Signaling For VoIP

Seminar Report
Dec - 2000

Kapil Bajaj
Roll Number : 00305905
Department of Computer Science & Engg., IIT Bombay
Guides: Prof. D. Manjunath and Prof. G. Sivakumar

November 3, 2000

Abstract

Voice over IP provides an opportunity to design a global multimedia communications system that may eventually replace the existing telephony system. This report discusses the advantages of VoIP over traditional telephony. It provides a overview of two protocols: Session Initiation Protocol(SIP) and H.323, which are currently prevalent for VoIP, focusing mainly on call setup and management. Finally, a comparison between two protocols is made.

1 Introduction

Voice Over IP(VoIP) or IP Telephony means providing traditional Plain Old Telephony Services(POTS) on the packet switched IP network. It is the means for voice transport over IP network. Signaling protocols for VoIP establish and control sessions or calls

Voice Over IP allows user to make long distance or international call at a much lesser price. IP Telephony integrates the voice with the data applications that can be transported on IP network. It provides a way of using same Local Area Network(LAN) in offices for data transport as well as for voice transport which can replace Private Branch Exchanges(PBX). VoIP aims to provide new services compared to the traditional POTS like broadcasting, multicasting media based endpoint selection, security by encryption and authentication of caller ID and even some services like caller ID, call forwarding, etc., at no extra cost. Graphical User Interface(GUI) for VoIP can be more readily customised. Signal flow in traditional telephony must go through all intermediate switches whereas IP Telephony can also provides direct communication between users.

Internet Telephony should provide at least the same level of Quality of Service(QoS) as PSTN provides. Delays in the communication should be less than a threshold value. Security, Interoperability within products from different vendors, Integration with PSTN are to be considered for VoIP to be able to replace POTS.

For two or more parties to communicate via IP Telephony there should be call setup, transport of voice/data over the network and management of call. Signaling Protocols in IP Telephony deals with the call setup and management, voice/data transport over the network is not the part of signaling protocols for IP Telephony. In this report I will mainly emphasize on how call is managed and setup in both the protocols.

The rest of the report is organised as follows: Section 2, gives a overview of both the protocols - H.323 and SIP, Section 3 describes comparison between two signaling protocols and contains concluding remarks.

2 Overview Of H.323 and SIP

Currently two families of protocols are prevalent for VoIP – H.323 family from the ITU consisting of H.245, H.225 & Q.931 and the IETF family of protocols consisting of RTP, SIP, MGCP, SDP and SAP.

International Telecommunications Union(ITU) developed H.323 for multimedia communications over Local area Network(LAN) that do not provide guaranteed Quality Of Service(QoS). H.323 represents a traditional circuit switched approach to IP Telephony, based on ISDN signaling protocol Q.931.

Session Initiation Protocol(SIP) is developed by Multiparty Multimedia Session Control(MMUSIC) group of Internet Engineering Task Force(IETF). It is a lightweight protocol based on Hyper Text Markup Language(HTML).

2.1 H.323 Overview

2.1.1 Network Architecture

H.323 contains four components : **Terminals**, **Gateways**, **Gatekeepers** and **Multipoint Control Unit(MCU)**.

Terminals

These are client endpoints on the network which are able to initiate and receive calls. These can be either software running in computer or a dedicated device. A Terminal should support H.245, Q.931, Registration Admission Status (RAS) and Real Time Transport Protocol(RTP).

Gateways

Gateways provide an interface between H.323 based network and Public Switched Telephone Network(PSTN), SIP or ISDN. They are used for translation between transmission format.

Gatekeepers

Gatekeeper is the main component and acts as a manager. H.323 network consists of zones which contains terminals, gateways and MCU. Gatekeeper is responsible for controlling a zone. Calls within a zone are controlled by Gatekeeper and for managing calls which are between different zones may requires several Gatekeepers.

Four main functions performed by Gatekeeper are :

- **Address Translation :** Gatekeepers translate alias addresses or phone numbers to transport addresses. It uses translation table for this purpose.

