Session Protocols

Kimmo Raatikainen

kimmo.raatikainen@cs.helsinki.fi

Lession Outline

- Session Initiation Protocol or SIP
  - Overview of SIP
  - Basic protocol architecture
    - architectural elements
    - signalling
    - session characteristics
    - security
  - Notes on
    - Telephony servives
    - Multiparti sessions
    - Mobility support

Overview

- Session Initiation Protocol (SIP) initiates, modifies, and terminates network sessions.
- Current applications of SIP focus on interactive multimedia sessions such as Internet phone calls or multimedia conferences
- In setting up sessions, SIP acts as a signaling protocol, offering services similar to telephony signaling protocols such as Q.931 or ISUP, but in an Internet context.
- SIP does not reserve resources or establish circuits (virtual or real) in the network.

Overall IETF Multimedia Architecture

- **Real-Time Transport Protocol** (RTP) for transporting audio, video and other time-sensitive data
- **Real-Time Streaming Protocol** (RTSP) for setting up and controlling on-demand media streams
- **Media Gateway Control Protocol** (MGCP) and **Megaco** (also known as H.248) for controlling media gateways
- **Session Description Protocol** (SDP) for describing multi-media sessions
- **Session Announcement Protocol** (SAP) for announcing multicast sessions
- **Telephony Routing over IP** (TRIP) for locating the best gateway between the Internet and the PSTN
- a suite of resource management and multicast address allocation protocols

Session Setup

- Discovery of a user wherever located so that a description of the session can be delivered to the user.
- The session description is delivered as an opaque body.
- SIP includes information about the caller, the purpose for the invitation, its urgency, and parameters of the session itself.

User Identity

- Users can maintain the same identifier even as they change attachment points to the network or use different devices (personal mobility).
- A single user identity may simultaneously be represented by a number of network terminals.
- A single network terminal or user may be reachable using multiple identifiers,
- Depending on logic in SIP network elements, SIP can deliver requests to any or all of these network locations.

Basic Architecture

- Four logical types of entities participate in SIP:
  - user agents, registrars, and proxy and redirect servers.
- User agents initiate requests and are usually their final destination.
- Registrars keep track of users within their assigned network domain
- Proxy servers are application-layer routers that forward SIP requests and responses.
- Redirect servers receive requests and then return the location of another SIP user agent or server where the user might be found.

Message Routes

- SIP messages originating at a user agent traverse one or more SIP proxy servers and then reach one or more SIP user agents.
- However, SIP user agents can also communicate directly with each other.
- Indeed, it is common that only the first request exchange travels along a chain of proxies
- all subsequent requests exchanged directly between the two user agents

Signalling

- SIP request URIs look similar to e-mail addresses, consisting of
  - a user name part and a host name part, plus a number of parameters.
- Each proxy or redirect server looks at the request URI and uses it, plus any other header fields it finds useful, to route the request.
- Typically, the server uses backend location servers to map the request URI to a new destination.
- The request URI is then rewritten to reflect the result of this lookup process.

Signalling

- A user agent may decide to send all requests to a fixed, typically local, outbound proxy server, which then routes the request according to the request URI.
- SIP messages can be transported on just about any communication protocol.
  - Different protocols can be used between proxies and user agents while forwarding a single message.
  - Generally, UDP is preferred since it avoids the TCP connection setup and teardown overhead.

Signalling

- SIP requests and responses are grouped into transactions.
- A transaction consists of
  - one request,
  - followed by zero or more provisional responses that indicate call progress,
  - a final response that indicates whether the request succeeded or failed, and,
  - for INVITE requests, an ACK request from the originator of the first request to confirm the arrival of the final response.

Signalling

- Responses in SIP are self-routing, tracing their way back through the same set of servers the requests visited.
- Each server records its address and port number in a Via header;
  - these are then reversed by the destination.

Describing and Changing Sessions

- Current implementations of SIP use SDP to describe multimedia sessions.
- SDP is a description format that allows each party to declare what media streams it wants to receive and its receive capabilities.
- These streams and capabilities are expressed in a simple textual format.
- SDP is carried as a message body in SIP, separated by a blank line from the SIP headers.

Describing and Changing Sessions

- Each media stream entry indicates the destination address and port number and lists the encodings supported by the receiver, among which the sender is allowed to alternate during the session.
- SDP also allows to schedule sessions into the future or describe recurrent sessions.
- If either party wants to add a media stream to the session, change the media destination network address, or drop a media stream, it simply sends another INVITE request to the other party.

