Session Initiation Protocol – SIP

Launching the IP Communications Revolution

AN ALCATEL WHITE PAPER

July, 2003

ARCHITECTS OF AN INTERNET WORLD ALCATEL

Overview

Voice over IP (VoIP), IP, and IP Communications have been extensively written about since 1995. Vendors, customers, analysts, consultants, trade journals, and service providers have all argued the merits of placing and receiving telephone calls within corporate LANs, through corporate WANs, and into the Internet itself.

Early adopters, innovators, and data vendors touted lower costs and trotted out dozens of ROI models, while traditional telephony managers and manufacturers either created fear of VoIP or scrambled to hit the market with IP adaptations to their traditional PBX products.

Missing from the discussion was how IP Communications would benefit an enterprise more than just providing packet based dial tone and PBX features. This paper discusses the Session Initiation Protocol (SIP) and how it could take communications beyond network-level convergence to a new era of application-level convergence and interactive communications. SIP could do for interactive communications what H.323 did for VoIP and IP telephony.

H.323 was the key enabling technology behind VoIP and IP telephony, providing a way to put telephone calls over an IP-based infrastructure. Because H.323 did not provide a standard way of delivering all the features enterprises had grown to expect, vendors augmented the H.323 feature-set via proprietary extensions and H.323 has become little more than an underlying enabler.

SIP, because of its basis in web technologies, is positioned not to replace H.323 in this enabler role, but to enable the next revolution where communications services will be integrated seamlessly and intuitively into the daily work environment of the employee. The web services and web-based technologies available will determine whether SIP is successful, both from an ease of use and an ease of deployment perspective.

CTI applications of the past were mostly customer care applications largely because their complexity and the associated costs to customize and integrate required a strong ROI. With SIP, communication services are integrated into web applications and environments more simply and by a broader set of web developers who don't need to be CTI experts.

In this paper, practical business applications are blended with technical material to show how SIP might be used to help an organization compete in today's business or eBusiness climate. Lastly, this paper examines how Alcatel has approached SIP with its OmniPCX Enterprise IP Communications platform and how SIP figures prominently in Alcatel's vision of unified and interactive communications.

Setting the Stage

People carry cell phones, palm or pocket PC devices, have at least two email addresses, several telephone numbers and lease IP addresses wherever they log in. Sitting in an airport lounge it is not unusual to see road warriors with laptops open, cell phones pressed to their ears, and palm devices on to check calendars or look up names. What if it was possible to combine these devices, or at least minimize the number needed? What would you use to communicate with and in what format?

What is SIP?

According to Jean-Francois Rey, the next evolutionary step after VoIP, networks, and IP Telephony, “SIP will bring Internet telephony to users’ desktops.” SIP represents the capability to reach someone regardless of location or device. Just as email can find you once you are logged in, SIP can do the same. SIP signals multiple devices until it finds you, or respects the fact that you do not want to be found.

SIP is a signaling protocol. Its job is to broker communications between two devices, and once that communication is set up, SIP backs out. SIP has many of the characteristics of HTTP. It looks and acts like a URL. SIP addresses may look like an email address or even a telephone number. In fact, your email address is one way for SIP to find you. Some examples include:

sip:john@alcatel.com

sip:+8185551234@gateway.com

tel:+8185551234

SIP resides at the application layer of the network and establishes, modifies, and terminates multimedia sessions between intelligent devices. This is an important concept in the world of convergence. Data networks have traditionally been unintelligent, but with very smart endpoints and devices. Telephone networks are inherently very smart, but with unintelligent endpoints (telephones).

SIP extends the intelligence of a data network out to the end user at the edge, while allowing the lesser intelligent core to forward communications requests without much effort. This makes the data network run more efficiently and effectively by putting intelligence where it is needed most.