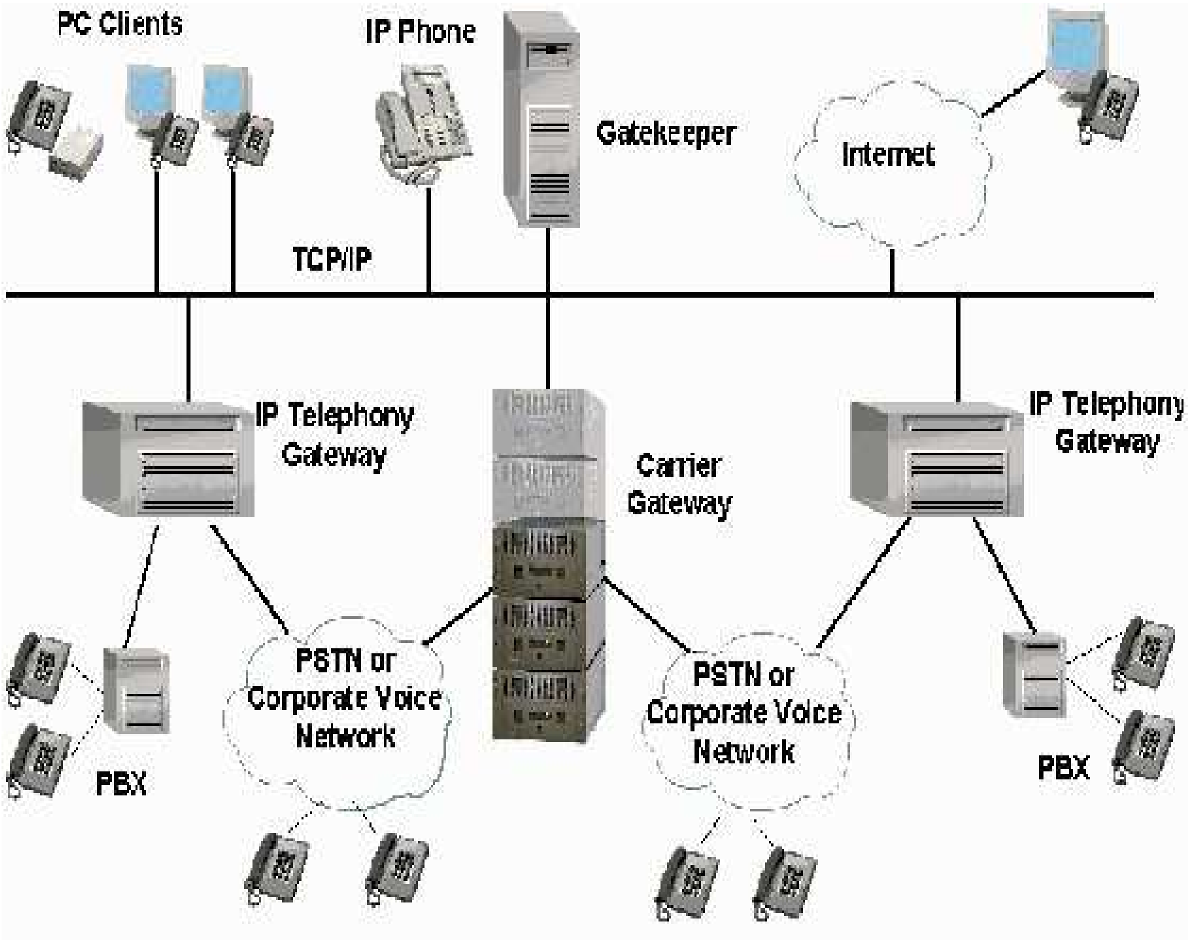

Figure 1: H.323 network architecture

- **Admission Control :** It allows the user to initiate or terminate a call based on authorisation.

- **Bandwidth control :** Certain amount of bandwidth is reserved for for H.323 traffic and is distributed among the connections. When the number of users exceed limit the Gatekeeper rejects the call due to bandwidth limitations.

- **Zone Management :** Gatekeeper maintains a registration table. Terminals within a zone register to the Gatekeeper which adds the corresponding address to the registration table.

Optional features that can be provided by Gatekeeper are : **Call authorisation**, **Call Control Signaling**, **Bandwidth Management** and **Call Management**.

Multipoint Control Units(MCU)

MCU provides three or more terminals to take part in conference. It consists of Multipoint Controller(MC) and zero or more Multipoint Processors (MP). The Multipoint Controller provides call setup, negotiation of capabilities at both end and Multipoint Processor performs necessary processing on media streams, such as audio mixing.

2.1.2 Protocol Architecture

The series H.32x contains following protocols:

- H.320 for narrow band switched ISDN.