SIP Message Format

- SIP uses a textual encoding of its messages with a syntax very similar to HTTP.
- The use of a plain text representation is often said to incur additional message length overhead and processing costs.

Forking

- SIP differs from other signaling protocols in that it allows a call request to fork
- a server can send out two or more requests to different destinations (branches) based on one incoming request, either at once or in sequence if an earlier request failed.
- This feature supports a number of advanced telephony services:
  - call forwarding to voice mail, automatic call distribution (ACD), and user location

Reliability

- SIP has to take care of reliability on its own.
- It has two reliability mechanisms, one for INVITE requests and one for all other requests.
- INVITE requests are different since the final response can be delayed by several tens of seconds from the arrival of the request.
- Clients retransmit INVITE requests until a provisional response arrives, and servers retransmit responses until confirmed by an ACK request.
- Clients retransmit other request methods until the final respons e arrives.

Locating Users

- Proxy and redirect servers use a logical entity called a location server to route requests.
- The location server itself can use any method to map an incoming call to a next-hop destination.
- Primary means of locating users is by using the location mappings installed through user registrations.
- User agents periodically register with their local registrar server
  - often collocated with the local proxy
- Registrations can also contain additional information about the capabilities and preferences of a user

Session Characteristics

- Sessions are described at two levels:
  - the overall characteristics and
  - if a media session, the session description.
- The overall session characteristics include
  - the caller' s name, address, and organization, the callee's name and address, the session's urgency, and its subject.
- The media in the session can be described in any mutually acceptable format;
  - only SDP is used at this point.
  - SDP was originally designed to describe multicast sessions, but is used in SIP for unicast sessions.
  - SDP indicates the receive capabilities and destination addresses and ports for any number of media streams.

Security

- Both signaling and media need to be secured against eavesdropping and alteration
- Authentication is important since there is no trusted third party (the phone company) to ensure the accuracy of the information contained in the session setup request.
- Also, proxy servers may only want to offer services to registered users, and registrations must be protected from malicious alteration.

Security

- SIP inherits the basic and digest authentication mechanisms from HTTP
- Basic authentication simply requires that the sender of a request provide a plain text password.
  - This is clearly picketfence security, but may be acceptable if SIP messages are carried using transport-layer or IP-layer security.
- Digest authentication uses a challenge-response approach that checks whether the originator of a request is privy to a shared secret.
- SIP requests can be signed with PGP
- Authentication using the CHAP challenge-response mechanism used by Point-to-Point Protocol (PPP) has been proposed.

Security

- SDP can convey session keys for media streams, as long as the signaling request is encrypted.
- Unless all proxies are trusted, only end-to-end encryption of the SIP message body can ensure confidentiality.
- Currently, public key cryptography using the PGP format has been defined, but any mechanism developed for e-mail should easily transfer to the SIP environment.
  - One difference is that part of the message needs to remain unencrypted to allow servers to forward the request appropriately.

Telephony Service

- Many traditional telephony services are provided by the baseline SIP specification
  - caller ID,
  - name/number mapping services (e.g., 800 and 900 services),
  - variations on call forwarding,
- SIP also attempts to provide building blocks that can be used to construct services, rather than including specific message headers or methods for each service.

Multiparty Sessions

- SIP sessions can use three different multiparty conferencing architectures:
  - Full mesh: In a full mesh, every participant builds a signaling leg with every other participant and sends an individual copy of the media stream to the others.
  - Mixer: A mixer or bridge takes several media streams and replicates them to all participants.
  - Network-layer multicast: Neither full mesh nor mixers scale to large conferences. These are most efficiently supported by network-layer multicast.

Mobility

- One of the central tasks of SIP is to locate one or more IP addresses where a user can receive media streams, given only a generic, location-independent address identifying a domain.
- This personal mobility allows a user to change communications devices without making the caller aware of these details.
- This mechanism makes it easy to offer precall terminal mobility, but does not directly support mid-call mobility

Mobility

- Standard mobile IP can be used to hide the change of IP addresses during calls, but address filtering and dog-legged routing are of particular concern for latency-sensitive voice calls.
- Another possibility is to use mid-session location updates using SIP.
- As long as media tools can change destination addresses in mid-session, this requires no further cooperation from the network.
- However, it does not work for long-lived TCP connections that may be part of a session (e.g., a chat session).