SIP also works between data and traditional telephony networks. It can broker a call between a manager’s analog home phone and an employee’s SIP enabled soft phone. If the employee is on a telephone call, it alerts the manager to another way of contacting the employee, such as sending an instant message using voice, perhaps via VxML, which is then converted to a simple text message and sent as an instant message to the employee’s PC. All the manager has to remember is the employee’s phone number.

SIP uses straightforward messages to set up, modify and terminate calls:

- **Invite** – just like it sounds
- **Ack** – call acceptance
- **Cancel** – terminate search
- **Register** – location of user (very important to the concept of presence)
- **Bye** – ends session

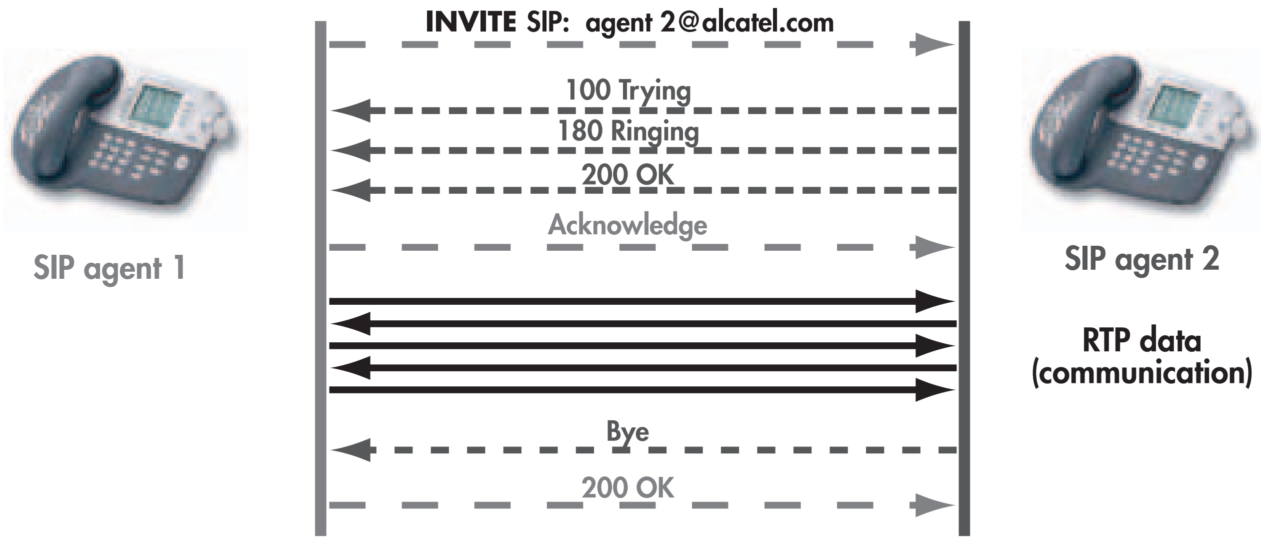
SIP also easily integrates with other protocols such as the Simple Mail Transport Protocol (SMTP) as well as working with Domain Name Servers (DNS) and the Lightweight Directory Access Protocol (LDAP). SIP can handle connectionless requests using the User Datagram Protocol (UDP), which is used for voice, or connection oriented requests using the Transmission Control Protocol (TCP), which would manage a chat session.

All these protocols form a communications fabric that makes it easier to locate someone, regardless of how they spell their name or what telephone number they are using. According to Frances Carincross, an industry visionary, the corporate communications network, along with the Internet, will be the “connective tissue” of an enterprise. This corporate connective tissue enables workers that are at home, in the office, or on the road to stay in touch with each other as well as customers and suppliers.

How Does it Work?

SIP provides connectivity and access. It enables communications between two SIP devices on a peer-to-peer basis, and acts as a client server by allowing SIP end users, called user agents (UA) to act as clients when initiating a request, or as servers when responding to a request. This could mean an attempt to call a SIP device results in a redirection to another device. The SIP infrastructure can have SIP servers to act as proxies, gateways, registrars and redirectors to bridge into multiple devices or other environments such as linking to the public switched telephone network (PSTN) or a corporate PBX.

