SIP Signaling and QoS for VoIP over IPv6 DVB-RCS Satellite Networks

M. Ali, L. Liang, Z. Sun (*Member IEEE)* and H. Cruickshank (*Member IEEE)*

Centre for Communication Systems Research, University of Surrey, Guildford, Surrey, GU2 7XH, UK
{m.ali, l.liang, z.sun, h.cruickshank@surrey.ac.uk}

Abstract—**With the rapid development of the Internet, new technologies and applications are emerging. One of the important applications is voice over IP. Satellites are playing an important role to provide VoIP services with their global coverage and on-board processing ability over IP networks. Satellite network environment, generally characterized by large delay and erroneous link, is considered to be unfriendly to VoIP. The performance of VoIP is adversely influenced by these demerits. The performance metrics of VoIP are signaling, bandwidth, delay, jitter and packet loss. Signaling plays a key role in call establishment and rest of the parameters signifies the quality of service (QoS). In this paper, the performance related issues of SIP-based VoIP over current, IPv4, and next generation, IPv6 satellites is studied. A comparative analysis is performed for different voice codecs. The experimentation is carried out on the satellite network testbed at Centre for Communication Systems Research (CCSR) at University of Surrey. The results show that delay, jitter and packet loss are quite comparable for both current and next generation satellites. SIP signaling performs poorly in IPv6 as compared to IPv4. IPv6 can be adapted for VoIP over next generation satellites, but with some modifications for SIP signaling.**

Index Terms—**QoS, SIP, IPv4, IPv6, Satellite network**

I. INTRODUCTION

Voice over IP being a real time application, is intolerant to long call setup time, delay, jitter and packet loss. Originally, H.323 was the key signaling protocol that allowed interoperability of VoIP products and introduced an alternative to the initial proprietary solutions [1]. Meanwhile, other VoIP signaling protocols, Media Gateway Control Protocol (MGCP), H.248 or *Megaco* and Session Initiation Protocol (SIP) were developed. SIP [2] is experiencing rapid growth after its proposal. Java has implemented the SIP protocol stack and Microsoft has included a SIP user agent in its operating systems, Windows XP and later versions. SIP is the main signaling protocol in 3G, IP Multimedia Subsystem (IMS). It is a key tool in building next-generation converged VoIP networks and integration of the PSTN with the VoIP networks.

The call setup time is an important factor in voice calls in

telecommunication networks. [3][4] provides a comparison of call setup in SIP on top of UDP, TCP, and stream control transmission protocol (SCTP) using ns-2. SIP call setup time and RTP one-way delay have been evaluated using IPv4, IPv6 and different IPv4-IPv6 transitioning mechanisms [5]. This study is continuation of our previous work [6], in which, we are focusing on these metrics along with other QoS parameters of VoIP over IPv6 satellites.

This paper is divided into the following sections: Section II is a description of IPv6 and its deployment issues in satellites; Section III is an overview of the SIP protocol; the audio compression techniques and quality of service (QoS) parameters have been briefly reviewed in Section IV; Section V describes the satellite network testbed used to perform the VoIP experiments; the outcome of these experiments is discussed in Section VI; and finally, Section VII concludes the paper based on the evaluation results.

II. IPV6 AND SATELLITES

The next generation network protocol, IPv6, is emerging to overcome the shortcomings of IPv4, the most important being the shortage of address space [6]. Several Internet Service Providers (ISPs) have offered IPv6 services in North America, Europe and Japan. Because of its huge address space, it will be adapted by mobile and vehicular networks. Although, IPv6 is proliferating and integrating in the existing networks but it has to coexist with IPv4 using dual-stack, protocol translation and tunneling approaches for a long transitional period [7].