- H.321 for narrow band switched digital ISDN & ATM.

- H.322 for guaranteed bandwidth packet switched network.

- H.323 for non-guaranteed bandwidth packet switched network.

- H.324 for Analog Phone System(POTS).

H.323 uses both reliable and unreliable communication. Data and control application packets uses reliable Transmission Control Protocol(TCP) while audio/video and registration uses User Datagram Protocol(UDP). Structure of H.323 terminal is shown in Fig 2.

In the lowest layer, H.225 defines stream synchronization, media and control message packetisation, and control message formats. It uses RTP to perform packetisation of media streams. It handles the call control to initiate and terminate calls. In next layer, H.245 defines negotiation of channel usage and capabilities exchange. It is used for opening and closing of channels for audio, video and data requests. Q.931 defines call signaling and call setup.

Control functions are provided by following signaling channels:

RAS Channel: Registration Admission Status(RAS) protocols is defined in H.225.0. This channel is used in communication with Gatekeeper. Through this channel endpoint register with the Gatekeeper and asks for permission to place a call. If permission is granted then the Gatekeeper returns the transport address for the call signaling channel of the endpoint.

Call signaling Channel: This channel is used to establish calls using Q.931 protocol.

H.245 Control Channel: This channel is a reliable channel which carries control messages regarding capability exchange, opening and closing of logical channels.

Logical Channels: This channel carries audio, video and other media information. It uses a pair of unidirectional channels for carrying each media type.

2.1.3 Call Setup and Management

Communications using H.323 is a five phase process: Call Setup, Capability Exchange, Audio/Video Communication, Call Services and Call Termination. If endpoint wishes to communicate with other which involves Gatekeeper in between then it first has to register with Gatekeeper.

