



ZULTYS

SMARTER COMMUNICATIONS

SIP 151

Author: Zultys Technical Support Department

This document tries to cover SIP, what it is, how it works, and why you will want to use it in your companies' communications plan. This document uses various sources which are not limited to

- SIP For Dummies (SIP Communications For Dummies(r),Wiley Publishing, Inc. ISBN: 978-0-470-38114-4)
- Convergence Technologies CTP Course (Comp-TIA ISBN: 978-0-470-38114-4)
- SDP: Session Description Protocol, Request for Comments 4566, Internet Engineering Task Force, April 2006.
- RTP: A Transport Protocol for Real-Time Applications, Request for Comments, Internet Engineering Task Force, July 2003.
- SIP: Session Initiation Protocol Request for Comments 3261, Internet Engineering Task Force, June 2002

1 Contents

2	<u>WHAT IS SIP, IN REFERENCE TO TELECOMMUNICATIONS?</u>	2
3	<u>SIP OPERATION</u>	3
3.1	USER AGENTS	3
3.2	SIP SERVERS	4
4	<u>PROTOCOLS</u>	5
4.1	SDP	5
4.2	RTP	6
4.3	RTCP	7
4.4	UDP	6
4.5	SIP	5
5	<u>SIP RESPONSE CODES</u>	7
6	<u>CODECS</u>	8
6.1	G711	8
6.2	G729	8
7	<u>PARAMETER DESCRIPTIONS</u>	8
8	<u>SIP CALL LIFE CYCLE</u>	9
8.1	EXAMPLE OF A SIP CALL	11
9	<u>SIP SECURITY</u>	12
10	<u>SIP SURVIVABILITY</u>	13
11	<u>SIP AND THE PSTN</u>	14
12	<u>IP TELEPHONY CHALLENGES</u>	15

2 *What Is SIP, in reference to Telecommunications?*

Session Initiation Protocol (SIP) is an open signaling protocol standard developed primarily by the *Internet Engineering Task Force (IETF)* in RFC 3261 for establishing, managing, and terminating real-time communications. SIP is an application layer peer-to-peer communication protocol. However, SIP does not transport the media itself: the transport is handled by codecs within the communications programs or devices.

SIP builds on a number of existing communications protocols and has rapidly become the standard for service integration.

- SIP has all but replaced H.323 for voice communications. H.323 is still being used in Video communications.
- SIP is modeled after HTTP, and in fact uses much of HTTP's semantics and syntax. Both SIP and HTTP use a plain text based language.
- SIP is very modular and extensible (like *XML*, or the *Extensible Markup Language*), allowing for integration with legacy systems and new and evolving technologies.
- These properties make SIP an ideal protocol for implementing a standards-based unified communications network.
- SIP is cross platform, therefore it works on virtually all operating systems
- SIP is open protocol; therefore a SIP network can be built using multiple manufactures for the Phone, Servers, and Network equipment. This allows the integrators to pick and choose the equipment that best fits their design and preferences.
- SIP is not limited to intranetworking, but can be used for internetworking, and joining multiple communication systems together that support SIP
- SIP Gateways can be implemented to communicate with legacy communication systems

The potential impact of SIP goes beyond internal communications within a business or enterprise. SIP has become a signaling standard for carrier networks. Service providers now provide SIP-based trunk services that can reduce costs and extend an enterprise's SIP environment into the public network.

3 SIP Operation

3.1 User agents

User agents (UAs) are applications installed on SIP endpoints, such as an IP phone, mobile phone, wireless device or PDA, or a laptop or desktop PCs that interface between the user and the SIP network. A UA can act as either a client or a server. When sending SIP requests, the UA acts as a *user agent client (UAC)*, and when servicing a request, it acts as a *user agent server (UAS)*. A *back-to-back user agent (B2BUA)* is an application that acts as an intermediary between two parties, but appears as an endpoint to both parties — like a middle man. It serves as both UAS and UAC simultaneously to process session requests.

3.2 SIP servers

SIP servers provide centralized information and services in a SIP network. There are several servers that make up what is commonly called a SIP server, their basic functions and names are listed below. Zultys rolls all of these servers in to a single server which could be the MX250 or MX30, commonly referred to the MX.