Trial deployments have demonstrated that IPv6 can be used over satellites [8]. Current Digital Video Broadcast (DVB) standards do not document much about satellite for IPv6. The default Multi-Protocol Encapsulation (MPE) does not have a source MAC address, and a payload Type field, presenting a hindrance in deployment of IPv6. Such encapsulations that do not have IPv6 support could use tunnel mode (e.g. IPv6 over IPv4), or a link layer encapsulation (LLC), but at the cost of increased overhead. IPv6-friendly encapsulation schemes, like, Unidirectional Lightweight Encapsulation (ULE) encapsulates the packets in the MPEG2 Transport Stream. The introduction of Generic Streams by DVB-S2 makes necessary a new encapsulation, the Generic Stream Encapsulation (GSE), that relies in some fundamental design choices of ULE. The next generation satellite system has to be designed to be as far as
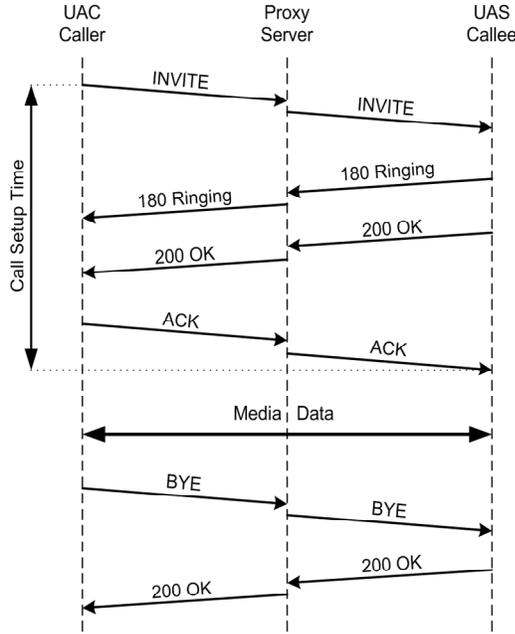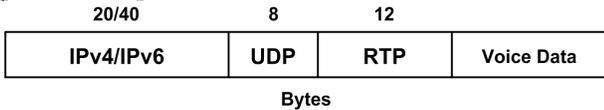
Fig. 1. SIP message flow

| 20/40 | 8 | 12 | |
| IPv4/IPv6 | UDP | RTP | Voice Data |

Bytes

Fig. 2. IPv4/IPv6 voice packet

TABLE I
VOICE CODECS AND BANDWIDTH CONSUMPTION

Voice codec	Bit rate (kbps)	Voice payload (bytes)	Bandwidth (kbps)		Difference IPv4 & IPv6	
			IPv4	IPv6	Calc	Meas
PCMA	64	240	74.667	80.000	5.333	5.440
G.721	32	80	48.000	56.000	8.000	8.160
GSM	13.2	33	29.200	37.200	8.000	8.160
LPC	2.5	7	16.786	23.929	7.143	7.200

main six methods are INVITE, REGISTER, BYE, ACK, CANCEL and OPTIONS. The response messages are 100 Trying, 180 Ringing, 200 OK and many others. The flow of messages during a call set up and tear down is as shown in Fig. 1. The sequence of messages to setup a call is INVITE - 200 OK - ACK. The voice data is carried by Real Time Protocol (RTP) [14] and it is exchanged between the caller and callee directly. In the same way, BYE - 200 OK are exchanged to tear down a session [2].

IV. CHARACTERISTICS OF VOICE OVER IP

A. Speech Coding

Speech coding is used to compress voice signal for transmission over long distance. It involves the process of transforming the analog signal (human voice) into digital signal, sending the digital data to the far end and regenerates the voice at the far end. Various speech codecs are being used in PSTN and Internet. A softphone, Ekiga [10], formerly known as Gnomemeeting, implements the codecs listed in Table I. We used it to generate voice streams with different encoding schemes.

Voice over IP uses Real Time Protocol (RTP) [14] to carry voice packets. RTP uses sequence numbers and time stamps to identify out of order packets. RTP is encapsulated in the unreliable User Datagram Protocol (UDP). So there is no guarantee of arrival of voice packets at the destination. If reliability has to be incorporated, it can be implemented in the application generating the voice packets. The voice payload with various headers is shown in Fig. 2. There is an extra overhead of 40 and 60 bytes with each IPv4 and IPv6 voice packet, respectively.