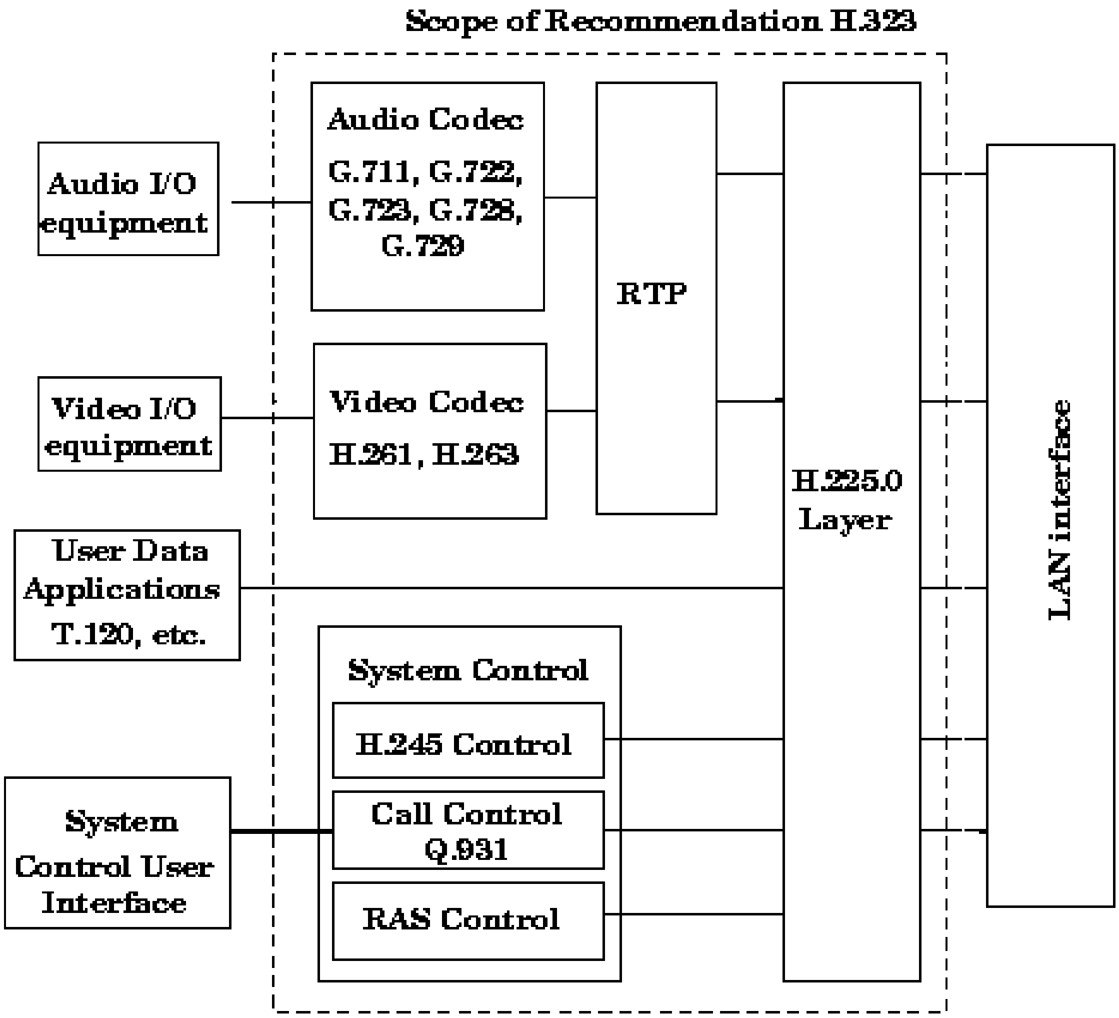

Figure 2: Structure of H.323 terminal

Gatekeeper registration procedure uses RAS signaling channel. The procedure consists of ARQ request to the Gatekeeper and in response Gatekeeper either sends an ACF confirmation or ARJ reject message.

After registration caller sends a *setup* message to the called endpoint. This call setup procedure uses H.225 protocol.When callee receives the request it responds with *call proceeding* message and when the user is informed then it sends *alerting* message to the caller. If call goes through a gateway then gateway sends *alerting* message when ring tone is heard. A *connect* message will be sent when the user answers. After accepting the call by callee media channels are setup using H.245 protocol. Then feature negotiation is performed between the two endpoints. This is done by sending *terminal-capability set* messages. Due to large number of different terminal types codecs and implementations available the two endpoints should first negotiates on capabilities and features which ensures that multimedia signals that are transmitted are processed appropriately on receiving side. Thus capability exchange and feature negotiation is a fundamental function in IP Telephony. To conserve resources and reduce call setup time H.245 messages can also be transported within Q.931 channel which is known as *Tunneling*
Call setup is shown in Fig 3

If during the conference another party wishes to enter conference than first feature negotiation and capability exchange is to be done before that party joins conference. Media parameters can be changed in H.323 during the conference by changing the bandwidth.

Call control services provided by H.323 are as follows:

Call Hold : It is defined as one party temporarily disconnecting voice communication without ending the call and reestablish the communication after some time. In the meantime, when call is on hold, some music can be played known as Music on Hold(MOH), informing party on hold about the liveliness of the call. H.323 provides two type of call hold service:

Near-End Call Hold: Hold is invoked at the holding endpoint as a local procedure.

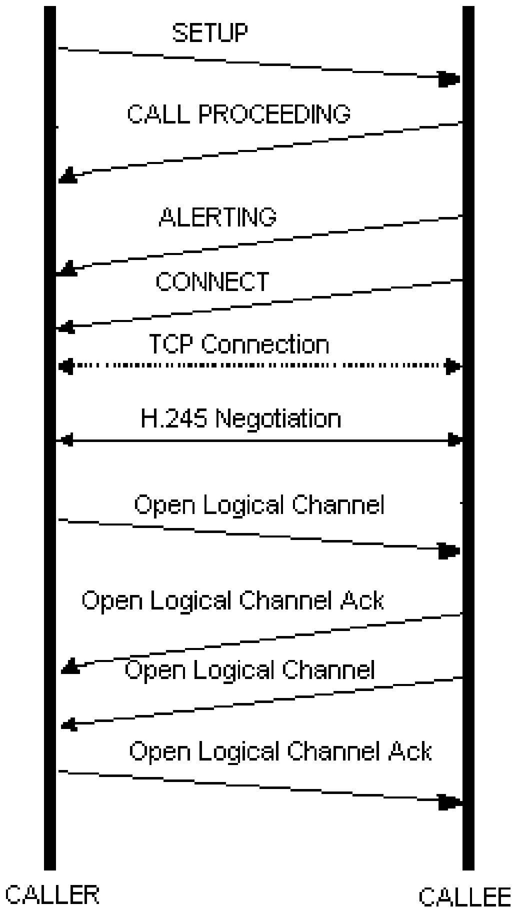

Figure 3: Setup of H.323 call

Remote-End Call Hold: Holding endpoint sends a hold request to remote endpoint requiring held endpoint to provide Music on Hold (MOH) to held user.

