SIP : Session Initiation Protocol

EFORT
http://www.efort.com

SIP (Session Initiation Protocol) as defined in IETF RFC 3261 is a multimedia signaling protocol used for multimedia session establishment, modification and termination. The session may be a voice session, a videotelephony session, a chat session, a fax session, or whatever multimedia session.
SIP is a component that can be used with other IETF protocols to build a complete multimedia architecture. These protocols are RTP/RTCP for transporting real-time data and providing QoS feedback, SDP for describing multimedia sessions and MSRP for transporting data messages in the context of a chat session. Therefore, SIP should be used in conjunction with other protocols in order to provide complete services to the users.
SIP incorporates elements of two important Internet protocols : HTTP (client/server design, URL, response codes) and SMTP (text-encoding scheme, header style).
SIP is not only a session control protocol but also a service control protocol. It can be used between call servers, Application Server (AS) and Media Resource Function (MRF) to provide multimedia services such as supplementary services of telephony, presence, multimedia conferencing, instant messaging, etc.
This tutorial presents the SIP capabilities, and the SIP protocol through its requests and answers.

1 The Internet Multimedia Stack

The above figure shows the four-layer Internet Multimedia Protocol stack.
The top layer is the application layer. This includes signaling protocols such as SIP and media transport protocols such as Real-time Transport Protocol (RTP). The figure includes H.323, which initially was an alternative signaling protocol to SIP developed by the International Telecommunication Union (ITU). Session Description Protocol (SDP) is shown above SIP in the protocol stack because it is carried in a SIP message body. SDP enables to describe the media involved in the multimedia session established with SIP. Audio, video are examples of media. Because SIP can use any transport protocol, it is shown interacting with both TCP, UDP and possibly with SCTP (Stream Control Transmission Protocol).

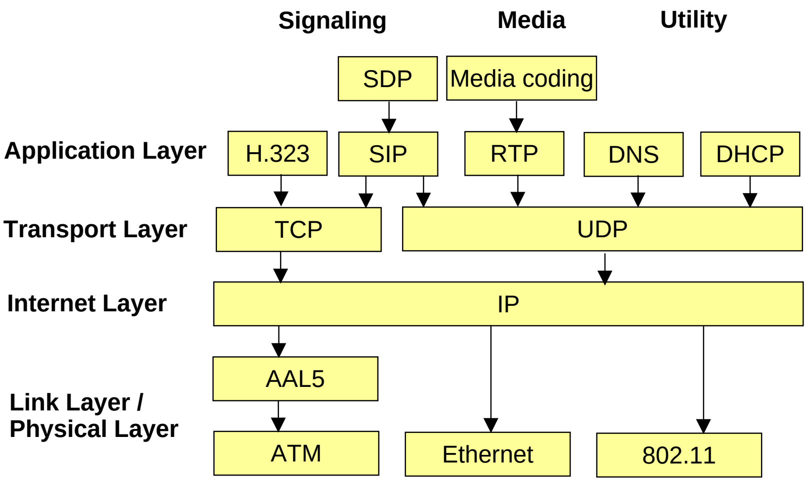

Figure 1 : The Internet Multimedia Protocol stack

2 What SIP is and is not

SIP has a number of capabilities :