- **Registrar Server.** The Registrar authenticates and registers users when they come online, and stores information on the users' logical identities and the communications devices or physical entities (IP address) of the communication devices they can use. The devices are identified by their URIs.
- **Location Service.** The location service is a database that keeps track of users and their locations when they move about the network. The location service gets its input from the Redirect Server.
- **Redirect Server.** If users are not in their home domains, sessions need to be redirected to them. The redirect server maps a SIP request destined for a user to the URI of the device "closest" to the user. For example, if a call is destined for Steve.Rothenburg@Zultys.com and the user is on the road, the company's redirect server may reply to the caller's UA (or to the requesting proxy server) with the contact address of the user's mobile phone, so that the incoming call can be redirected to the mobile phone.
- **Proxy Server.** A proxy server takes SIP requests, processes them, and passes them downstream while sending responses upstream to other SIP servers or devices. A proxy server may act as both a server and a client, and may modify certain parts of a SIP request before passing it along. A proxy is involved only in the setup and teardown of a communication session. After user agents establish a session, communications occur directly between the parties. It is important to note that a proxy server is used for both intranet (local communications) as well as internet (NAT) communications
- **Presence Server.** Presence servers accept, store, and distribute presence information that allows users to see the availability of people they want to contact. The presence server has two distinct sets of clients:
 - *Presentities* (producers of information) provide presence information about themselves to the server to be stored and distributed.
 - *Watchers* (consumers of information) receive presence information from the server. Watchers can subscribe to certain users, much like instant messaging users choose which "buddies" to add to their list.

4 Protocols

4.1 SIP

Session Initiation Protocol (SIP) initiates and manages sessions between two or more parties; SIP is a signaling protocol only it cannot deliver media. SIP widely used for setting up and tearing down multimedia communication sessions such as voice and video calls over the Internet. SIP runs at the application-layer, and is end-to-end oriented protocol. It's Default port is 5060.

4.2 SDP

SIP commonly makes use of the Session Description Protocol (SDP) to describe the attributes of SIP sessions. SDP does not provide the content of the media form itself but simply provides a negotiation between two end points to allow them to agree on a media type and format. SDP parameters are encapsulated as the message body of a SIP request. SDP plays a similar role as that of H.245 in the H.323 world. Like SIP, SDP headers are encoded with ASCII text. The SDP headers are of the simple form <type>=<value>. The <type> is always a single character and <value> is a text string whose format is dependent on <type>. SDP is not really a protocol as much as it is a format for describing multimedia sessions. SDP headers specify:

- Session name and purpose
- Time(s) the session is active
- The media comprising the session
- Transport address and media format of the session
- Bandwidth to be used by the session
- Contact information for the person responsible for the session

A key component of SDP is the description of the media of the session. SDP media descriptions include:

- Media type (audio, video)
- Transport protocol (UDP, TCP, RTP)
- Media format (H.261, MPEG, etc.)
- Multicast address for IP multicast sessions
- Transport port for IP multicast sessions
- Remote address for IP unicast sessions
- Transport port for IP unicast sessions
- Session start and stop times

4.3 UDP

User Datagram Protocol (UDP) applications can send messages, sometimes known as datagrams, to other hosts on an Internet Protocol (IP) network without requiring prior communications to set up special transmission channels or data paths.

UDP uses a simple transmission model without implicit hand-shaking dialogues for guaranteeing reliability, ordering, or data integrity. Thus, UDP provides an unreliable service and datagrams may arrive out of order, appear duplicated, or go missing without notice. UDP assumes that error checking and correction is either not necessary or performed in the application, avoiding the overhead of such processing at the network interface level. UDP does not retransmit packets (voice communications can stand up to 5% loss of packets and remain intelligible). Its true benefit is its speed.

4.4 RTP

SIP sessions use Real-time Transport Protocol (RTP) for end-to-end transport of audio and/or video information. RTP defines a standardized packet format for delivering audio and video over the Internet. RTP can be delivered over TCP and/or UDP, most applications use UDP due to its speed. The main features of RTP are

- **Sequence number:** Used by the receiving client to detect lost packets and to play the audio or video packets in the correct order. This is important since, with UDP, there is no guarantee that the packets will arrive at the receiving client in the same order as they were transmitted (if they arrive at all).
- **Timestamp:** Used by the receiving client to play the packet stream using the same timing that was used during transmission. This is critical as the packets may experience varying amounts of delay as they are forwarded through the network. Delay variations, or jitter, result in decreased audio or video quality.
- **Payload type:** Indicates the encoding technique that was used to encode the audio or video information. The encoding technique is chosen to optimize quality or bandwidth usage.
- **Delivery Monitoring:** Monitor that a packet is delivered via RTCP.

