A Novel Approach of Next Generation Signaling Bandwidth Optimization

Munir Sayyad[#], SanjayNalbalwar [$]
[#$]Dr.B.A. Technological University, Lonere
[1]munirsayyad@gmail.com,[2]nalbalwar_sanjay@yahoo.co.in,

***Abstract*- Session Initiation Protocol is a signaling protocol and it is an integral part of IP Multimedia Subsystem to support functions such as establishment, control and termination of the sessions. ASCII signaling messages sent using the SIP protocol pass through intermediate SIP nodes in such a way that the message size grows as a result of additional data being added to the message. In this paper a flow model analysis is presented to describe the size of the expected signaling flows and we have also described a method for compression on retransmitted messages.**

***Keywords*- Session Initiation Protocol (SIP), IP Multimedia Subsystem (IMS), Voice over Internet Protocol (VoIP), Internet Engineering Task Force (IETF), Internet Protocol (IP), Single Space (SP), Third Generation Partnership Project (3GPP).**

I. Introduction

SIP is designed to be a part of the IETF multimedia data and control architecture. SIP is used in conjunction with several other IETF protocols, such as the Session Description Protocol (SDP), the Real-Time Streaming Protocol (RTSP), and the Session Announcement Protocol (SAP). SIP is designed to initiate interactive sessions on an IP network. Programs that provide real-time communication between users can use SIP to set up, modify, and terminate connection between two or more computers, allowing them to interact and exchange data. SIP supports additional functions, such as call waiting, call transfer, and conference calling, by sending the necessary signals to enable and disable this functions.SIP is an ASCII or text based protocol. SIP is a request-response protocol, which means that it makes a request of a server, and awaits a response. Once a session is established, other protocols handle such tasks as negotiating the type of media to beexchanged, and transporting it between the endpoints [3].

II. SIP BREIFING

A. SIP Functions and Features

SIP is developed to support five specific elements for setting up and tearing down communication sessions. These facets of the protocol are:

1. User location- the endpoint of a session can be identified and found, so a session can be established.
2. User availability- the participant who is called has the opportunity and ability to indicate whether he or she wishes to participate in the communication.
3. User capabilities- the media that will be used in the communicationis established, and the parameters of that media are agreed.
4. Session setup- the parameters of the session are negotiated andestablished.
5. Session management- the parameters of the session are modified,data is transferred, services are supported, and the session isterminated.

B. SIP Architecture

There are two fundamental components that are used by the Session InitiationProtocol:

1. User agents- this are endpoints of a call (i.e., each of the participantsin a call)
2. SIP servers- It is a network element that receives requests in order to service them and sends back responses to those requests.

1) User Agents

User agents are the device that is being used to make a call, and the targetdevice that is being called. This makes the two endpoints of thecommunication session. There are two components to a user agent named a clientand a server. When a user agent makes a request (like initiating a session),it is the User Agent Client (UAC), and the user agent responding to therequest is the User Agent Server (UAS).

2) SIP Server

The SIP server is used to resolve usernames to IP addresses, so that requestssent from one user agent to another can be directed properly.A user agent registers with the SIP server, providing it with their username and current IP address and hence establishing their latest location on the network. A request is made by the user agent to theSIP server to invite another user into a session. The SIP server then identifieswhether the target agent is currently online, and if so, it compares the username totheir IP address to determine their location. If the user isn't part of that domain, and hence uses a different SIP server, it will also pass on requests toother servers.

In performing these various tasks of serving client requests, the SIP server have elements listed as:

a. Registrar server
b. Proxy server
c.Redirect server

a) *Registrar Server*

Registrar servers are used to register the location of a user agent who haslogged onto the network. It obtains the IP address of the user and associates itwith their username on the system.This creates a directory of all those whoare currently logged onto the network, and where they are located. When any user wishes to establish a session with one of these users, the Registrarserver's information is referred to, thereby identifying the IP addresses of thoseinvolved in the session. The figure 1 below describes the registration process.