- session control for multimedia sessions (multimedia session may be voice, video, data or whatever session). An example of a data session is a chat session. SIP enables establishing, modifying and terminating a multimedia session. A session may be started with the audio media only and be modified to support audio and video (videotelephony).
- Service control in an IN-like architecture : The extensive capabilities provided by the existing IN architecture can be supported by SIP to enable the introduction of enhanced services such as customized ring back tone, IP centrex, SIP trunking, Rich communications suite, etc.
- Supplementary services : Most supplementary services can be supplied by the SIP protocol. Examples of supplementary services are call forward, call transfer, call hold, calling line identity presentation, calling line identity restriction, call completion on busy subscriber, conferencing. In a PSTN network, ISUP signaling protocol provides supplementary services.
- Presence and instant communications : Instant messaging allows users to exchange short text messages in real time, in contrast to E-mail that is a store an forward communication, well suited for large messages and attachments. While some E-mail may take considerable time to reach the destination, instant messages are transferred in real time and are well suited for interactive. The advantage of using SIP for presence and instant messaging lies in the common infrastructure with many other communication services.
- User mobility, service mobility and session mobility : User mobility is one of the unique features of SIP and is basically a by-product of the similarity of SIP with HTTP and SMTP. By redirecting incoming calls to any of the user devices, users can be reachable irrespective of location and device. Service mobility means that the user can get his services from any registration point even when he is in a visited network. The services follow the user movement. Session mobility means that the session can be moved from one network to another network. For example a 4G user may start his VoIP session with 4G coverage which continues in 2G or 3G where a circuit switched domain need to handle that call.
- Authentication and charging information : SIP carries all the information to authenticate and charge the user.

However, the following capabilities are not supported by SIP :

- SIP is not a resource reservation protocol, so it cannot assure QoS. It is a session control protocol and not a bearer control protocol.
- SIP is not a transfer protocol such as HTTP, designed to carry large amounts of data. It is designed to transport small amounts of data required to setup multimedia communications. Small amounts of data not related to call setup, such as short text messages for instant messaging are also well suited for SIP.
- SIP is not designed to manage interactive sessions, once sessions are established. For example, it is not designed to exercise floor control in conference sessions. Extra control protocols are required.
- SIP does not handle terminal mobility. This is a feature directly handled by the mobile packet switched domain.

3 SIP entities

A SIP architecture consists of the following entities :

- A User Agent is an application that contains both a User Agent Client (UAC) and a User Agent Server (UAS).
- A UAC is an application that initiates a SIP request.
- A UAS is an application that contacts the user when a SIP request is received and that returns a response on behalf of the user. The response accepts, rejects, or redirects the request.
- A Proxy Server acts as both a client and a server. It accepts requests from other clients, either responding to them or passing them on to other servers, possibly after modification.
- A Redirect Server accepts requests from clients and returns zero or more addresses to that client. Unlike a Proxy server it does not initiate requests. Unlike a User Agent Server it does not accept calls.
- A Registration server is a server that accepts REGISTER requests. A Registration server is typically co-located with a Proxy or Redirect Server.
- A Back-To-Back User Agent is a type of SIP device that receives a SIP request, then reformulates the request and sends it out as a new request. Responses to the request are also reformulated and sent back in the opposite direction.
- A SIP Gateway translates SIP signaling into ISUP/H.323/Q.931/etc. signaling and converts voice signals into RTP packets.