Call Transfer : It enables a user to transfer an established call to third party

Call Waiting : By this feature called party can know if any new party is trying to connect with it while it is busy with another party.

Call Forwarding : It permits a called party to forward the call to other address.

Finally, the call is terminated by the caller or callee by sending *RELEASE COMPLETE* message.

2.2 SIP Overview

2.2.1 Network Architecture

SIP consists of two components:

User Agent:

It is an application in the end system which is acting on behalf of a user. It contains User Agent Client(UAC) and User Agent Server(UAS).

User Agent Client: also known as calling user agent is used to initiate a SIP request.

User Agent Server: also known as called user agent, contacts user receiving SIP request and sends responses on behalf of user.

Network Servers:

SIP structure contains three type of servers:

Registration Server: also called Registrar, is a server that accepts registration requests. It maintains the current location of users registered with it.

Proxy Server: It receives request and forward it to the next - hop server or to the called party whichever comes next in the route. It can act both as a client and a server.A Proxy Server can be stateless or stateful.

Stateless Proxy Server forgets SIP request after forwarding it.

Stateful Proxy Server remembers the request after it has been forwarded. It can be a forking proxy server which sends more than one request in response to receiving the request.

redirect Servers: A redirect Server on receiving SIP requests, maps the address into zero or more new addresses and return the result to client. It does not forwards a request.

2.2.2 Protocol Architecture

SIP is independent from underlying network topology and it can be used with several protocols like TCP and UDP. Since SIP does not require reliable transport protocol User Datagram Protocol(UDP) can be used to implement simple client. SIP retransmits request after every 0.5 seconds to achieve reliable communication until response is returned. TCP allows easier passage through firewall. UDP allows to perform parallel search without maintenance of TCP connection state for each outstanding request. With TCP, SIP can be used with one or more connections to contact a user or to modify parameters of an existing conference.

SIP uses *Session Description Protocol(SDP)* for describing capabilities and media type that are supported by the terminals. SDP messages list the features that must be implemented in the endpoints. These messages are used in capability exchanges and feature negotiation. SIP also uses *Session Announcement Protocol(SAP)* to advertise the session in large public conferences and broadcast streams.

2.2.3 Messages used in SIP

INVITE : for initiating a call.
ACK : message sent after caller gets a response for INVITE.
REGISTER : message sent by user to Registrar giving information about its location.
BYE : for terminating a connection between endpoints.
CANCEL : terminating the search for a user.
OPTIONS : message sent for obtaining information about the capabilities of a call.

Call Setup and Management

Session Initiation Protocol hosts are identified by SIP URL which is of the form *sip:username@host*. Caller can send a request either directly to IP address of the callee or it has to sent the request to locally configured SIP proxy.

For initiating a call, caller sends a *INVITE* message. The invite request contains a description of the session to be used written in SDP format. This description is used by called party to know about what media types and formats caller is willing to use. If callee accepts the description it responds to the call by returning similar description. The communication can be made through proxy server or redirect server.

Operation in Proxy Mode

- Caller send *INVITE* message, containing destination address in the TO header field, to Proxy Server.
- Proxy Server contact Location Service.
- Location Service returns zero or more address.
- Proxy Server make the requests to the address(es) returned by Location Service.
- On receiving request the User Agent Server returns a success indication to Proxy Server.