On a peer-to-peer basis, SIP makes an invitation, which is either acknowledged or cancelled. Once the set up is complete, the Real Time Protocol (RTP) takes over to provide the actual communications. When the call is completed, the user hangs up and the SIP “Bye” message is sent out to end the session. A diagram of a simple SIP peer-to-peer call follows.



Calling into a non-SIP environment requires registering to be accessible on more than one device, or for changed locations. In this case, a SIP proxy server provides services such as redirecting a call to where you have moved, launching parallel requests (forking) to more than one device or connecting into a non-SIP device such as an analog or digital PBX telephone (SIP gateway).

SIP proxy servers also track users by dynamically updating changes and noting when the user is on line. This is the concept of “presence” and it enables a user to be found regardless of physical location.

The primary function of SIP is to set up a connection. How communications take place depends on the device being used, the user’s status, and location in the network (presence) and the preferences listed when the user registered. For example, the manager might be on a conference call at his/her desk. If the second line rings and the manager puts the conference call on hold to speak with a client, the other participants have to tolerate music on hold until the conference operator can mute the manager’s port.

Instead of disrupting a conference call the manger’s profile would state that if a call comes in from this particular client, identified by using caller ID, the caller has the option of engaging the manager in a chat session. An instant message on the manager’s PC screen might appear saying a client is trying to reach the manager (an “invite”) and he/she could accept (an ACK). The client could receive a message that says the manager is available to chat and see on his/her PC that a chat session has been set up. All of this depends on the manager’s network, devices, and desktop capabilities.

“There is already rapid growth in the number of novel and useful functions in new SIP-based systems,” said Jay Batson, founder and Chairman of Pingtel Corp., a pioneer and leader in SIP phones and technology. “Emerging SIP products are highly flexible and customizable, and are also rigorously tested for multi-vendor interoperability. This combination enables customers to find and deploy products that enhance their businesses much more simply and easily than with traditional voice systems.”

SIP is an evolving standard that is slowly gaining in acceptance and use in the telecommunications industry. The number of popular PBX features it supports is also growing. The cost of SIP phones, although currently higher than digital or IP phones, will decrease as demand increases and soft phones offer more cost-effective ways to realize the benefits of SIP in locations where their use is practical (e.g., call centers). Only recently have SIP telephones begun to reach price points similar to today's digital PBX phones.

Security, as with any IP-based device or service, is an issue, although there are security measures that can be taken such as embedding a firewall and NAT within the SIP proxy (Windows XP does this) or using stand-alone gateways.

Telecommunications service providers must still address the issue of mapping PSTN numbers and SIP addresses. Recently, the Department of Commerce said it would support the electronic numbering system, known as ENUM to allow consumers to specify a single identification system for their telephone numbers, e-mail, fax numbers, cell phone numbers and instant messaging addresses. The ENUM standard, known as E.164.arpa translates telephone numbers to Internet addresses and vice versa.

Enhanced 911 services are also a challenge with not just SIP, but all IP telephony devices due to the inherent mobility of IP. Since these devices essentially float in the network, ensuring E911 location information becomes a challenge. Within the SIP community several proposals have been offered ranging from installing GPS chip sets in the devices to having users simply log on to specify their location. Because a growing number of states require location information for E911 calls, this issue calls for some type of standardization across the SIP product line.

Why You Should Care

As workers become more mobile and IP Telephony, wireless LANs and WANs become reliable and stable, SIP offers the possibility of true enterprise roaming, access to multiple directories, find me or hide me services, choice of media, language and device, and the list goes on. As the communications fabric of an enterprise grows more reliable and flexible, workers will find themselves working more from home and on the road. Today's enterprise has an increasingly global element of employees, suppliers, and customers. The world is both the storefront and the workplace.