B. Performance metrics

Following are the main performance metrics of voice over IP and SIP:

Bandwidth required depends on the voice codec and its algorithmic complexity. The IP bandwidth consumed by a voice call can be computed by the following formulae [11].

$$Total\ packet\ size = IP/UDP/RTP\ header + Voice\ payload\ size$$

$$PPS = Codec\ bit\ rate / Voice\ payload\ size \qquad (1)$$

$$Bandwidth = Total\ packet\ size \times PPS$$

where PPS is the number of packets needed per second to deliver the codec rate. The IP/UDP/RTP header is fixed and its
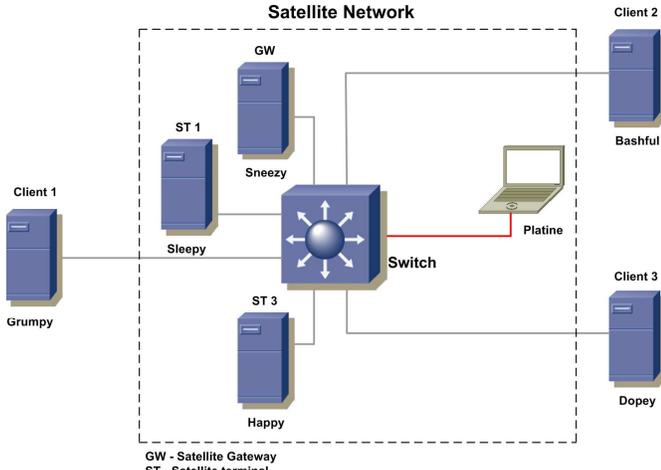
compatible with GSE/ULE. However, as vast majority of existing satellite systems use MPE, the transition may take a long time.

There is currently little (or no) support for IPv6 in commercial two-way DVB-RCS terminals. Some projects have carried out trials demonstrating issues relating to IPv6 provision in DVB-RCS (e.g., the EC SATSIX Project [16]). An upgrade to IPv6 needs some amendment in the standards, especially to ensure that control and management planes are IPv6-enabled.

III. SESSION INITIATION PROTOCOL (SIP)

Session Initiation Protocol (SIP) is a signaling protocol [2]. It is used in call setup, tear down and management of session parameters. Its architecture consists of user agents (UA), registrar, location, proxy and redirect servers. User agent is a telephony like application. Registrar server offers registration services and it updates its database as a new user arrives. Location server keeps track of the location of the users. It is updated by the registrar server on new registration. Proxy server forwards the requests and responses from the user agents. Redirect server forwards the request to possible proxy servers or user agent if the requested URI is not in its database. Mostly registrar, proxy and redirect servers are available in one software package, one example being SER [17].

SIP signaling comprises of exchange of request and response messages among the user agents through one or more SIP servers. These request messages are known as methods. The

Fig. 3. Satellite network testbed

Fig. 4. Satellite network emulator

GW - Satellite Gateway
ST - Satellite terminal

length is 40/60 bytes. The detail of the codecs and their bandwidth consumption is given in Table I. The bandwidth required by each call is computed using (1).

Delay is the one way delay between the source and destination. In geostationary (GEO) satellite systems, this delay is dominated by the propagation delay which is approximately 250-270 ms. VoIP is a real-time application, which cannot tolerate longer delays as the users will loose interactivity. According to ITU-T recommendations [12][13], one-way delay follows these constraints:

- Under 150 ms: acceptable.
- 150 to 400 ms: acceptable with limitations; and
- Over 400 ms: unacceptable.

Voice packets are transmitted by RTP. RTP identifies a voice stream by its unique Synchronization Source Identifier (SSRC). Additionally, individual packets can be identified by the port numbers, sequence numbers and timestamps [14]. The time difference of the same packet at source and destination results in its one-way delay as suggested in [5].