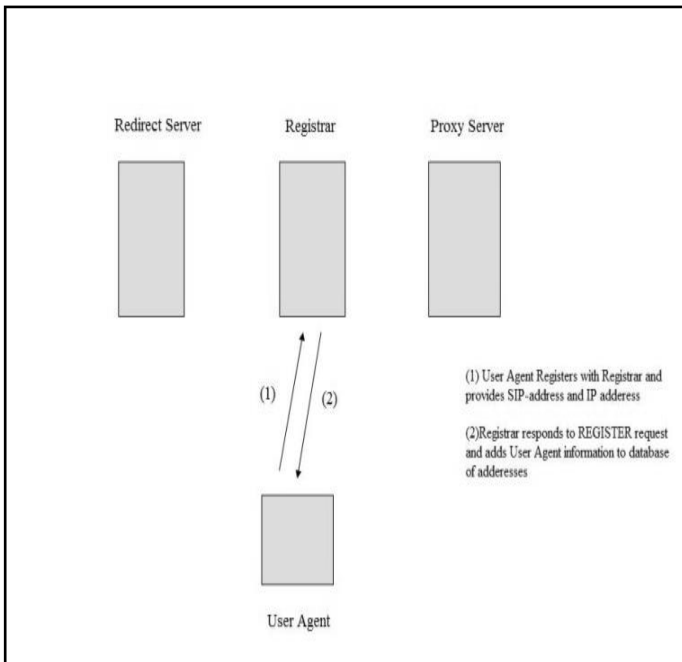

Figure 1: registration process.

b) *Proxy Server*

Proxy servers are the elements that route SIP requests to the user agent servers and SIP responses to user agent client. While functioningas a proxy server, the SIP server can provide functions such as network accesscontrol, security, authentication, and authorization. The following figure 2 describes the working of proxy server.

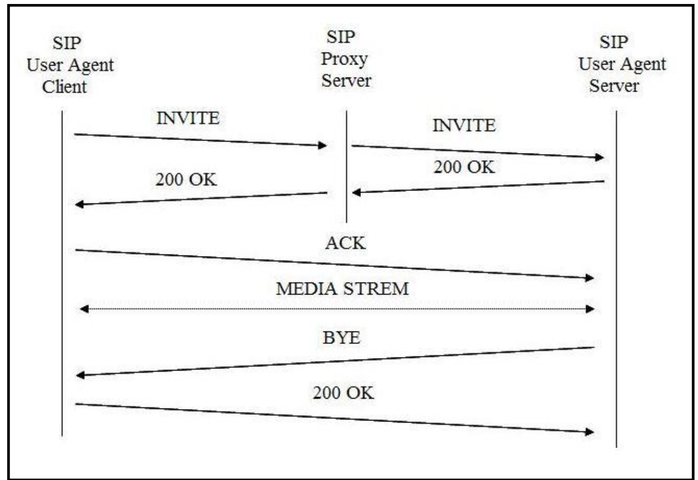

Figure 2: Proxy server example.

c) *Redirect Server*

The Redirect servers are used to redirect clients to the user agent theyare attempting to contact. Redirection allows servers to push routing information for a request back in a response to the client, thereby taking themselves out of the loop of further messaging for this transaction while still aiding in locating the target of the request. When the originator of the request receives the redirection, it will sent a new request. This is differentfrom a Proxy server, which forwards the request on user's behalf, as theRedirect server essentially tells you to contact them yourself.

The figure 3 shown below describes the redirect server.

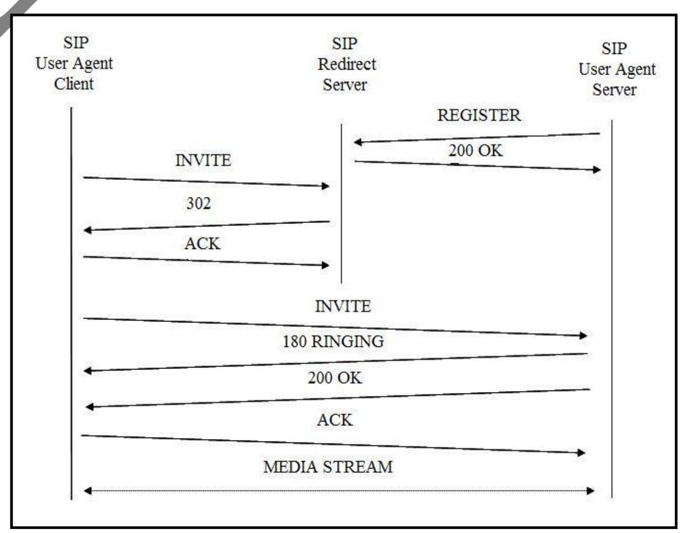

Figure 3: Request redirection.