Service providers are also working to deliver SIP services, further extending the richness and reach of an enterprise. In their book Blown to Bits, How the New Economics of Information Transforms Strategy, Philip Evans and Thomas Wurster talk about the competitive power of information and how the explosion of connectivity and the dissemination of standards will help extend headquarters capabilities to every corner of the enterprise.

This connectivity revolution will mean that corporations can be more loosely structured, reduce costs by accessing workers knowledge where they live, creating a more dynamic collaboration environment for product development, shortening time to market, delivering timely innovations and serving customers in the ways they wish to be served, all through the enterprise's communications fabric. SIP provides a first step toward making this a reality, using the existing IP infrastructure and blending with the current telephony infrastructure. However, people still need to receive and work with voice mail and email. The basics don't go away, there are just more choices of how to gain access to these important capabilities.

Getting Started With SIP in Your Enterprise

The PBX is the core communications device providing mission critical voice services to an enterprise network and early applications of SIP will be used mostly for the flexibility it provides in the delivery of voice services to remote and mobile users, or for serving small locations or new facilities. In a paper written for the SIP Forum, Davis Sennreich, Pete Davis and Peter Blatherwick explain how the traditional PBX needs to integrate into the existing data network infrastructure to communicate with SIP devices. The challenge is how to make the migration from this traditional world into the richer converged world of IP Communications.

Most IT managers faced with traditional PBX platforms typically do not install IP-PBX products except in green field implementations. Traditional PBX platforms can be upgraded by adding gateways permitting communications with the data infrastructure to support IP phones. This is a first step.

The next step is to verify if the PBX can provide some type of SIP proxy or gateway services since SIP users will want to communicate with non-SIP or IP devices and make use of the company's circuit switched communications facilities for making and receiving calls. SIP devices use URLs not telephone numbers and these must be converted back and forth as SIP devices communicate with non-SIP devices. For the PBX administrator, it will be important to be able to assign extension numbers to SIP devices and "see" them as part of the PBX platform.

Another important consideration is how to serve small branch locations. While SIP IP phones could be installed and linked to the local data switch/router back to the HQ PBX, local dial tone will often be required along with fail over capabilities. Local PSTN gateways and a branch's ability to survive a WAN link failure and isolation from the HQ PBX is critical.

Not all PBX vendors support SIP and for some, it is only with the vendor's proprietary devices. Another question is will the PBX support SIP soft phones, such as the one embedded in Windows XP?

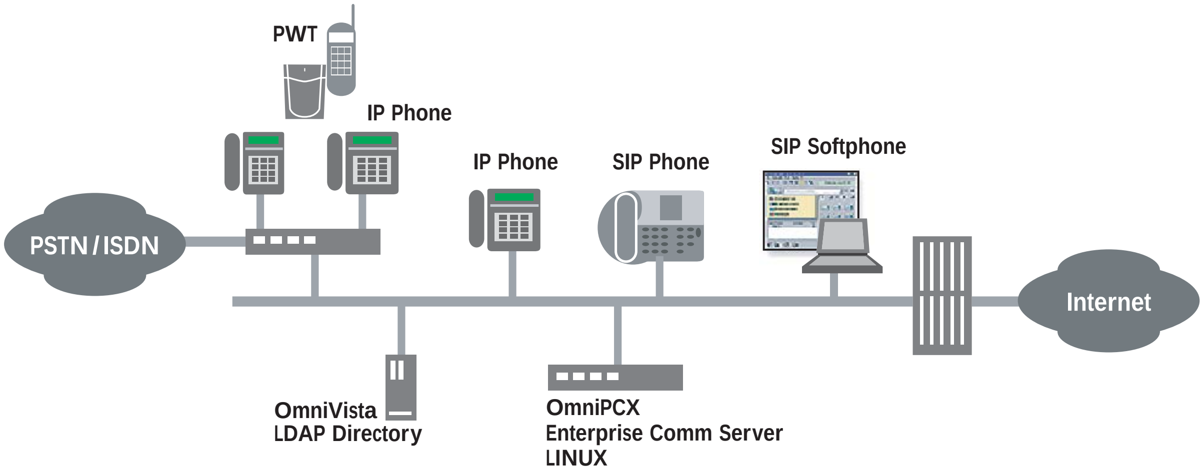
Understanding the capabilities and capacities of the enterprise network is critical. Although this paper does not address LAN and WAN issues, it should be noted that the corporate data fabric must be completely evaluated as to its ability to manage real time voice services. Issues such as resiliency, reliability, QOS, bandwidth management and security must each be reviewed with an eye on latency, jitter, and availability.

