an ACK request to confirm reception of the final response to the INVITE request. If the original INVITE request did not contain the session description, the ACK request must include it.
- 1*xx*, 202, 2*xx*, 3*xx*, 4*xx*, 5*xx*, 6*xx* Response Codes—Used to indicate the status of a transaction. Header fields are modified by endpoints and servers.
- OPTIONS—The User Agent (UA) obtains information about the capabilities of the SIP proxy. A server responds with information about what methods, session description protocols, and message encoding it supports. Often used as a SIP keepalive for trunks.
- BYE—Any user will send a BYE request to terminate a session.
- CANCEL—A user sends a CANCEL request to cancel a pending INVITE request. A CANCEL request has no effect if the SIP server processing the INVITE had sent a final response for the INVITE before it received the CANCEL.
- INFO—Used to communicate mid-session signalling information.
- SUBSCRIBE—Used to request current state and state updates from a remote node or server.
- NOTIFY—Sent to inform subscribers of changes in state to which the subscriber has a subscription such as presence, voicemail, BLF etc. .

- REFER—Used to refer the recipient (identified by the Request-URI) to a third party by the contact information provided in the request.
  For example, if user A in a private network refers user B, in a public network, to user C, who is also in the private network, the SIP Application Layer Gateway (ALG) allocates a new IP address and port number for user C so that user C can be contacted by user B. If user C is registered with a registrar, however, its port mapping is stored in the ALG Network Address Translation (NAT) table and is reused to perform the translation.
- UPDATE—Used to open pinhole for new or updated SDP information. The Via:, From:, To:, Call-ID:, Contact:, Route:, and Record-Route: header fields are modified.

ALG DoS Attack Protection

Denial-of-service (DoS) protection features enables SBC, routing and firewall devices to monitor INVITE requests and generate proxy server replies to them. If a reply contains a 3*xx,* 4*xx*, or 5*xx* response code other than 401, 407, 487, and 488 that are not real failure responses, then the request should normally be blocked.

{discuss}

SIP ALG Unknown Message Types

The ALG allows configuration which enables you to specify how unidentified Session Initiation Protocol (SIP) messages are handled. The default is often to drop unknown messages.

Permitting unknown messages can compromise security. However, in a lab, secure test or controlled production environment, this configuration can be useful for resolving interoperability issues with various vendors equipment. Permitting unknown SIP messages can help you get your network operational so you can later analyse the VoIP traffic.

SIP ALG Call Duration and Timeouts

The call duration and timeout features give you control over Session Initiation Protocol (SIP) call activity and help you to manage network resources.

Typically a call ends when one of the clients sends a BYE or CANCEL request. The SIP ALG intercepts the BYE or CANCEL then requests and removes all media sessions for that call. If there are problems preventing clients in a call from sending BYE or CANCEL requests, for example, a power or circuit failure, the call might go on indefinitely and consume resources on the networks/servers/trunks and costs through providers etc.

A call may have one or more media/audio channels active. Each active channel has two sessions, one for Real-Time Transport Protocol (RTP) traffic and one for Real-Time Control Protocol (RTCP) signalling. Timeouts can be sued to ensure calls do not go on indefinitely.

The following parameters usually control SIP call activity:

- `RTP-timeout or inactive-media-timeout` = maximum length of time a call can remain active without any media/audio (RTP or RTCP) traffic. Each time an RTP or RTCP packet occurs within a call, this timeout resets. When the period of inactivity exceeds this setting, the temporary openings (pinholes) in the firewall the SIP ALG opened for media are closed
- `Media/audio session timeout or maximum-call-duration` = absolute maximum length of a call. When a call exceeds this parameter setting, the SIP ALG tears down the call and releases the media sessions.
- RFC 3261 SIP-T1 or `t1-interval` = specifies the roundtrip time estimate, in seconds, of a transaction between endpoints. SIP timers often scale with the t1-interval (as described in RFC 3261), when value of the t1-interval timer is changed, other SIP timers also are adjusted
- `RFC 3261 SIP-T4 or t4-interval` = specifies the maximum time a message remains in the network. The default is 5 seconds and the range is 5 through 10 seconds. SIP timers scale with the t4-interval (as described in RFC 3261), when you change the value of the t4-interval timer, other SIP timers also are adjusted
- `c-timeout` = specifies the INVITE transaction timeout at the proxy, in minutes; the default is 3. As the SIP ALG is in the middle, instead of using the INVITE transaction timer value B (default is (64 * T1) = 32 seconds), the SIP ALG gets its timer value from the proxy