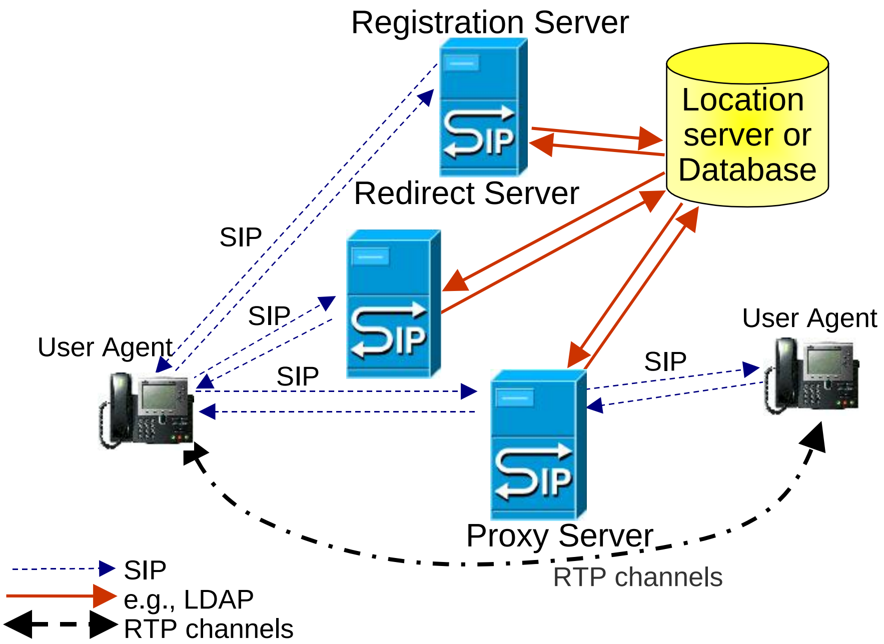

Figure 2 : SIP entities

4 SIP requests

There is a growing number of SIP request types, known as methods.

- The INVITE method is used to establish media sessions between user agents. In telephony, it is similar to a Setup message in Q.931, or IAM in ISUP. A final response to an INVITE is always acknowledged with the ACK method.

- The ACK is used to acknowledge final responses to INVITE requests. Final responses to all other requests are never acknowledged. Final responses are defined as 2XX, 3XX, 4XX, 5XX and 6XX class responses. For 2XX responses, the ACK is end-to-end, but for all other final responses, it is done on a hop-by-hop basis when stateful proxies are involved.
- The BYE method is used to terminate an established media session. A session is considered established if an INVITE has received a success class response (2XX) and an ACK has been sent. A BYE is only sent by user agents participating in the session, never by proxy or other third parties. It is an end-to-end method so responses are only generated by the other agent. A user agent responds with a 486 Unknown Call Leg to a BYE for an unknown Leg. A BYE can be sent by either the calling party or called party in the session.
- The OPTIONS method is used to query a user agent or server about its capabilities and discover its current availability. The response to the request lists the capabilities of the user agent or server. A proxy never generates an OPTIONS request.
- The CANCEL method is used to terminate pending searches or call attempts. It can be generated by either user agents or proxies. A user agent uses the method to cancel a pending call attempt it had earlier initiated. A proxy can use the method to cancel pending parallel branches after a successful response has been proxied back to the user agent client.
- The REGISTER method is used by a user agent to notify a SIP network of its current IP address and the URLs for which it would like to receive calls. SIP registration looks similar to cell phone registration on initialization. REGISTER is used to register and deregister. An expires duration is specified corresponding to the registration duration. If set to 0, it means de-registration.