Alcatel's Approach to SIP

The Alcatel OmniPCX Enterprise IP Communications platform was designed to integrate easily into the data infrastructure. The Linux based communication server has an integrated SIP proxy and gateway such that SIP devices can be assigned a directory number and become part of the enterprise dial plan. SIP users can be placed in the OmniPCX directory. Since the OmniPCX Enterprise has LDAP directory capabilities, SIP users also become part of the larger corporate directory infrastructure and can be dialed by name, a unique capability of the OmniPCX Enterprise. These same SIP users can be "seen" as extensions within the OmniPCX and be assigned a class of service, linked to call accounting and tied into voice mail. Non-SIP extensions can be routed to SIP devices, including client, gateway, and proxy, using automatic route selection (ARS). Alcatel supports third party SIP devices through the OmniPCX Enterprise's gateway and proxy servers.

The OmniPCX Enterprise SIP gateway provides connections to analog, digital and IP telephones as well as PSTN access. If the SIP devices support features such as caller or called party ID, hold, transfer, forwarding, etc., the gateway ensures these are supported on the OmniPCX Enterprise as well. Again, SIP supports an evolving set of PBX features. The gateway can provide a message waiting indication to the SIP device when voice mail has been delivered.

The OmniPCX Enterprise SIP proxy server provides dynamic location and routing for SIP communications. The proxy implements parallel forking to enable a user to be called simultaneously on several SIP devices. The OmniPCX SIP registrar dynamically updates the location database, when it receives notification that users are online. Because the proxy server "speaks" with the enterprise's Domain Name Server (DNS), it handles connections to other SIP proxies in the same or different domain. This builds the "connective tissue" of an enterprise communications infrastructure. The SIP proxy also handles TCP and UDP network transport.



SIP is also a signaling component in Alcatel's e-Communications Center environment, which combines unified messaging, personal information management, personal routing management and Web-enabled soft phone capabilities. SIP and RTP ensure the signaling connection from the OmniPCX Enterprise to the Media Server, which drives the SIP user agent and VxML functions.

Conclusion

Today, according to Reye and Thyrland, equipment vendors, service providers, and application developers are all investing in SIP technology as it has become the standard for real-time converged communications in the IP world, as well as for presence and instant messaging.

The richness of connectivity and access promised by IP Communications is seeing its first demonstration in the use of SIP. Today's businesses are faced with the challenges of both a downturn in the economy and preparing for the recovery. This means having the flexibility to reach workers where they live, having flexible and virtual teams on the fly, keeping customers in touch and knowledgeable, and moving corporate knowledge and innovation around the enterprise as business needs change.

Vinton Cerf (Denning, 1999) said it well in a speech he gave in 1997

> ***"What we're trying to do is take the canvases on which we've been painting telephony applications and services for 125 years, and canvases on which we've been painting Internet technology, products and services for maybe 10 years and tear those two canvases apart. Then we want to reweave them together into one common canvas on which we'll now paint products and services that could not have been built except for the commingling of the two."***

Bibliography

Caircross, Frances. *The Death of Distance – How the Communications Revolution Is Changing Our Lives.* Boston: Harvard Business School Press,2001.

Denning, Pete. *Talking Back to the Machine – Computers and Human Aspiration*. New York. Copernicus, Springer-Verlag, 1999.

Evans, Philip and Wurster, Thomas. *Blown to Bits – How the New Economics of Information Transforms Strategy.* Boston: Harvard Business School Press, 2000.