SIP ALG Hold Resources

There are two main methods of handling SIP call holding - When a user puts a call on hold, the SIP ALG releases SDP media/audio resources, such as pinholes and translation contexts. When the user resumes the call, an INVITE request message negotiates a new SDP offer and answer and the SIP ALG reallocates resources for the media/audio stream. This can result in new translated IP address and port numbers for the media description even when the media description is the same as the previous description. This is compliant with *RFC 3264 An Offer/Answer Model with the Session Description Protocol (SDP)*.

Alternatively, some SIP implementations have designed call flows so that the User Agent (UA) module ignores the new SDP INVITE offer and continues to use the SDP offer of the previous negotiation. To accommodate this method, configure the ALG to retain SDP media resources when a call is put on hold for reuse when the call is resumed.

NOTE – BLF Scaling - UDP-Based SIP ALG

FYI - If/When busy lamp field (BLF) is configured, the phone subscribes to a resource list available on the voice server to be notified of status information for other extensions. BLF works through the Session Initiation Protocol (SIP) and uses the SUBSCRIBE and NOTIFY messages. Usually, the phone is the subscriber and the IP PBX/voice server is the notifier.

Some implementations of SIP ALG only support 3000-byte SIP messages – therefore, if there are too many instances of BLF in the message body, the payload will not be changed and the translation for the BLF packets will fail through ALG.

More up-to-date routers (Junos OS Releases 12.3X48-D15 and above, Junos OS Release 17.3R1 and above, Cisco IOS XE >16 etc.) support 65,000-byte SIP messages on the UDP protocol. In the scaling BLF application, if every instance is around 500 bytes, the SIP ALG supports 100 instances in one SIP UDP message = jumbo SIP through NAT.

SIP ALG and NAT

The Network Address Translation (NAT) protocol enables multiple hosts in a private subnet to share a single or pool of public IP addresses to access the Internet. For outgoing traffic, NAT replaces the private IP address and port of the host in the private subnet with the public IP address and port from PAT allocation. For incoming traffic, the public IP address is converted back into the private address and port, and the message is routed to the appropriate host and port in the private subnet.

Using NAT with the Session Initiation Protocol (SIP) service is more complicated because SIP messages contain IP addresses within the SIP headers as well as within the SIP body - these addresses and porta are different to those on the IP packet headers. When using NAT with the SIP service, the SIP headers contain information about the caller and the receiver {locally significant addresses and ports}. The SIP body contains the Session Description Protocol (SDP) information, which includes IP addresses and port numbers for transmission of the media/audio.

How IP addresses and port numbers in SIP messages are replaced depends on the direction of the message. For an outgoing message, the private IP address and port number of the client are replaced with the public IP address and port number {protocol port defaults or as configured} of the firewall/router. For an incoming message, the public address of the firewall/router is replaced with the private address of the client and the public port address is replaced from the port address translation table/database or static configuration.

When an INVITE message is sent out across the firewall/router, the SIP ALG collects information from the message header into a call table, which it uses to forward subsequent messages to the correct endpoint. When a new message arrives, for example an ACK or 200 OK, the ALG compares the “From:, To:, and Call-ID:” fields against the call table to identify the call context of the message. If a new INVITE message arrives that matches the existing call, the ALG processes it as a SIP REINVITE.

When a message containing SDP information arrives, the ALG allocates ports and creates a NAT mapping between them and the ports in the SDP {ALG provides consecutive even-odd ports for RTP/RTCP}. If it is unable to find a pair of ports, it discards the SIP message.

IPv6 is usually supported on the SIP ALG along with NAT-PT mode and NAT64 address translation.

ALG for VoIP Overview

Now in more detail we will discuss the following;

- Outbound Calls
- Inbound Calls
- Forwarded Calls
- Call Termination
- Call Re-INVITE Messages
- Call Session Timers
- Call Cancellation
- Forking
- SIP Messages
- SIP Headers
- SIP Body
- SIP NAT Scenario
- Classes of SIP Responses
- NAT Mode in Pure IPv6 Mode (NAT66) for SIP IPv6 ALG
- NAT-PT
- NAT64
- STUN and SIP ALG

Outgoing Calls to the public network

When a SIP call is initiated with a SIP request message **from the internal to the external** network, NAT replaces the IP addresses and port numbers in the SDP and binds the IP addresses and port numbers to the firewall/router. Via, Contact, Route, and Record-Route SIP header fields, if present, are also bound to the firewall/router IP address. The ALG stores these mappings for use in retransmissions and for SIP response messages.