C) SIP Request

As SIP is an ASCII protocol. A SIP message sent from a server to a client, for the purpose of invoking a particular operation. SIP request have a distinguishing feature of having a Request-Line for a start-line. A Request-Line contains a method name, a Request-URL, and the protocol version separated by a SP character. The Request-Line ends with the CRLF. The signaling commands that might be used are:

1. REGISTER- Used for registering contact information.
2. INVITE-Used to invite user(s) to a session.
3. ACK- Acknowledgment of an INVITE request.
4. BYE-For terminating sessions.
5. OPTIONS-Querying servers about their capabilities.
6. CANCEL-Cancel a pending request.
7. SUBSCRIBE- This is used to acquire updated informationon whether a User agent is online, busy, offline, and so on.
8. NOTIFY-Used to send updated information on a User agent's currentstatus.

C. SIP Responses

A SIP message sent from a server to a client, for indicating the status of a request sent from the client to the server. SIP Responses are distinguished from requests by having a Status-Line. A Status-Line consists of the protocol version followed by a numeric Status-Code and its associated textual phrase, with each element separated by a single SP character. The Status-Code is a 3-digit integer result code that indicates the outcome of an attempt to satisfy a request.

The various categoriesand their response code prefixes are as follows:

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

D. SIP Session setup

SIP Session setup is shown in figure 4. The process starts with SIP INVITE message, which is used from the calling party to the called party. The message invites the called party to participate in a session i.e. a call. The called party answers the call, which generates an OK response back to the caller.

The calling client acknowledges that the called party has answered by issuing an ACK message. At this point, media are exchanged. Finally one of the party hangs up, which causes a BYE message to be sent. The party receiving the BYE message sends OK to confirm receipt of the message. At this point the session is over.

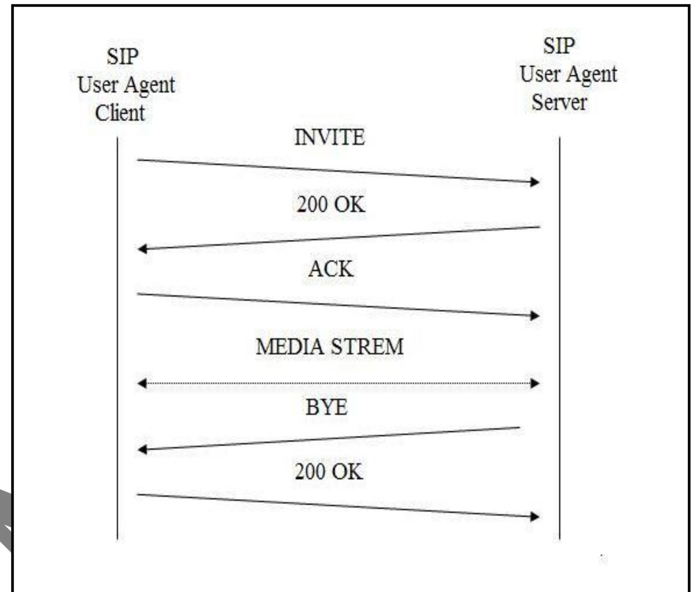

Figure 4: SIP Session setup.

III. MATHEMATICAL MODELING

Lost Message Model

The SIP protocol is transport protocol independent. Since the only mandatory transport protocol for SIP is UDP, it needs to incorporate its own end-to-end reliability mechanism. In particular the SIP extension is known as Reliability of Provisional Responses" introduces an additional reliability mechanism for the provisional response in a SIP call flow, which is used by 3GPP. This section discusses a flow model to take the reliability mechanism into account. The modeling approach of this behavior is based on the model for repeated attempts in.

A message flow M between two nodes has to be transmitted over a link thatis assumed to have an error probability PE(M). This link has to accommodatethe original message flow M. Consequently, a flow of (M.PE(M)) will be loston the link due to the message error and has to be retransmitted on this samelink. This new flow (M.PE(M)) is subjected to loss once again with probabilityPE(M). So the lost flow in this instance is then (M.PE(M).PE(M)). If amessage is resent n times this yields Equation (1).

$$F\,1 = M + M.PE(M) + M.PE(M)^2 + \ldots + M.PE(M)^n \qquad -(1)$$

Where F1 is the total flow on the link. This well-known geometric series can besummed as shown in Equation (2).

$$F1 = \frac{M(1-PE(M)\hat{}(n+1)}{1-PE(M)} \ldots (2)$$

For an infinite number of retransmissions a simplification of Equation (2) ispossible as:

$$F1 = \frac{M}{1-PE(M)} \ldots\ldots (3)$$

The SIP protocol specifies that the messages are resent with a maximum numberof reattempts n = 7. Under certain conditionsthe number of reattemptscan be reduced to n=4. A message that is lost n times will cause a termination of the connection. Since the error is very small the formula for the limiting case is used for simplicity.