4.5 RTCP

RTCP provides out-of-band control information for an RTP flow. It partners RTP in the delivery and packaging of multimedia data, but does not transport any data itself. It is used periodically to transmit control packets to participants in a streaming multimedia session. The primary function of RTCP is to provide feedback on the quality of service being provided by RTP. This is done with periodic reports on

- Fraction of packets lost since the last report
- Total number of packets lost since the last report
- Delay variation (jitter).

Clients use this information to control adaptive encoding algorithms. RTCP does not actually transmit any data itself; this is the job for UDP and RTP, RTCP provides Quality of Service.

5 SIP Response Codes

- 1xx Informational trying, ringing, call is being forwarded, queued
 - 100 Trying
 - 180 Ringing
 - 181 Call is being forwarded
- 2xx Success OK
 - 200 OK
- 3xx Redirection Moved permanently, moved temporarily, etc
 - 300 Multiple Choices
 - 301 Moved Permanently
 - 302 Moved Temporarily
- 4xx Client error Bad request, unauthorized, not found, busy, etc
 - 400 Bad Request
 - 401 Unauthorized
 - 482 Loop Detected
 - 486 Busy Here
- 5xx Server error Server error, not implemented, bad gateway, etc.

- 500 Server Internal Error
- 6xx Global failure Busy everywhere, does not exist anywhere, etc
 - 600 Busy Everywhere

6 Codecs

An audio codec is an application that compresses/decompresses digital audio data according to a given streaming audio format. The MX supports G.711 and G.729 codecs.

6.1 G.711

G.711, also known as Pulse Code Modulation (PCM), is a very commonly used waveform codec. There are two main compression algorithms defined in the standard, the μ -law algorithm (used in North America & Japan) and A-law algorithm (used in Europe and the rest of the world).

- 3.5-kHz bandwidth at 8kHz
- Amplitude of 8 bits/sample
- Transmission rate of 64 Kbps
- Compression delay of 0.75ms

6.2 G.729

G.729 is a compressed codec requiring the following

- Transmission rate of 8 Kbps
- Compression delay of 10ms

7 Parameter Descriptions

- **Call-ID:** Uniquely identifies a particular session.
- **CSeq:** A monotonically increasing sequence number used to identify the sequence of requests associated with a given Call-ID. From A SIP URL that identifies the initiator of the request. May include a "friendly name" (e.g. John Smith).

- **To:** A SIP URL that identifies the recipient of the request. May include a “friendly name”.
- **Via:** Indicates the path taken by the request so far. The Via parameter is used to prevent looping of requests, assures that replies take the same route as requests and assists in unusual routing situations.

Messages in SIP are not directly sent to the user. They are sent to proxy server, which is responsible for routing and delivering messages to the called party. The proxy servers relay call signaling. There are several types of proxy servers, like Call-stateful, transaction-stateful and stateless proxies.

- **Call-stateful** proxies track call state and provide a lot of services but they can be slower.
- **Transaction-stateful** proxies track the requests and responses but not the call state or session.
- **Stateless** proxies just receive requests, forward them and forget them.

Proxy servers also provide forking, which is used for trying different locations for a request. There are also Redirect Servers. These servers redirect the requestor to the other servers instead of forwarding them. Redirection is useful if a user moves or changes the provider. There are also SIP Registrars, which accept the registration request of the users.

8 SIP Call Life Cycle

The life cycle of a SIP based session is broken down into several parts, first is to initiate a session. In order to initiate a session SIP has to locate the user. The user can use different user agents in different places. To find the user, the server relies on the location service. After locating the user SIP delivers a description of the session in order to inform the user agent. SIP only transmits the descriptions SIP does not know anything about the session itself. Session Description Protocol (SDP, RFC 2327) is most common protocol to describe sessions. After locating the user and conveying the description of the session SIP conveys the response of the user agent. The user agent can accept, reject, or forward the session. If the session is accepted then an active session will be initiated.

SIP is based on the request-response paradigm. The methods are

- **Invite:** The Invite method indicates that the user is invited to a session. A session description is also included in the message body
- **Ack:** The Ack method is used to confirm a session establishment. This method can only be used with Invite requests.
- **Bye:** The Bye method terminates the sessions.
- **Cancel:** The Cancel method is used to cancel a pending Invite.