The SIP ALG then opens pinholes in the firewall/router to allow media/audio through on the dynamically assigned ports negotiated based on information in the SDP and the Via, Contact, and Record-Route header fields. The pinholes also allow incoming packets to reach the Contact, Via, and Record-Route IP addresses and ports. For the return traffic, the ALG inserts the original Contact, Via, Route, and Record-Route SIP fields back into packets.

Incoming Calls from the public network

Incoming calls either via a static public addressed interface, VRRP, GLB etc. from the public network can be statically configured IP addresses that point to internal hosts. More often the configuration will have voice servers behind NAT which originate the sessions to external public addressed services and trunks. All SIP IP addresses are dynamically recorded by the ALG as it monitors REGISTER messages sent by internal hosts and servers to the SIP registrar at the provider or remote organisation.

The ALG examines the SIP request message (initially an INVITE or REGISTER) and, based on information in the SDP, opens the pinholes for outgoing media/audio. When a 200 OK response message arrives back from the remote/provider, the SIP ALG performs NAT on the IP addresses/ports and opens pinholes in the outbound direction. If not using statically configured

signalling to the remote/provider, the ALG opens gates to allow the outbound traffic with a short time-to-live, and they time out if a 200 OK response message is not received quickly.

When a 200 OK response arrives, the SIP proxy examines the SDP information and reads the IP addresses and port numbers for each media/audio session. The SIP ALG performs NAT on the addresses and port numbers, opens pinholes for outbound traffic, and refreshes the timeout for gates in the inbound direction.

{DISCUSS}

When the ACK arrives for the 200 OK, it also passes through the SIP ALG. If the message contains SDP information, the SIP ALG ensures that the IP addresses and port numbers are not changed from the previous INVITE {if they are changed, the ALG deletes old pinholes and creates new pinholes to allow media/audio to pass through}. The ALG also monitors the Via, Contact, and Record-Route SIP fields and opens new pinholes if it determines that these fields have changed.

Forwarded Calls

If, for example when configured for 'direct-media', user {X} outside the enterprise network calls a user {Y} inside the enterprise network - user forwards the call to user {Z} outside the network. The SIP ALG processes the INVITE from user {X} as a normal incoming call, then when the ALG examines the forwarded call from {Y} to {Z} outside the network and notices that {Y} and {Z} are outside networks, it does not open pinholes in the firewall, because media/audio will flow directly between user {Y} and user {Z}.

If, for example, user {A} outside the enterprise network calls a user {B} inside the enterprise network - user forwards the call to user {C} also inside the enterprise network. The SIP ALG processes the INVITE from user {A} as a normal incoming call, then when the ALG examines the forwarded call from {B} to {C} inside the enterprise network, it opens pinholes in the firewall to allow IP addressing and ports for user {C} and media/audio will flow between user {B} and user {C} through the ALG configured NAT.

If 'no direct-media' is configured or there is a need for voice recording, calls are configured to flow through a voice server. The voice server provides voice recording and CODEC translation between two or more users.

Call Re-INVITE Messages - Call Transfer

If configured for 'direct-media', Re-INVITE messages add new media sessions to a call and remove existing media sessions = new pinholes are opened in the firewall and new address bindings are created through NAT. The process is identical to the original call setup. When all the media sessions or media pinholes are removed from a call, the call is removed when a BYE message is received.

If configured for 'no-direct-media' within the enterprise, Re-INVITE messages add new media sessions to a call and remove existing media sessions = pinholes remain open in the firewall and address bindings remain mapped through NAT.

{DISCUSS different topologies and location of voice server in relation to direct-media}

Public addressed SBC calls internal phone – transfers to internal phone

Public addressed SBC calls internal phone – transfers to external SIP trunk

Internal phone calls external SIP trunk - transfers to different SIP trunk

Internal phone calls external SIP trunk - transfers to ISP provider cloud

Call Termination

The BYE message terminates a call and ALG translates the header fields much the same as it does for the other messages, except the ALG then delays call teardown for 5 seconds to allow time for transmission of the 200 OK to acknowledge the BYE message.