Jitter is the variation in delay of the successive voice packets. Jitter occurs because different packets suffer different delays in the network. Jitter contributes in the overall delay of the voice packets. It is an estimate of the inter arrival time of RTP packets and that's why it's referred as the inter arrival jitter [14]. If R represents the arrival time of a packet and S represents the RTP timestamp, then the inter arrival difference D(i,j) between two packets i, and j, can be calculated as,

$$D(i, j) = (R_j - R_i) - (S_j - S_i) = (R_j - S_j) - (R_i - S_i) \qquad (2)$$

Packet loss is also a dominant factor in degrading the voice quality. It is intolerable in time constrained applications like VoIP. Packet loss is devastating because voice packets are carried by User Datagram Protocol (UDP) which does not retransmit lost packets. Packet loss is due to congestion, interference, noise and buffer overflow at the receiver. A packet arriving after a certain scheduled play out time is also discarded. Packet loss can be reduced using forward error correction (FEC) by transmitting redundant information and
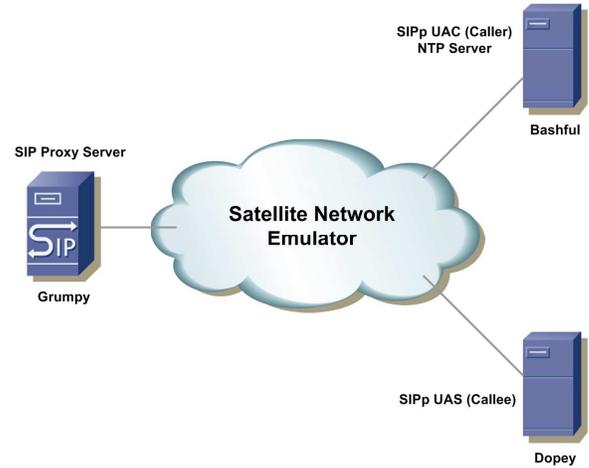
interleaving the packets. A packet loss up to 10% is acceptable in VoIP [13].

V. SATELLITE NETWORK TESTBED

The satellite network testbed at Centre for Communication Systems Research (CCSR) in University of Surrey consists of one switch and six Linux servers. The servers are named grumpy, sneezy, sleepy, happy, bashful and dopey. Sneezy, sleepy and happy emulate a satellite gateway and two satellite terminals respectively as shown in Fig. 3. A laptop with satellite emulator software, Platine [15], is connected to the switch. Now, this laptop, gateway and the two satellite terminals emulate a satellite network as depicted in Fig. 4. Grumpy, bashful and dopey are three clients where users can install their software for testing the network. After starting the emulator the clients are at a one-way delay of 250-270 ms from each other.

Platine [15] is a satellite emulator, used in the IST SATSIX project [16] to test and demonstrate various features of satellite networking. Many real satellite networks do not allow performing experiments on their infrastructure, so emulators can be used as alternatives for those real networks. Platine is designed and developed to replicate a real DVB-RCS (Digital Video Broadcasting – Return Channel via Satellite) / DVB-S2 (Digital Video broadcasting – Second). It can be configured as a satellite hub or a regenerative payload with onboard processing.

SER [17] is used as a proxy server to route VoIP calls from caller to callee and is installed on grumpy as shown in Fig. 4. SER is a high performance SIP server which can support a large number of simultaneous calls.

To automate voice calls, a freely available call generator for SIP, known as SIPp [18], has been used. SIPp has built-in XML-based caller (UAC) and callee (UAS) scenarios that can be used to generate calls at a specific rate, calls per second (cps). It can dump the running calls statistics in comma separated value (CSV) format. SIPp supports advanced features like support for IPv6, TLS, UDP retransmission and

most important being media transmission, both audio and video. It is installed on bashful and dopey as a caller and callee respectively, as shown in Fig. 4.

Network Time Protocol (NTP) [19] is used to synchronize computer clocks and promises a precision of 1 microsecond. Most of the measurements in our experiment are related to time, so all the systems should be synchronized. To meet this requirement, we installed NTP server on bashful and synchronized grumpy and dopey with its clock. They have been connected through a separate local area network (LAN) for the rapid exchange of the NTP messages. It ensures the precision of the time measurements.