Rey, Jean-Francois and Tyrland, Claude. *SIP Technology in the Enterprise.* Alcatel Telecommunications Review, 4th Quarter 2002.

Sinnreich, Henry, Davis, Pete and Blatherwick, Peter. *Enterprise IP Communications, A SIP Perspective on the New Converged World.* SIP Forum, http://www.sipforum.org

The SIP Center - a portal for the commercial development of the Session Initiation Protocol. http://www.sipcenter.com

Acknowledgements

This document was produced by Alcatel Internetworking, Inc. Authored and researched by Rudy Mazza, edited by Vicki Vaughn, outlined, reviewed, and approved by Alan Amrod, with significant contributions by Jeanne Bayerl, Marc Dayton, Erin Cavanaugh, and Ed Youngblood.

Technical Appendix: Session Initiation Protocol

What is SIP?

SIP (Session Initiation Protocol) is a signaling protocol used to establish the IP addresses and port numbers that end systems use to send and receive data. SIP is not a transport protocol (it doesn't transport data) – and the actual data packets do not follow the same path as the SIP packets. However, a mechanism does exist that allows a photo, business card info, or a Web page to be transported with the signaling packets.

The SIP function involves signaling only, not the actual data being communicated. The philosophy behind SIP is to keep the communications needed to provide interoperability between Internet entities separate and simple, building additional services on top of it.

SIP is an end-to-end, client-server session signaling protocol that allows two or more Internet entities to find one another. It creates the signaling to establish paths and then tears them down when the session is over. Since SIP is not a transport protocol, RTP (Real time Transport Protocol) is needed.

SIP does not play any part in ensuring the required QoS (quality of service) is available for the actual data transfer that follows as there is no way to synchronize SIP signaling with QoS requirements.

SIP depends on the Session Description Protocol (SDP) to carry out negotiation and identification. SIP supports session descriptions that allow participants to agree on a set of compatible media types. It also supports user mobility by proxy and redirecting requests to the user's current location.

The services that SIP provides include:

- **User location:** determination of the end system to be used for communications.
- **Call set-up:** ringing and establishing call parameters for both called and calling parties.
- **User availability:** determination of the willingness /availability of the called party to engage in communications.
- **User capabilities:** determination of the media and media parameters to be used.

Features of SIP include the following capabilities:

- **User mobility** (the use of the URL Internet address format) – sip:john.smith@alcatel.com

  SIP allows a session to be established without knowledge of an absolute IP or MAC address. This provides user mobility through the user himself declaring (programming) where he is physically located at one time, e.g., one day located at a phone in New York, the next at a phone in Boston.

- **Manual / automatic call control on top of the protocol.** With simple, dynamic programming SIP users can:

  – Redirect calls from unknown callers to an assistant

  – Reply with a Web page

  – Send a JPEG image with the session invitation to allow the called party to see who is calling

SIP is Internet telephony, not VoIP.

SIP Services

SIP allows tailoring applications to the needs of individual users. Some of the basic services that SIP can provide are detailed below.

Presence services are a new form of communication. Presence is possible due to the datagram nature of the Internet. Presence can provide information about:

- Presence on the Internet (which parties are available)
- Location: office, home, traveling
- Call state: ready, on another call, etc.
- Willingness: available, in meeting, etc.
- Preferred medium: text, voice, video, email
- Instant messaging (text chat and voice chat)

SIP allows additional features to be used, for example, sending a JPEG image and / or business card with the signaling – so that a called party can see who is calling. Two major features that differentiate SIP are:

Splitting / forking a call

SIP provides the ability to split an incoming call so that several extensions ring at the same time from the same single originating call. The first extension to answer gets the call. This feature is useful if the called party is identified as being in one of several locations, each with a separate phone number, e.g., at the main office, branch office, factory, or at home. This feature can also be used to call one number and ring both a manager and his or her assistant.