Call Session Timers

Watch out for call timers within ALG, voice servers and NAT tables. If keepalives are not used and the endpoints are configured to supress silence packets, the session can timeout during long silence periods (such as during conference calls when on mute etc.).

Most SIP ALG implementations use configured timeout values to set the maximum amount of time a call can exist. Voice servers will also have configuration for maximum silence or RTP inactivity before assuming the call has terminated. Be aware should one of these events occur:

- Network, Trunk or circuit failures - no BYE message
- End systems crash during a call – no BYE message
- Malicious send of a BYE in an attempt to attack a SIP ALG
- Poor implementations of SIP proxy fail to process Record-Route so no BYE message

Call Cancellation

Either party can cancel a call by sending a SIP CANCEL message - the SIP ALG then closes pinholes through the firewall and releases address NAT bindings. Before releasing the resources, the ALG delays the control channel age-out for approximately 5 seconds to allow time for the final 200 OK to pass through. The call is usually terminated when the 5-second timeout expires, regardless of whether a 487 or non-200 response arrives.

Forking

Not supported in all environments, forking enables a SIP proxy or voice server to send a single INVITE message to multiple destinations simultaneously. When the multiple 200 OK response messages arrive for the single call, the SIP ALG will update call information with the first 200 OK messages it receives.

SIP Messages in Detail

The SIP message format consists of a SIP header section and the SIP body. In the request messages, the first line of the SIP header section is the request line, which contains the method type, request-URI, and protocol version etc. In the response messages, the first line is the status line, which contains a status code. SIP headers also contain IP addresses and port numbers used for signalling. The SIP body, separated from the header section by a blank line, is reserved for SDP, which is optional.

SIP Headers

In the following example SIP request message, NAT replaces the IP addresses in the header fields to translate to outside network addressing.

INVITE 212@10.10.10.5 SIP/2.0

Via: SIP/2.0/UDP 10.66.66.122:16382

From: 733@10.66.66.122

To: 212@10.10.10.5

Call-ID: 56ac6612@10.66.66.122

Contact: 733@10.66.66.122

Route: <sip:acme@10.66.66.122:5060>

Record-Route: sip:acme@10.66.66.122:5060

The way an IP address translation is performed depends on the type and direction of the SIP message. Discuss the following examples ;

{Assume example using a requesting messages with NAT Table}

Inbound Request	To:	Replace domain with local address
(from public to private)	From:	None
	Call-ID:	None
	Via:	None
	Request-URI:	Replace ALG address with local address
	Contact:	None

	Record-Route:	None
	Route:	None
Outbound Response (from private to public)	To:	Replace ALG address with local address
	From:	None
	Call-ID:	None
	Via:	None
	Request-URI:	N/A
	Contact:	Replace local address with ALG address
	Record-Route:	Replace local address with ALG address
	Route:	None
Outbound Request (from private to public)	To:	None
	From:	Replace local address with ALG address
	Call-ID:	None
	Via:	Replace local address with ALG address
	Request-URI:	None
	Contact:	Replace local address with ALG address
	Record-Route:	Replace local address with ALG address
	Route:	Replace local address with ALG address
Outbound Response (from public to private)	To:	None
	From:	Replace ALG address with local address
	Call-ID:	None
	Via:	Replace ALG address with local address
	Request-URI:	N/A
	Contact:	None
	Record-Route:	Replace ALG address with local address
	Route:	Replace ALG address with local address

SIP Body

The SDP information in the SIP body includes IP addresses the ALG uses to create channels for the media/audio streams.

The following excerpt from a sample SDP section shows the fields that are translated for resource allocation.

o=user 3524322 86249321 IN IP4 10.66.66.122

c=IN IP4 10.66.66.122

m=audio 32833 RTP/AVP 0

SIP messages can contain multiple media/audio streams. In this example, an INVITE message sent from a SIP client to a SIP server might have the following fields:

o=user 3524322 86249321 IN IP4 10.66.66.122

c=IN IP4 10.66.66.122

m=audio 32833 RTP/AVP 0

c=IN IP4 10.66.66.122

m=audio 32835 RTP/AVP 0

c=IN IP4 10.66.66.122

m=audio 32837 RTP/AVP 0

Cisco IOS, Cisco CUBE and Junos OS supports up 12 channels per call.