If every time there is need of retransmission and message is transmitted as it is then it will increase flow on the link and it will require more bandwidth. This will increase transmission overheads. Hence we are proposing compression of the messages every time there is a retransmission.Now consider the equation 1:

$$F1 = M + M.PE(M) + M.PE(M)2 + \ldots + M.PE(M).$$

Now suppose the message is compressed by the factor 2 i.e. compressed to 50% every time when it is retransmitted then the equation becomes

$$F2 = \frac{M}{1} + \frac{M.PE(M)}{2} + \frac{M.PE(M)\hat{}2}{4} + \cdots + \frac{M.PE(M)\hat{}n}{2\hat{}n}$$

By using geometric series the flow can be calculated as:

$$F2 = \frac{2M}{2-PE(M)} \quad \ldots(4)$$

Note: Any compression technique as per the application can be used and accordingly the retransmitted message compression factor can be chosen [4].

IV. RESULT

This section discusses results that have been found by applying the above model.The presented results are preliminary with a focus on the effective utilization of bandwidth during retransmission of messages.Considering the following example, where a SIP retransmitted messagehaving an extremely high bit error rate of $PE=10^{-2}$ [1] and message size (M) of 8 bytes.

Equating the equation (3) F1 comes out to be 8.8889 and the equation (4) F2 comes out to be 8.4211.

Hence we have reduced the flow on the link by 5% (≈) during retransmission of message.Following graph illustrates the simulation results. The X-axis shows Message flow on the link and the Y-axis represents Message length. The message flow for compressed retransmitted messages is shown by dotted lines while the uncompressed retransmitted messages are shown by the continuous line.

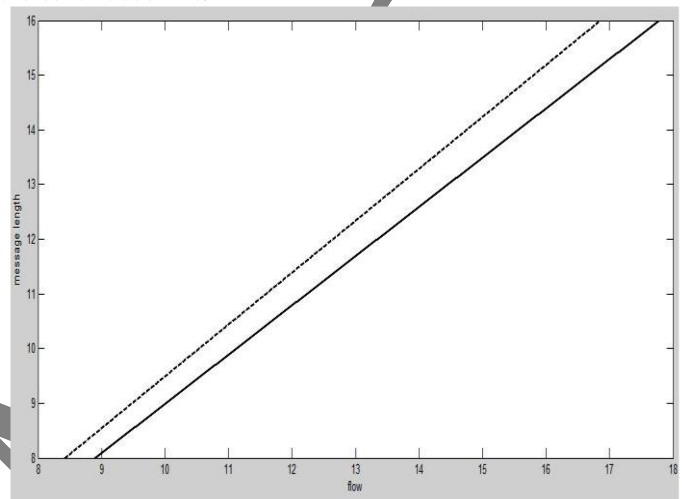

Figure 5: Message length versus message flow.

V. CONCLUSION

Inthis paper we have presented the role of SIP, used for establishment, control and termination of session where we have analyzed a method for calculating the message flow on the SIP connection andthe effective utilization of bandwidth during retransmission of message is described. This paper analyzesthe flow model to describe the size of the expected signaling flows and also describes a method for compression on retransmitted messages.

REFERENCES

[1] Alexander A. Kist1 and Richard J. Harris1- A Simple Model for Calculating SIP Signaling Flows in 3GPP IP Multimedia Subsystems, RMIT University, BOX 2476V, Victoria 3001, Australia.
[2] Carrier Grade Voice Over IP- Daniel Colling- MCGRAW-HILL PROFESSIONAL TELECOM. ISBN 0-07-136326-2.
[3] Henning G. Schulzrinne and Jonathan D. Rosenberg- The Session Initiation Protocol : Providing Advanced Telephony Services Across the Internet.
[4] John Sweeney, Valerie Kennealy, Dirk Pesch, Craig Purcell- Efficient SIP based Presence and IM Services with SIP message compression in IST OPIUM.
[5] Putting VoIP to work- Bill Douskalis-Prentice Hall PTR, ISBN 0-13-040959-6