Understanding SIP

Today's Hottest Communications Protocol
Comes of Age

Ubiquity

Understanding SIP

Today's Hottest Communications Protocol Comes of Age

Introduction

The growing thirst among communications providers, their partners and subscribers for a new generation of IP-based services is now being quenched by SIP – the Session Initiation Protocol. An idea born in a computer science laboratory less than a decade ago, SIP is the first protocol to enable multi-user sessions regardless of media content and is now a specification of the Internet Engineering Task Force (IETF).

Today, increasing numbers of carriers, CLECs and ITSPs are offering such SIP-based services as local and long distance telephony, presence & Instant Messaging, IP Centrex/Hosted PBX, voice messaging, push-to-talk, rich media conferencing, and more. Independent software vendors (ISVs) are creating new tools for developers to build SIP-based applications as well as SIP software for carriers' networks. Network equipment vendors (NEVs) are developing hardware that supports SIP signaling and services. There is a wide variety of IP phones, User Agents, network proxy servers, VOIP gateways, media servers and application servers that all utilize SIP.

Gradually, SIP is evolving from the prestigious protocols it resembles – the Web's Hyper Text Transfer Protocol (HTTP) formatting protocol and the Simple Mail Transfer Protocol (SMTP) email protocol – into a powerful emerging standard. However, while SIP utilizes its own unique user agents and servers, it does not operate in a vacuum. Comparable to the converging of the multimedia services it supports, SIP works with a myriad of preexisting protocols governing authentication, location, voice quality, etc.

This paper provides a high-level overview of what SIP is and does. It charts SIP's migration from the laboratory to the marketplace. It describes the services SIP provides and the initiatives underway that will spur its growth. It also details the key features that distinguish SIP among protocols and diagrams how a SIP session takes place.

A New Generation of Services

Flexible, extensible and open, SIP is galvanizing the power of the Internet and fixed and mobile IP networks to create a new generation of services. Able to complete networked messages from multiple PCs and phones, SIP establishes sessions much like the Internet from which it was modeled.

In contrast to the longstanding International Telephony Union (ITU) SS7 standard used for call setup and management and the ITU H.323 video protocol suite, SIP operates independent of the underlying network transport protocol and is indifferent to media. Instead, it defines how one or more participant's end devices can create, modify and terminate a connection whether the content is voice, video, data or Web-based.

SIP is a major upgrade over protocols such as the Media Gateway Control Protocol (MGCP), which converts PSTN audio signals to IP data packets. Because MGCP is a closed, voice-only standard, enhancing it with signaling capabilities is complex and at times has resulted in corrupted or discarded messages that handicap providers from adding new services. Using SIP, however, programmers can add new bits of information to messages without compromising connections.

For example, a SIP service provider could establish an entirely new medium consisting of voice, video and chat. With MGCP, H.323 or SS7, the provider would have to wait for a new iteration of the protocol to support the new medium. Using SIP, a company with locations on two continents could enable the medium, even though the gateways and devices may not recognize it.

Moreover, because SIP is analogous to HTTP in the way it constructs messages, developers can more easily and quickly create applications using popular programming languages such as Java. Carriers who waited years to deploy call-waiting, caller ID and other services using SS7 and the Advanced Intelligent Network (AIN) can deploy premium communications services in just months with SIP.

This level of extensibility is already making its mark in growing numbers of SIP-based services. Vonage, a service provider targeting consumer and small business customers, delivers over 20,000 lines of digital local and long distance calling and voice mail to over customers using SIP. Deltathree, which provides Internet telephony products, services and infrastructure for service providers, offers a SIP-based PC-to-Phone solution that lets PC users call any phone in the world. Denwa Communications, which wholesales voice services worldwide, delivers PC to PC and Phone to PC caller ID, voice mail as well as conference calling, unified messaging, account management, self-provisioning and Web-based personalized services using SIP.

Understanding SIP

Today's Hottest Communications Protocol Comes of Age

While some pundits predict that SIP will be to IP what SMTP and HTTP are to the Internet, others say it could signal the end of the AIN. To date, the 3G Community has selected SIP as the session control mechanism for the next-generation cellular network. Microsoft has chosen SIP for its real-time communications strategy and has deployed it in Microsoft XP, Pocket PC and MSN Messenger. Microsoft also announced that its next version of CE.net will include a SIP-based VoIP application interface layer, and is committed to deliver SIP-based voice and video calls to consumers' PCs.