VI. RESULTS AND ANALYSIS

The performance of different codecs described in Table I is measured for IPv4 and IPv6, using the satellite network testbed. SIPp [18] is used to generate calls at a rate of 0.1 cps (calls per second) from bashful (caller) to dopey (callee) through grumpy acting as the proxy server. The default call generation rate of SIPp is 10 cps but we reduced it because of the low-bandwidth satellite link. In total, 100 calls are generated. Wireshark [20], formerly known as ethereal, a packet analyzer is used to capture packets at the caller, callee and proxy server to measure and calculate different performance metrics. The bandwidth consumed on the IPv4 and IPv6 experimental setup for PCMA is shown in Fig. 5. Most of the calls consume same bandwidth, but, there are some peaks due to call setup delays. The average bandwidth consumed over IPv4 and IPv6 for all the codecs has been tabulated in Table I and their measured difference tallies the calculated difference.

The call setup time for each call is measured by finding the difference between the absolute times of INVITE and corresponding ACK messages of SIP. The cumulative sum of call setup time for PCMA is shown in Fig. 6. Rest of the codecs also exhibit similar behavior, so they have not been shown. The statistical variation of the call setup time of 100 calls for all
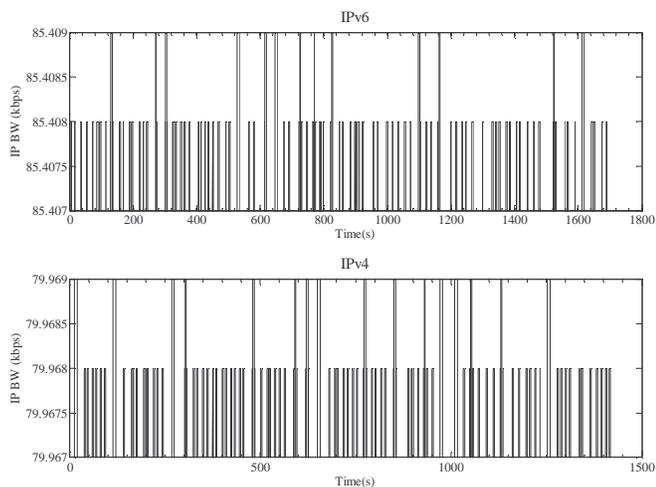
TABLE II
SIP call setup time (s)

Voice codec	IPv4		IPv6	
	Mean	Std Dev	Mean	Std Dev
PCMA	7.381	5.982	13.860	9.956
G.721	7.334	6.778	12.142	9.182
GSM	7.647	7.526	11.565	8.989
LPC	7.042	6.491	10.946	9.241

TABLE III
RTP one-way delay (ms)

Voice codec	IPv4		IPv6	
	Mean	Std Dev	Mean	Std Dev
PCMA	322.845	18.882	324.654	16.585
G.721	319.986	18.856	321.449	17.499
GSM	321.511	19.172	319.804	19.808
LPC	321.633	17.502	322.117	17.018

TABLE IV
Maximum jitter (ms)

Voice codec	IPv4		IPv6	
	Mean	Std Dev	Mean	Std Dev
PCMA	1.149	0.112	1.133	0.098
G.721	6.607	0.059	6.611	0.052
GSM	6.609	0.061	6.609	0.062
LPC	8.017	0.076	8.007	0.088

Std Dev = standard deviation

codecs is tabulated in Table II. The results show that 70% of the calls setup in 6-8 s for all codecs for IPv4 and in 12-18 s for IPv6. The main factor for higher call setup time in IPv6 is the large header size. By default, SIPp uses UDP as the transport layer protocol to carry SIP signaling messages. SIP over UDP [2] associates a timer with every SIP message starting from 500 ms and doubling after each retransmission. This results in retransmissions after 1, 2, 4, 8, 16 and 32 s. After six retransmissions, the destination is assumed unreachable. Some of the calls are established in 30 - 45 s because of these retransmissions of SIP messages. The high call setup time is alarming and some modifications in the SIP signaling are required to reduce it.