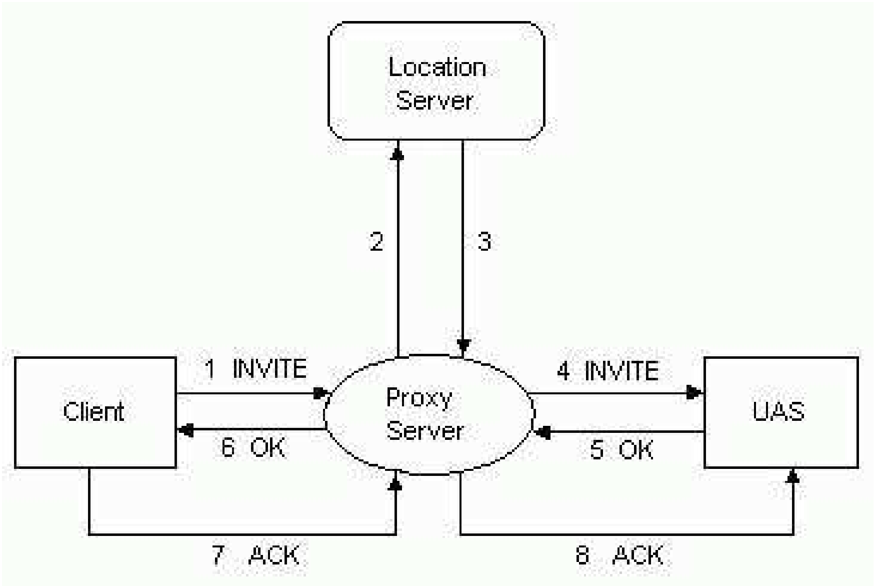

Figure 4: Operation in Proxy Mode

- Proxy Server, in turn returns it to caller.
- Caller sends *ACK* message which can also be sent directly to callee.

After which the call is established. This operation is shown in Fig 4.

Operation in Redirect Mode

- Caller send *INVITE* message, containing destination address in the TO header field, to Redirect Server.
- Redirect Server contact Location Service.
- Location Service returns zero or more address.
- Redirection Server instead of contacting itself returns the address(es), returned by Location Service, to caller.
- Caller acknowledge by sending *ACK* message to Redirection Server.

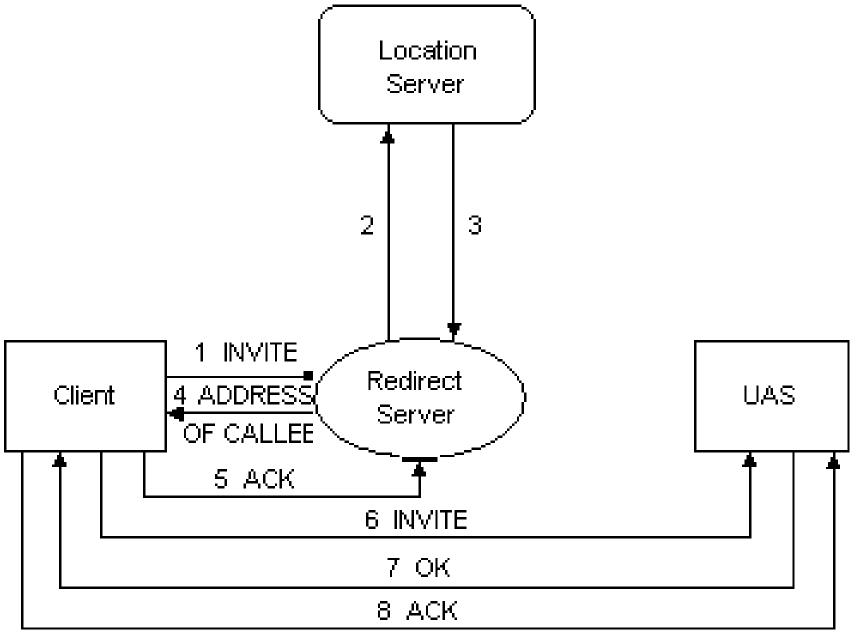

Figure 5: Operation in Redirect Mode

- Caller issues new request with same Call ID to addresses returned by Redirection Server.
- Callee sends success response after accepting session description sent along with the message.
- Caller sends ACK message which can be sent directly to callee.

The operation is shown in Fig 5.

SIP provides mobility of callee. The locations to which callee moves can be registered dynamically with Location Server. When it is queried about the location of the callee, it returns a list of possible locations.

Responses passed in SIP message contains three digit codes which are classified according to first digit.

- 1xx - Indicate progress of call and are always followed by other responses indicating final result.
- 2xx - Indicate success.
- 3xx - Indicate redirection.
- 4xx - Indicate client error.
- 5xx - Indicate server error.
- 6xx - Indicate global failure.

SIP message contains following header fields:

From :- Indicate originator of the call.
To :- Address of the callee.
media capabilities :- Gives description of the session.
Call ID :- Unique identifier created by caller and is same for all the messages in the call.
CSeq :- Sequence number used in messages.
Via :- It defines the path traversed by the request so far.

