INTEROP LABS

What is SIP?

SIP is the Session Initiation Protocol. In IP and traditional telephony, network engineers have always made a clear distinction between two different phases of a voice call. The first phase is "call setup," and includes all of the details needed to get two telephones talking. Once the call has been setup, the phones enter a "data transfer" phase of the call using an entirely different family of protocols to actually move the voice packets between the two phones. In the world of VoIP, SIP is a call setup protocol that operates at the application layer. You may have also heard of H.323, an ITU protocol with similar function.

SIP is a very flexible protocol that has great depth. It was designed to be a general-purpose way to set up real-time multimedia sessions between groups of participants. For example, in addition to simple telephone calls, SIP can also be used to set up video and audio multicast meetings, or instant messaging conferences. In this document, we'll focus on SIP's capabilities for VoIP, and how it sets up calls that then use RTP (the Real-time Transport Protocol) to actually send the voice data between phones.

SIP also has great breadth as it does more than just handle call setup. The table below shows the five major functions within SIP from a VoIP point of view.

Function	Description
User location and registration	End points (telephones) notify SIP proxies of their location; SIP determines which end points will participate in a call.
User availability	SIP is used by end points to determine whether they will "answer" a call.
User capabilities	SIP is used by end points to negotiate media capabilities, such as agreeing on a mutually supported voice codec.
Session setup	SIP tells the end point that its phone should be "ringing;" SIP is used to agree on session attributes used by the calling and called party.
Session management	SIP is used to transfer calls, terminate calls, and change call parameters in mid-session (such as adding a 3-way conference).

One of the wonderful things about SIP is that it is a text-based protocol modeled on the request/response model used in HTTP. This makes it easy to debug because the messages are easy to construct (if you're a developer) and easy to see (if you're a network manager). Contrasted with H.323, SIP is an exceedingly simple protocol. Nevertheless, it has enough powerful features to model the behavior of a very complex traditional telephone PBX.

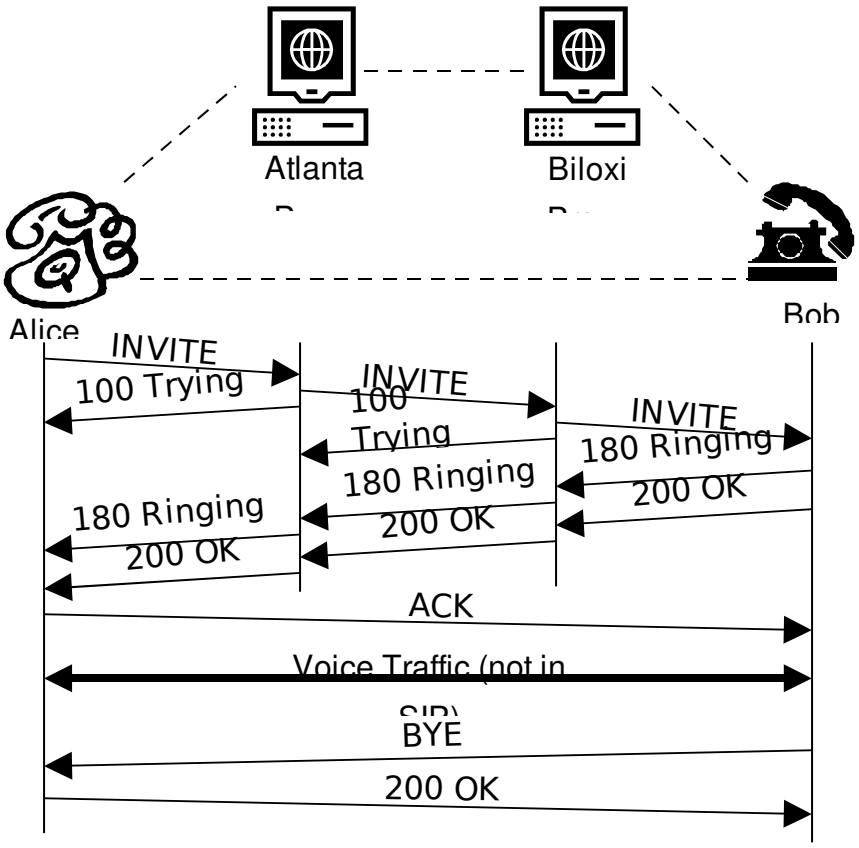
SIP can run over IPv4 and IPv6 and it can use either TCP or UDP. The most common implementations, though, use IPv4 and UDP. This minimizes overhead, thereby speeding performance.