One-way delay for RTP packets puts a limitation on the real-time VoIP application for its useful operation. SIPp sends 236 RTP packets in one call for PCMA. We captured same number
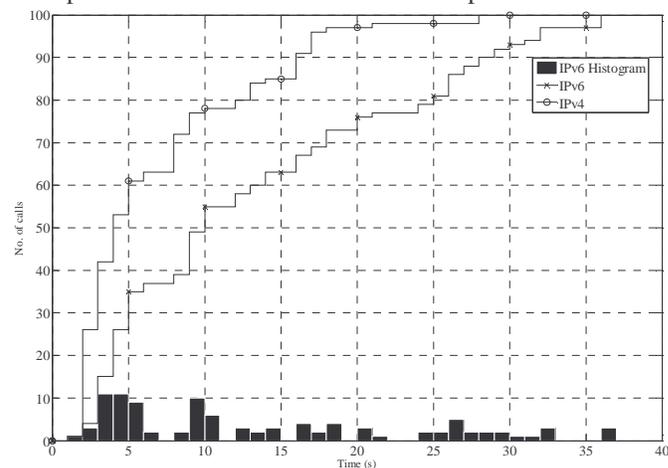
Fig. 5. Bandwidth consumption in PCMA measurement

Fig. 6. Call setup time for PCMA for 100 calls
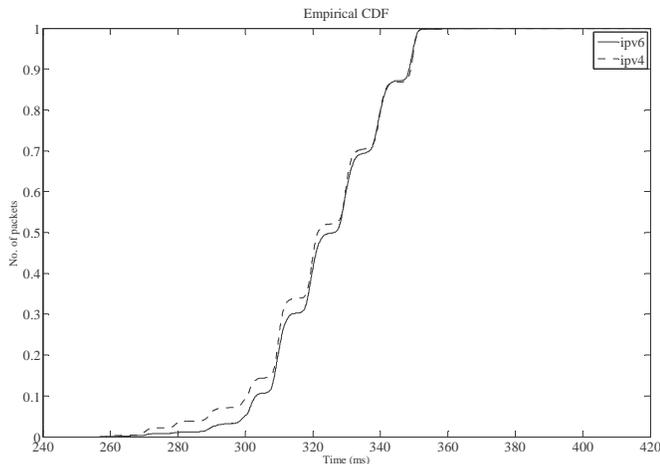
Fig. 7. RTP one-way delay for 23600 packets

Fig. 8. Maximum jitter for PCMA for 100 calls

of packets from voice streams of other codecs. So, for 100 calls, 23600 RTP packets are used for delay calculations. The statistical parameters of the delay are depicted in Table III. The cumulative distribution function (CDF) of the delay is plotted for PCMA, for both, IPv4 and IPv6, as shown in Fig. 7. This delay ranges from 250 to 550 ms. The delay for IPv4 and IPv6 is quite comparable. So, the performance of the network layer protocols is same for RTP packets carrying voice data. Therefore, IPv6 can rightfully, take over IPv4 in next generation satellite networks.

Wireshark [20] implements (2) to calculate jitter. The statistical variation of maximum jitter computed for each call has been shown in Table IV and plotted for PCMA in Fig. 8. Other codecs' plots have not been shown due to space limitation. In both IPv6 and IPv4, maximum jitter varies between 0.8 and 8.3 ms and it is the least in PCMA as compared to other codecs. It is comparable in G.721 and GSM. LPC has the highest maximum jitter. Highly compressed codecs exhibit higher jitter due to extra CPU overhead. In all the codecs, maximum jitter is tolerable for voice over IP

Packet loss of RTP packets in all these measurements was less than 1%, for both, IPv4 and IPv6, as there is no other traffic in the network. This much packet loss is tolerable in VoIP.