In addition, MCI is using SIP to deploy advanced telephony services to its IP communications customers. Users will be able to inform callers of their availability and preferred method of communication, such as email, telephone or Instant Message. Presence will also enable users to instantly set up chat sessions and audio-conferences. With SIP, the possibilities go on and on.

A Historical Snapshot

SIP emerged in the mid-1990s from the research of Henning Schulzrinne, Associate Professor of the Department of Computer Science at Columbia University, and his research team. A co-author of the Real-Time Transport Protocol (RTP) for transmitting real-time data via the Internet, Professor Schulzrinne also co-wrote the Real Time Streaming Protocol (RTSP) – a proposed standard for controlling streaming audio-visual content over the Web.

Schulzrinne's intent was to define a standard for Multi-party Multimedia Session Control (MMUSIC). In 1996, he submitted a draft to the IETF that contained the key elements of SIP. In 1999, Shulzrinne removed extraneous components regarding media content in a new submission, and the IETF issued the first SIP specification, RFC 2543. While some vendors expressed concerned that protocols such as H.323 and MGCP could jeopardize their investments in SIP services, the IETF continued its work and issued SIP specification RFC 3261 in 2001.

The advent of RFC 3261 signaled that the fundamentals of SIP were in place. Since then, enhancements to security and authentication among other areas have been issued in several additional RFCs. RFC 3262, for example, governs Reliability of Provisional Responses. RFC 3263 establishes rules to locate SIP Proxy Servers.

RFC 3264 provides an offer/answer model and RFC 3265 determines specific event notification.

As early as 2001, vendors began to launch SIP-based services. Today, the enthusiasm for the protocol is growing. Organizations such as Sun Microsystems' Java Community Process are defining application program interfaces (APIs) using the popular Java programming language so developers can build SIP components and applications for service providers and enterprises. Most importantly, increasing numbers of players are entering the SIP marketplace with promising new services, and SIP is on path to become one of the most significant protocols since HTTP and SMTP.

The SIP Advantage: Open, Extensible Web-Like Communications

Like the Internet, SIP is easy to understand, extend and implement. As an IETF specification, SIP extends the open-standards spirit of the Internet to messaging, enabling disparate computers, phones, televisions and software to communicate. As noted, a SIP message is very similar to HTTP (RFC 2068). Much of the syntax in message headers and many HTTP codes are re-used. Using SIP, for example, the error code for an address not found, "404," is identical to the Web's. SIP also re-uses the SMTP for address schemes. A SIP address, such as sip:guest@sipcenter.com, has the exact structure as an email address. SIP even leverages Web architectures, such as Domain Name System or Service (DNS), making messaging among SIP users even more extensible.

Using SIP, service providers can freely choose among standards-based components and quickly harness new technologies. Users can locate and contact one another regardless of media content and numbers of participants. SIP negotiates sessions so that all participants can agree on and modify session features. It can even add, drop or transfer users.

However, SIP is not a cure-all. It is neither a session description protocol, nor does it provide conference control. To describe the payload of message content and characteristics, SIP uses the Internet's Session Description Protocol (SDP) to describe the characteristics of the end devices. SIP also does not itself provide Quality of Service (QoS) and interoperates with the Resource Reservation

Setup Protocol (RSVP) for voice quality. It also works with a number of other protocols, including the Lightweight Directory Access Protocol (LDAP) for location, the Remote Authentication Dial-In User Service (RADIUS) for authentication and RTP for real-time transmissions, among many others.

SIP provides for the following basic requirements in communications:

1. User location services

2. Session establishment

3. Session participant management

4. Limited feature establishment

An important feature of SIP is that it does not define the type of session that is being established, only how it should be managed. This flexibility means that SIP can be used for an enormous number of applications and services, including interactive gaming, music and video on demand as well as voice, video and Web conferencing.

Below is are some of other SIP features that distinguish it among new signaling protocols

• SIP messages are text based and hence are easy to read and debug. Programming new services is easier and more intuitive for designers.

• SIP re-uses MIME type description in the same way that email clients do, so applications associated with sessions can be launched automatically.