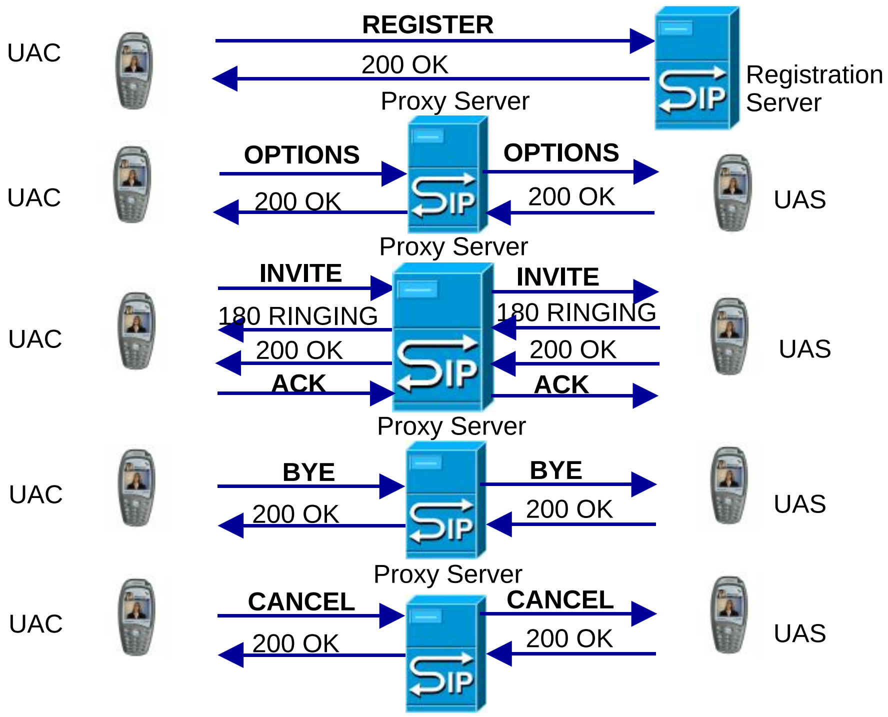

Figure 3 : SIP requests

5 SIP Responses

SIP Response messages indicate either call progress information or final status information. Response messages contain a Status-Code and a Reason-Phrase.
The Status-Code is a three digit integer that indicates the outcome of the request. The Reason-Phrase provides a textual description intended for humans.

- 1xx: Provisional -- request received, continuing to process the request;
- 2xx: Success -- the action was successfully received, understood, and accepted;
- 3xx: Redirection -- further action needs to be taken in order to complete the request;
- 4xx: Client Error -- the request contains bad syntax or cannot be fulfilled at this server;
- 5xx: Server Error -- the server failed to fulfill an apparently valid request;
- 6xx: Global Failure -- the request cannot be fulfilled at any server.

6 Format of SIP requests and answers

A SIP request consists of a request line, several headers (e.g., From, To, Call-ID, etc), an empty line and a message body. The message body is optional; some requests do not carry it.
A Request line has three elements : method, Request URI, and protocol version. The method indicates the type of request. The request-URI indicates the next hop, which is where the request has to be routed. Finally, the protocol version is SIP/2.0.
INVITE sip:mark.rich@tn.com SIP/2.0
BYE sip:+33672727272@tn.com SIP/2.0
A SIP response consists of a status line, several headers, an empty line, and a message body. The message body is optional; some responses do not carry it.
SIP/2.0 180 Ringing
SIP/2.0 486 Busy Here

7 Additional SIP requests

Eight additional SIP requests have been defined to extend SIP capabilities :

- The INFO method is used by a user agent to send call signaling information to another user agent with which it has an established media session. The request is end-to-end and is never initiated by proxies. An INFO method contains signaling information such as midcall event as DTMF tones.
- The PRACK method is used to acknowledge the receipt of provisional responses (1xx) except the 100 Trying response. The final responses (2XX, 3xx, 4xx, 5xx, 6xx) are acknowledged by the ACK method in the case where the initial method is INVITE. In the case where a provisional response such as 180 Ringing is critical in determining the call state, the confirmation of this provisional response may be necessary, using the PRACK message.
- To request and receive a notification when a certain event occurs is supported in SIP by the SUBSCRIBE and NOTIFY request types. For example, the CCBS (Call Completion on Busy Subscriber) service can be handled by these messages : When the calling party sending an INVITE message is informed that the receiver is busy, the calling party sends a SUBSCRIBE request to the callee to request notification when the callee is available to establish a session. When the called party sends a NOTIFY request indicating that the user is now available, the calling party immediately establishes the session.

- PUBLISH is used for publishing event state. PUBLISH is similar to REGISTER in that it allows a user to create, modify, and remove state in another entity which manages this state on behalf of the user.
- The REFER method allows a third party such as a controller to request the caller setup a call with a resource. It can be used for call transfer.
- The MESSAGE method is used to transport instant messages (IM) using SIP. IM usually consists of short message exchanged in near-real time by participants engaged in a “conversation.” MESSAGEs may be sent within a dialog or outside a dialog, but they do not establish a dialog by themselves. The actual message content is carried in the message body as a MIME attachment.
- The UPDATE method is used to update the media description before the session starts.

References

- J. Rosenberg, H. Schulzrinne, G. Camarillo, A. Johnston, J. Peterson, R. Sparks, M. Handley, E. Schooler, “ SIP : Session Initiation Protocol ”, RFC 3261, June 2002.