Technical Publications

- **Options:** The Options method is used to query the server for its capabilities.
- **Register:** The Register method is used to bind a permanent address to the current location of the user.

In order to establish sessions the caller has to send an Invite to the user with whom the caller wants to talk to. This request is then sent to the address of the user. The addresses in SIP are like email addresses. For example if the caller wants to talk to the user support@Zultys.com then the SIP address for the user is sip:Support@Zultys.com. A session is established only after the following events have happened

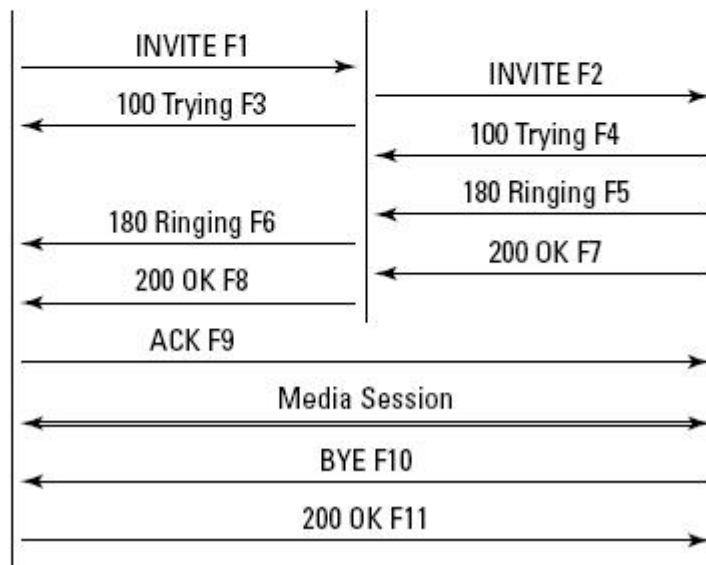
- An invite is sent to the user by the caller
- The invite is received by the user
- The user sends an ACK back to the caller

8.1 Example of a SIP Call


Michael's softphone



John's SIP phone



1. Michelle@smallcompany.com (the UAC) initiates a session by inviting Tony@bigcompany.com and sending this request to the proxy server at small company.com. Michelle's UA generates an INVITE request, which is sent to the proxy at smallcompany.com. The INVITE message contains *Session Description Protocol (SDP)* parameters that define the types of media she is capable of accepting and where she wants the media to be sent.
2. The proxy at smallcompany.com performs a DNS SRV record lookup for SIP services at bigcompany.com since bigcompany.com is a foreign domain. This record lookup returns proxy.bigcompany.com, which is then resolved to a physical IP address by DNS. Michelle's INVITE request is then forwarded to the proxy server at bigcompany.com.

3. The bigcompany.com proxy server receives and processes the invitation, and looks up Tony's contact in the location database of the Registrar (physical IP address of the UA).
4. The location database of the Registrar returns host@ 192.168.1.100 where Tony is currently located.
5. The bigcompany.com proxy server forwards the INVITE request to Tony's UA at host@192.168.1.100.
6. The UAS at host@192.168.1.100 asks Tony whether he wants to accept the call. Tony may hear a ring, see a text message, or see a blinking LED.
7. Tony's acceptance is sent back through the big company.com proxy, which forwards it to the small company.com proxy, which forwards it to Michelle's UA. The body of Tony's acceptance includes SDP parameters defining the selected media chosen from what Michelle had originally offered and where Tony wants the media to be sent.
8. Michelle's UA responds to the acceptance with an ACK (acknowledgement) directly to Tony's UA, which tells Tony's UA that Michelle is ready to start the call.
9. At the end of the conversation, Tony hangs up his phone. His UAC sends a BYE message directly to Michelle's UA.
10. Michelle's UAC responds with a 200-OK message directly to Tony's UA, which ends the session.

From this simple example you can see that some messages are handled by the Proxy server and some are handled by the UA itself. Another note from this is that all communications require a response, usually in the form of an ACK or an OK.