• SIP re-uses several existing and mature internet services and protocols such as DNS, RTP, RSVP etc. No new services have to be introduced to support the SIP infrastructure, as much of it is already in place or available off the shelf.

• SIP extensions are easily defined, enabling service providers to add them for new applications without damaging their networks. Older SIP-based equipment in the network will not impede newer SIP-based services. For example, an older SIP implementation that does not support method/ header utilized by a newer SIP application would simply ignore it.

• SIP is transport layer independent. Therefore, the underlying transport could be IP over ATM. SIP uses the User Datagram Protocol, (UDP) as well as the Transmission Control Protocol (TCP) protocol, flexibly connecting users independent of the underlying infrastructure.

• SIP supports multi-device feature levelling and negotiation. If a service or session initiates video and voice, voice can still be transmitted to non-video enabled devices, or other device features can be used such as one way video streaming.

The Anatomy of a SIP Session

SIP sessions utilize up to four major components: SIP User Agents, SIP Registrar Servers, SIP Proxy Servers and SIP Redirect Servers. Together, these systems deliver messages embedded with the SDP protocol defining their content and characteristics to complete a SIP session. Below is a high-level description of each SIP component and the role it plays in this process.

SIP User Agents (UAs) are the end-user devices, such as cell phones, multimedia handsets, PCs, PDAs, etc. used to create and manage a SIP session. The User Agent Client initiates the message. The User Agent Server responds to it.

SIP Registrar Servers are databases that contain the location of all User Agents within a domain. In SIP messaging, these servers retrieve and send participants' IP addresses and other pertinent information to the SIP Proxy Server.

SIP Proxy Servers accept session requests made by a SIP UA and query the SIP Registrar Server to obtain the recipient UA's addressing information. It then forwards the session invitation directly to the recipient UA if it is located in the same domain or to a Proxy Server if the UA resides in another domain.

SIP Redirect Servers allow SIP Proxy Servers to direct SIP session invitations to external domains. SIP Redirect Servers may reside in the same hardware as SIP Registrar Severs and SIP Proxy Servers.

The following scenarios demonstrate how SIP components work in harmony to establish SIP sessions between UAs in the same and different domains:
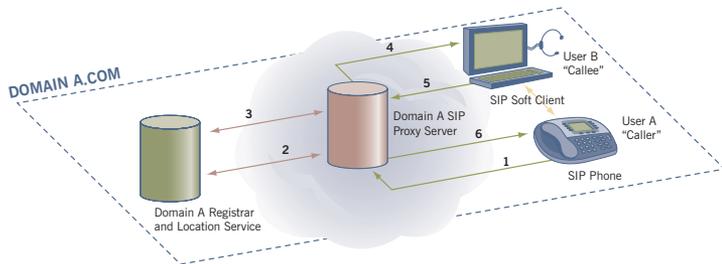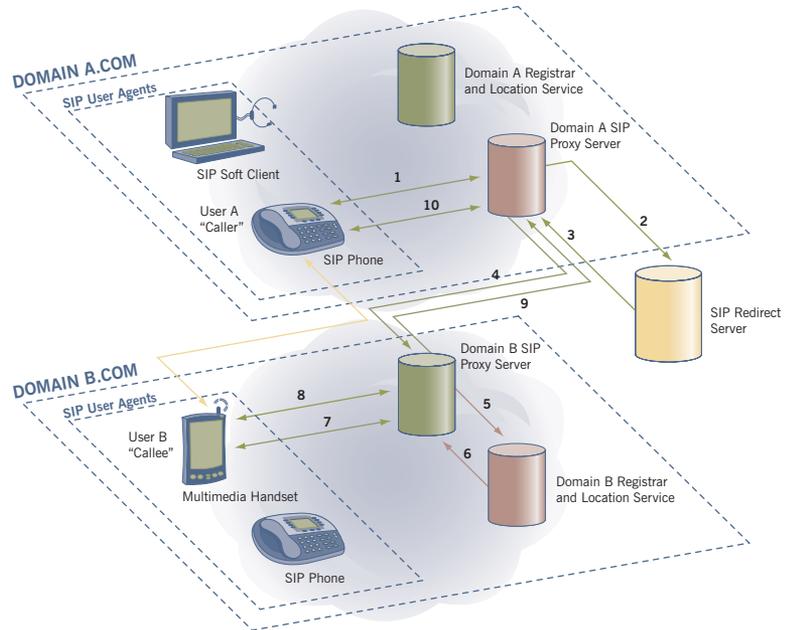
Establishing A SIP Session Within the Same Domain
The diagram below illustrates the establishment of a SIP session between two users who subscribe to the same ISP