Call Services provided by SIP are :

Call Hold : To achieve *Near-end Call Hold* client just continually receives media stream from server but does not generate any response.
For *Remote-end Call Hold* holding side send INVITE message to other side indicating a NULL set of receiving capability for any kind of media.

Call Forwarding : SIP messages contains LOCATION header fields which contains destination to which the call is to be forwarded.

Call Waiting : When caller A calls B and if B is already busy with another client then it will return status code "*181 Queued*" in its response. Afterward when it becomes free it will return appropriate status response. This allows caller A to remain connected even when callee B is already busy.

Third Party Control : It provides a party to set up call between two other parties without participating in it.

Finally, a SIP call can be terminated by caller or callee by sending *BYE* request.

3 Comparison Of H.323 & SIP and Discussions

Complexity

H.323 uses a binary representation for message, based on Abstract Syntax Notation(ASN.1) and Packet encoding Rules(PER), which requires special code-generators to parse.Whereas SIP encodes message as text similar to Hyper Text Transfer Protocol(HTTP). This leads to simple parsing and generation. Moreover, SIP being similar to HTTP can also reuse code and also existing HTTP parsers can be easily modified for SIP usage.

Moreover, H.323 uses several subprotocols which includes H.225 for call signaling, H.245 for call control, H.332 for large conferences, H.235 for security and encryption, H.246 for interoperability. H.323 contains hundreds of elements while SIP defines 37 headers. This all makes H.323 more complex than SIP.

Transport Protocol

SIP can be used with any protocol. It allows support for UDP which results in lower call setup time whereas H.323 requires reliable protocol like TCP.

Loop Detection

SIP provides loop detection by using *Via* header field in the message. Proxy Server, before forwarding a request checks *Via* field and if it find its name then a loop must have occurred. H.323 does not provides any provision to prevent loops but it defines a *PathValue* field which indicates maximum number of Gatekeeper that message should traverse before being discarded.

Fault Tolerance

H.323 provides better fault tolerance than SIP by redundant Gatekeepers and endpoints. During registration process a endpoint may give alternate address where it can be found or Gatekeeper may provide endpoint with alternate Gatekeeper which can be contacted in the event of failure.

Server Processing

SIP transaction through servers and gateways can be either stateful or stateless. Stateful model provides the sufficient information to forward response correctly whereas stateless model simplifies memory management. H.323 v1 and v2 requires stateful server processing where TCP is the transport protocol. H.323 v3 supports stateless processing model.

Ease Of Customization

H.323 requires more interactions between the its subprotocols. Whereas SIP uses its relatively simpler header fields for customization and also text based encoding makes it easier to customize.

Interoperability Among Versions

H.323 provides full backward compatibility. Whereas in SIP a newer version may discard some old features that are not expected to be implemented any more. Full backward compatibility provided by H.323 has advantages that if either of Client/Server side is upgraded then the other side need not be upgraded but backward compatibility always increments the size of encoding, making it more complex.

Scalability

H.323, because of complex architecture, no provision for loop detection and requires stateful Gatekeepers, is less scalable as compared to SIP.

Future demands would require IP Telephony to be integrated in many applications. New services based on speech transmission would require lightweight signaling protocols like SIP that can be easily be integrated into the software. Hence because of simplicity and extensibility SIP is a better protocol than H.323.

References

[1] M. Handley, H. Schulzrinne, E. Scholler, J. Rosenberg, "*SIP: Session Initiation Protocol*", RFC2543, IETF, March 1999.

[2] H. Schulzrinne, J. Rosenberg, " *Signaling for Internet Telephony*", Technical Report CUCS–005-98, Columbia University, New York, Feb. 1998.

[3] H. Schulzrinne, J. Rosenberg, "*A Comparison of SIP and H.323 for Internet Telephony*", Network and Operating System Support for Digital Audio and Video 154(NOSSDAV), Cambridge, England, July 1998.

[4] Nicklas Beijar, "*Signaling Protocols for Internet Telephony — Architectures based on H.323 and SIP*", Helsinki University of Technology.

[5] Ismail Dalgic, Hanlin Fang, "*Comparison of SIP and H.323 for Internet Telephony Signaling*".

[6] Lasse Huovinen, Shuanghong Niu, "*IP Telephony*", Department of Computer Science and Engg., Helsinki University of Technology.
http://www.tml.hut.fi/Opinnot/Tik-110.551/1999/papers/04IPTelephony/voip.html