Although two SIP devices can talk directly to each other, they generally will use an intermediary system that acts as a SIP proxy. Note the SIP proxy only participates in the SIP messages---once the call is set up, the phones send their voice traffic directly to each other without involving the proxy. SIP proxies are very helpful in offloading tasks and simplifying implementation of end station telephones. As an example, a SIP phone might want to make a call to another phone at extension 4094. Although the phone could have some magical way of translating that 4094 into an IP address or location, it typically will simply send its call request to its own SIP proxy. The job of the SIP proxy is to know what "4094" really means: Is that a phone? An auto-attendant system? Perhaps several phones, all to be rung at once?

We normally think in telephony of phones as having numeric addresses. In SIP, an end station has a SIP URI (a form of URL) that identifies it and is used in the SIP protocol. Because phones generally have numeric keypads, the phone is responsible for translating what you dial (such as extension 4094) into a SIP URL (such as sip:4094@voip.ilabs.interop.net). You can learn more about how SIP URIs, traditional telephone numbers, DNS, and IP addresses all interact in our white paper on "ENUM."

This diagram shows a SIP dialog involving two parties (Alice and Bob) and their SIP proxy servers, the Atlanta and Biloxi. In this case, the SIP messages have been heavily abbreviated to show the flow of traffic.

Although the diagram here shows that the proxies do not participate in the SIP protocol once Alice acknowledges that Bob has picked up the phone, not every call will work that way. A proxy may elect to "stay in the middle" of the conversation even after the call is connected to provide some mid-call features, such as conferencing services, or accounting. Note that even if the proxy is in the middle of the call, we're still only talking about the SIP part of the call---the voice traffic will generally go directly from one phone to another once the call is set up.



Another common operation in SIP is called Registration. In our example call, this might be how the Biloxi proxy learned where Bob was located. The registration capability is especially useful in an environment where phones do not have static IP addresses (such as a DHCP environment or when a phone travels with its owner). In SIP, the registration server can be co-located with the proxy server or they could be different systems. Bob is also not limited to registering from a single location. He could have SIP phones at home and at the office that both register with the SIP server. Then, it is the responsibility of the proxy server to decide which phones to "ring" when a call comes in for Bob. With SIP, that could mean selecting a single phone to ring, or just ringing all the phones at once.

Because SIP is used for call control, features such as voice mail and auto-attendant are not part of the SIP protocol itself. Instead, they are provided by end points that send and receive calls themselves. This means that a VoIP network based on SIP has no real parallel to the "PBX" in traditional telephony. You may hear the term "SIP Server" or "SIP PBX" used to describe the SIP proxy server, but the functionality is quite different. However, it is possible to integrate some traditional PBX features, such conferencing into a SIP proxy server. For example, the Asterisk SIP proxy server used as part of our InteropLabs demonstration can include both voice mail and auto-attendant. In other cases, such as a conferencing server with its heavy digital signal processing requirements, you might want a separate dedicated device.

To give you an idea of how simple SIP is, we've included a SIP message here: an idea of what Alice's original INVITE to Bob might look like. In this message, the Session Description Protocol (RFC 2327) part of the INVITE is not shown; SDP is where the voice traffic characteristics, such as choice of audio encoder, would be indicated. SIP's easy-to-read format has made implementation and debugging of SIP easier than other similar protocols, such as H.323.

```
INVITE sip:bob@biloxi.com SIP/2.0
Via: SIP/2.0/UDP pc33.atlanta.com;branch=z9hG4bK77ds
Max-Forwards: 70
To: Bob <sip:bob@biloxi.com>
From: Alice <sip:alice@atlanta.com>;tag=1928301774
Call-ID: a84b4c76e66710@pc33.atlanta.com
CSeq: 314159 INVITE
Contact: <sip:alice@pc33.atlanta.com>
Content-Type: application/sdp
Content-Length: 142
```