and, hence, use the same domain. User A relies on a SIP phone. User B has a PC running a soft client that can support voice and video. Upon powering up, both users register their availability and their IP addresses with the SIP Proxy Server in the ISP's network. User A, who is initiating this call, tells the SIP Proxy Server he/she wants to contact User B. The SIP Proxy Server then asks for and receives User B's IP address from the SIP Registrar Server. The SIP Proxy Server relays User A's invitation to communicate with User B, including – using SDP – the medium or media User A wants to use. User B informs the SIP Proxy Server that User A's invitation is acceptable and that he/she is ready to receive the message. The SIP Proxy Server communicates this to User A, establishing the SIP session. The users then create a point-to-point RTP connection enabling them to interact.

1. Call User B
2. Query "Where is User B?"
3. Response "User B SIP Address"
4. 'Proxied' Call
5. Response
6. Response
7. Multimedia Chanel Establised

Non-SIP Queried
(i.e. Database Lookup)

SIP Signaling

RTP

Establishing A SIP Session In Dissimilar Domains

The difference between this scenario and the first is that when User A invites User B – who is now using a multimedia handset – for a SIP session the SIP Proxy Server in Domain A recognizes that User B is outside its domain. The SIP Proxy Server then queries the SIP Redirect Server – which can reside in either or both Domain A or B – for User B's IP address. The SIP Redirect Server feeds User B's contact information back to the SIP Proxy Server, which forwards the SIP session invitation to the SIP Proxy Server in Domain B. The Domain B SIP Proxy Server delivers User A's invitation to User B, who forwards his/her acceptance along the same path the invitation travelled.

1. Call User B
2. Query "How do I get to User B, Domain B?"
3. Response "Address of Proxy Controller for Domain"
4. Call 'Proxied' to SIP Proxy for Domain B
5. Query "Where is User B?"
6. User B's Address
7. Proxied Call
8. Response
9. Response
10. Response
11. Multimedia Channel Established

Non-SIP Queried
(i.e. Database Lookup)

SIP Signaling

RTP

Seamless, Flexible, Extensible: Looking Ahead With SIP
Able to connect users across any IP network (wireline LAN and WAN, the public Internet backbone, mobile 2.5G, 3G and Wi-Fi and any IP device (phones, PCs, PDAs, mobile handsets), SIP opens the door to a wealth of lucrative new possibilities that improve how businesses and consumers communicate. Used alone, SIP-based applications such as VOIP, rich media conferencing, push-to-talk, location-based services, Presence and IM offer service providers, ISVs, network equipment vendors and developers a plethora of new commercial opportunities. However, SIP's ultimate value lies in its ability to combine these capabilities as sub-sets of larger, seamless communications services.

Using SIP, service providers and their partners can customize and deliver a portfolio of SIP-based services that let subscribers use conferencing, Web controls, Presence, IM and more within a single communications session. Service providers can, in effect, create one flexible application suite that addresses many end user needs instead of installing and supporting discrete, "stovepipe" applications that are tied to narrow, specific functions or types of end devices.

By consolidating their IP-based communications services under a single, open standards-based SIP application framework, service providers can dramatically lower the cost of designing and deploying innovative new IP-based hosted services to their customers. This is the power SIP's extensibility can bring to the industry and the marketplace and the promise it holds out for us all.

About Ubiquity

Ubiquity Software Corporation develops and markets SIP-based communications software, including its award-winning SIP Application Server (SIP A/S) and Speak Conference Director. The SIP A/S is both a carrier-class deployment platform and a programmable, standards-based application-creation environment (ACE) that allows providers to develop and deploy next-generation converged communications services. Speak Conference Director is a rich media conferencing application built using the SIP A/S and sold to service providers worldwide. Use of the SIP A/S is extended to service providers and Independent Software Vendors (ISVs) through open, standards-based application programming interfaces (APIs). Ubiquity has corporate offices in the US, UK and Canada.

For more information, please visit www.ubiquitysoftware.com or email info@ubiquitysoftware.com.

www.ubiquitysoftware.com