VII. CONCLUSION

This paper presented the evaluation of the SIP signaling and QoS for VoIP over IPv4 and IPv6 for DVB-RCS satellite networks. The results show that delay, jitter and packet loss are quite comparable for both current and next generation satellites. In terms of jitter, there is a tradeoff among the performance metrics of codecs with different compression level. Highly compressed codecs consume less bandwidth but the jitter experienced by voice packets is more. On the other hand, non-compressed or less compressed codecs are more bandwidth consuming, but their jitter is lesser. The call setup time is quite large in IPv6 in comparison with IPv4. It can be
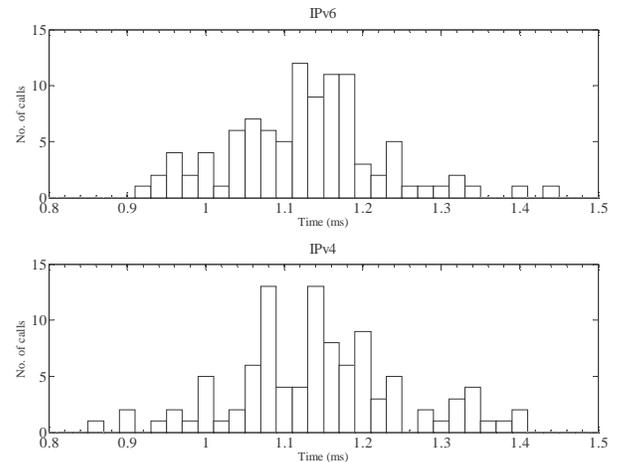
reduced with alteration in SIP timers and optimization of the algorithms for retransmission of SIP messages.

REFERENCES

[1] H. Liu and P. Mouchtaris, "Voice over IP Signaling: H.323 and Beyond," *IEEE Communications Magazine,* vol. 38, pp.142-148, October 2000.
[2] J. Rosenberg, et. al., "SIP: Session Initiation Protocol", RFC 3261, June 2002.
[3] G. Camarillo, R. Kantola and H. Schulzrinne, "Evaluation of Transport Protocols for the Session Initiation Protocol," *IEEE Network*, vol. 17, pp. 40-46, September/October 2003.
[4] M. Lulling and J. Vaughan, "A Simulation-based Comparative Evaluation of Transport Protocols for SIP," *Computer Communications*, vol. 29, pp. 525-537, February 2006.
[5] T. Hoeher, M. Petraschek, S. Tomic, and M. Hirschbichler, "Evaluating Performance Characteristics of SIP over IPv6," *Journal of Networks*, vol. 2, pp. 40-49, August 2007.
[6] M. Ali, L. Liang, Z. Sun and H. Cruickshank, "Evaluation of SIP Signaling and QoS for VoIP over Satellite Networks," *IEEE International Conference on Communications*, Germany, June 2009. *Accepted for publication.*
[7] Alex Vallejo, et. al., "State of the Art of IPv6 Conformance and Interoperability Testing," *IEEE Communications Magazine*, October 2007, pp.140-146.
[8] Dr. G. Fairhurst, "IPv6 – The Network Protocol of the Future," *IEEE Advanced Satellite Mobile Systems*, 2008.
[9] "Preparation for IPv6 in Satellite Communications: IABG," ESA Contract 17629/03/NL/ND, 2004.
[10] Ekiga. [Online]. Available: http://www.gnomemeeting.org
[11] Voice over IP – Per Call Bandwidth Consumption, [Online]. Available: http://www.cisco.com/en/US/tech/tk652/tk698/technologies_tech_note0 9186a0080094ae2.shtml
[12] ITU-T Rec. G.114, "One-way transmission time," 2003.
[13] P. Metha and S. Udani, "Voice over IP," *IEEE Potentials*, vol. 20, pp. 36-40, Oct-Nov 2001.
[14] H. Schulzrinne, S. Casner, R. Frederick, and V. Jacobson, "RTP: A Transport Protocol for Real-Time Applications," RFC 3550, July 2003.
[15] C. Baudoin, et. al., "PLATINE: DVB-S2/RCS enhanced testbed for next generation satellite networks" *IP Networking over Next-Generation Satellite Systems, International Workshop*, Budapest, July 2007.
[16] IST SATSIX Project. [Online]. Available: http://www.ist-satsix.org
[17] SIP Express Router. [Online]. Available: http://www.iptel.org/ser
[18] SIPp. [Online]. Available: http://sipp.sourceforge.net
[19] David L. Mills, "NTP: Network Time Protocol (NTP) Version 3 Specification, Implementation and Analysis," RFC 1305, Mar. 1992.
[20] Wireshark. [Online]. Available: http://www.wireshark.org