SIP NAT Scenario

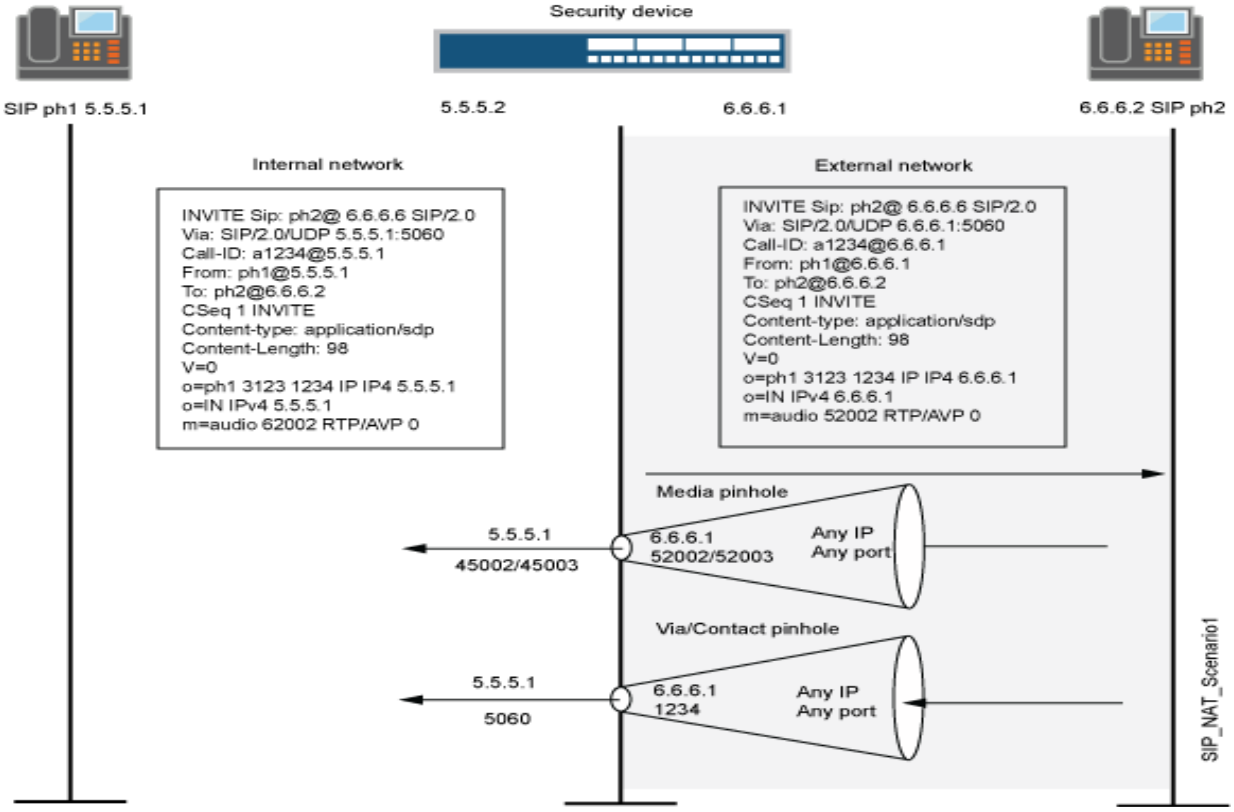
Diagrams below show a SIP call INVITE and 200 OK. ph1 sends a SIP INVITE message to ph2. Note how the IP addresses in the header fields—shown in bold font—are translated by the device. {DISCUSS}

The SDP section of the INVITE message indicates where the caller is willing to receive media. Note that the Media Pinhole contains two port numbers, 52002 and 52003, for RTCP and RTP. The Via/Contact Pinhole provides port number 5060 for SIP signaling.