Replying with alternative media

When a SIP server receives a connection request, it can return a Web page to the calling SIP client. The page can be used to supply alternative numbers to call in the case of vacation time or absence due to illness, for example. The caller only has to click on the supplied Web icon to call the alternative number or party.

End users and / or third parties may program SIP. This adds tremendous flexibility in personalizing SIP communication services, allowing the user to dynamically change the response to calls. Examples include:

- Distinctive ringing: ring selected based on who is calling
- Discard all calls from a particular person / number outside business hours
- Redirect authenticated friends to my cell phone, everyone else to my assistant

SIP Issues

SIP packets use textual encoding, which may be considered inefficient, but it does allow much simpler debugging, and no special monitors are necessary to examine the SIP signaling packets.

Billing and accounting features are not yet defined since the current trend is to charge a flat rate for the SIP service, as with email. Issues about how to bill for calls that originate or end in a legacy telephony environment such as PSTN remain unresolved.

Emergency services (911 calls) are not included and are still under discussion.

Currently under consideration is for the transport of DTMF (dual tone multi frequency) to be in RTP and not SIP. This is because IVR (interactive voice response) is too difficult for SIP to handle, as would be the case, for example, in a flight reservation request over the phone: "Press 1 for…, press 2 for…, etc."

The supporters of SIP are playing this issue down by declaring that in the near future, this operation will be performed using a browser or softphone to access a Web site. However, IVR response using a normal telephone handset is not likely to disappear in the one to five year timeframe.

Since SIP is simple and effective, it will play a role in the near future for Internet communications.

SIP and QoS

SIP does not provide any form of QoS. In a practical sense RSVP could be used to provide the resources necessary for real-time media transfer; however, RSVP is not being deployed everywhere. A more realistic approach is to over-provision bandwidth or provision MPLS tunneling.

SIP and H.323

H.323 was originally designed for video conferencing and LAN telephony, whereas SIP is designed for multimedia Internet communication. Both SIP and H.323 define mechanisms for call routing, call signaling, capabilities exchange, media control, and supplementary services.

The advantage of SIP is that it is backed by IETF, one of the most important standards bodies, while the advantage of H.323 is that it has a much larger piece of the current market.

The following is a comparison of the two protocols:

- H.323 defines hundreds of elements, while SIP has only 37 headers, each with a small number of values and parameters.
- H.323 uses a binary representation for its messages, which is based on ASN.1, while SIP encodes its messages as text, similar to HTTP.
- H.323 is not very scalable as it was designed for use on a single LAN and has some problems in scaling. Newer versions have suggested techniques to get around this problem.
- H.323 is limited when performing loop detection in complex multi-domain searches. It can be done statefully by storing messages, but this technique is not very scalable. On the other hand, SIP uses a loop detection method by checking the history of the message in the header fields, which can be done in a stateless manner.

SIP Conclusion

SIP is promising because it is a simple, open protocol that can be deployed in carrier and enterprise networks and because it enables new multimedia applications in voice, data, and video. For service providers, H.323 is a legacy technology that will eventually give way to SIP because of the inherent simplicity of SIP and potential for new media-blending services (Internet telephony).

However, in the enterprise market, H.323 has gained significant support because of its manageability, reliability, and interoperability with the PSTN and will continue to do so for some time. Most serious implementers of SIP are considering a SIP-to-H.323 inter-working function / gateway as an essential element of any network.

There is a general consensus among standards organizations, companies, and technology experts that standardized procedures need to be specified to allow seamless inter-working between the two protocols. Bodies such as TIPHON (ETSI), IMTC, and IETF are working to address this topic.

www.alcatel.com/enterprise

Alcatel

26801 West Agoura Road
Calabasas, CA 91301 USA

Contact Center
(800) 995-2612 US/Canada
(818) 880-3500 Outside US

www.alcatel.com/enterprise

P/N 031347-01. 7/03

ARCHITECTS OF AN INTERNET WORLD