9 SIP Security

Because SIP relies heavily on an IP-based network and utilizes a plain-text language similar to HTTP, SIP-enabled applications are potentially vulnerable to many of the same security threats that plague corporate networks and the Internet today. These may include

- Authentication

- Authorization
- Privacy issues
- Denial of service
- Buffer overflow attacks
- *SPIT (SPAM over Internet Telephony)*

Fortunately, many of the same solutions to these problems are effective for securing SIP implementations as well. These may include:

- HTTP Digest Authentication using MD5 for challenge/ response user authentication
- Signaling channel encryption using *Transport Layer Security (TLS)* for end-to-end session security
- *Certificate Authorities (CA)* for authentication in networks using SRTP and TLS
- *Secure Real-Time Transport Protocol (SRTP)* or *IP Security (IPSEC)* using *Advanced Encryption Standard (AES)* encryption to provide authentication, confidentiality, and integrity for protection of the media (to prevent eavesdropping, for example) Other security concerns are not necessarily unique to SIP, but are nonetheless threats that must be addressed with innovative solutions.

10 SIP Survivability

Another major concern for enterprises is survivability of their telecommunications and network systems. Traditional TDM telephone equipment and the PSTN are commonly perceived as highly reliable, dedicated networks and systems compared to distributed, “best effort” networks such as the Internet. Although a high degree of redundancy is built into the individual components of expensive TDM PBX systems, having redundant PBX systems in remote locations capable of providing seamless failover during a major systems malfunction or failure, or a catastrophic event is rather uncommon. By comparison, IP-based communications systems and networks, in addition to having built-in redundant components, are commonly deployed as duplexed systems or server farms in multiple locations. SIP survivability and redundancy includes benefits like re-routing DID’s and quickly redirecting calls to alternate data centers (or having the Service Provider do so after a specified timeout) — something traditional PSTN circuits don’t provide. Even SIP endpoints can failover to third party SIP gateways as part of the “survivable intelligent edge.”

11 SIP and PSTN

Gateways link the SIP calls with PSTN Trunks; they are widely deployed and used by VoIP users every day. Before SIP and VoIP, enterprises connected their internal PBX-based telephone systems to carriers via dedicated TDM (Time Division Multiplexing) trunks. Companies paid for them whether they were idle or busy, and incurred toll and tariff charges, particularly expensive for long-distance calls. Today, many companies integrate voice and data over IP networks and link their sites using wide area networks to reduce communications costs within the enterprise. An enterprise SIP proxy joins with a carrier SIP proxy, or even PSTN circuits via a gateway with the appropriate security protections established between them. The IP circuit continues to carry e-mail, Web, data, and other corporate traffic as it does today, and voice is simply added to the mix as another IP application. On-net calls ride the carrier's VoIP backbone. Off-net calls ride the carrier IP network until the "last mile" where a gateway converts VoIP to TDM for calls to PSTN parties.

Gateways may be internal to the SIP server such as the MX250 or MX30, or it may be a standalone device such as the MX25 which is strictly a Gateway device used to convert FXO and PCM trunks to SIP or the reverse. The MX25 can also be used to bring SIP services to a TDM switch that cannot support SIP.

SIP trunks change how you make connections to carriers. SIP trunks offer a number of advantages, including:

- **PSTN origination/termination and cost savings:** Many SIP service providers support origination/termination services directly to the PSTN from their SIP networks. This allows the enterprise to reduce monthly recurring costs associated with multiple TDM circuits by deploying a single IP pipe to the provider network. Companies may cut back on multiple TDM circuits by upstreaming calls, that is smaller locations may only have a few trunks for local dial tone, but outgoing calls are made and received from a larger branch that has PCM or SIP circuits
- **DID and Toll-free Number Mobility:** These features take advantage of the fact that SIP is geographically agnostic. Calls destined to local or toll-free numbers can be automatically rerouted over the service provider SIP network to another enterprise location. For enterprises, this system offers great flexibility in providing a local presence in all their markets while routing calls to a centralized call center for more efficient service.

12 IP Telephony Challenges

There are three main challenges in the path from the Internet technologies of today to a VoIP telephone system.

- First is that call-signaling capability needs to be brought to packet switching.
- The second is that quality of service must be controlled.
- The third challenge is building a converged PSTN/VoIP network. The transition from PSTN to VoIP will be a gradual process because of significant technical and business issues to be solved. Since VoIP and PSTN networks will coexist for many years, a converged network will need to be built in order to bridge the gap between the two. This converged network will allow calls to originate on a VoIP network and terminate on a PSTN network (and vice versa). This converged network will make use of VoIP gateway devices to bridge the two networks.