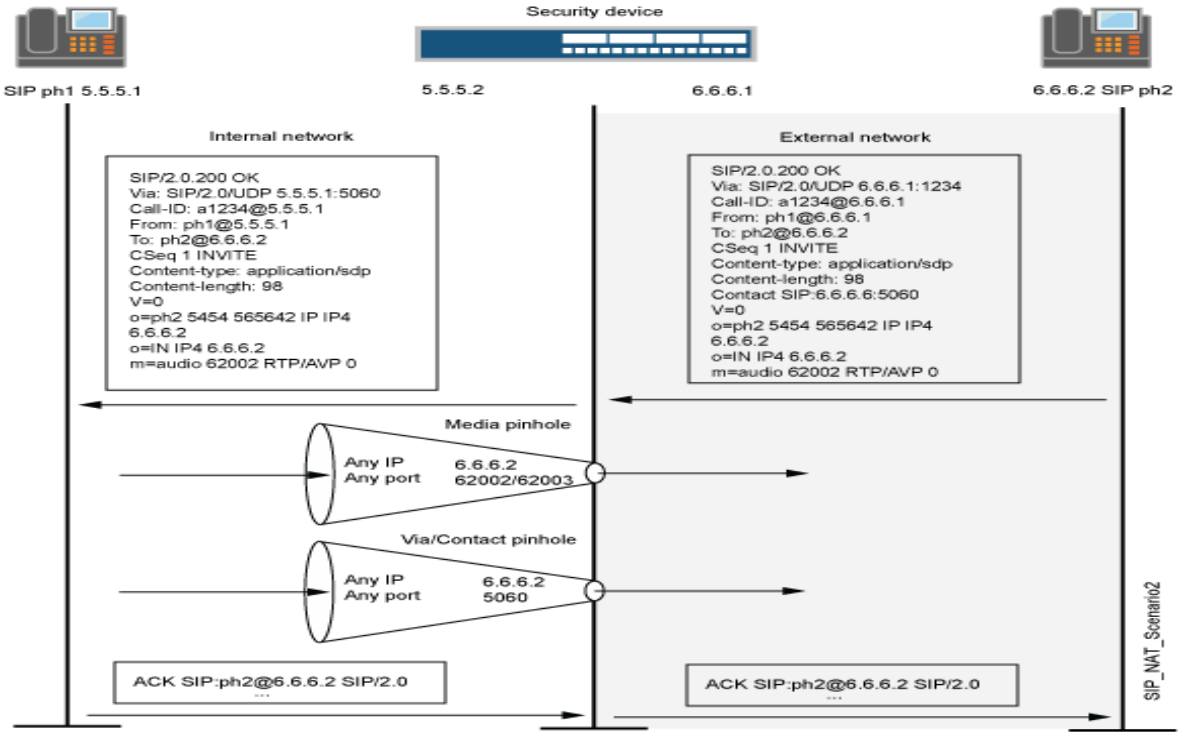
Observe how, in the 200 OK response message, the translations performed in the INVITE message are reversed. The IP addresses in this message, being public, are not translated, but gates are opened to allow the media stream access to the private network.

ALG for VoIP Overview

SIP NAT Scenario 1



SIP NAT Scenario 2



ALG for VoIP Overview

Classes of SIP Responses

SIP responses provide status information about SIP transactions and include a response code and a reason phrase. SIP responses are grouped into the following classes:

- Informational (100 to 199)—Request received, continuing to process the request.
- Success (200 to 299)—Action successfully received, understood, and accepted.
- Redirection (300 to 399)—Further action required to complete the request.
- Client Error (400 to 499)—Request contains bad syntax or cannot be fulfilled at this server.
- Server Error (500 to 599)—Server failed to fulfil an apparently valid request.
- Global Failure (600 to 699)—Request cannot be fulfilled at any server.

SIP Responses

Informational	100 Trying	180 Ringing	181 Call is being forwarded
	182 Queued	183 Session progress	
Success	200 OK	202 Accepted	
Redirection	300 Multiple choices	301 Moved permanently	302 Moved temporarily
	305 Use proxy	380 Alternative service	
Client Error	400 Bad request	401 Unauthorized	402 Payment required
	403 Forbidden	404 Not found	405 Method not allowed
	406 Not acceptable	407 Proxy authentication required	408 Request time-out
	409 Conflict	410 Gone	411 Length required
	413 Request entity too large	414 Request URL too large	415 Unsupported media type
	420 Bad extension	480 Temporarily not available	481 Call leg/transaction does not exist
	482 Loop detected	483 Too many hops	484 Address incomplete
	485 Ambiguous	486 Busy here	487 Request canceled
	488 Not acceptable here		
Server Error	500 Server internal error	501 Not implemented	502 Bad gateway