A PRACTICAL QUALITY OF SERVICE EVALUATION FOR VOICE OVER IP: INTSERV APPROACH

Virgil DOBROTA, Daniel ZINCA, Cristian Mihai VANCEA, Gabriel LAZAR
Technical University of Cluj-Napoca, Department of Communications
26-28 Baritiu Street, Room 364, 3400 Cluj-Napoca, Romania
Phone: +40-64-195699/ext.208, +40-64-413038 Fax: +40-64-197083
E-mail: Virgil.Dobrota@com.utcluj.ro, Daniel.Zinca@com.utcluj.ro

Abstract: This paper presents the preliminary results of a practical QoS (Quality of Service) evaluation for Voice over IP. The IntServ (Integrated Services) approach is based on both H.323 (which is an ITU-T recommendation) and SIP (Session Initiation Protocol), developed by IETF (Internet Engineering Task Force) in RFC 2543. An optical fiber-based infrastructure provided by CAMAN (Cluj-Napoca Academic Metropolitan ATM Network) has been used to determine the technical challenges of the telecommunications and data networks convergence. The experiments discussed herein are related to the main topology of the network, as well as the evaluation of the VoIP parameters: packet loss ratio, transfer delay and delay variation (jitter). The testbed demonstrator included Alcatel OmniPCX 4400, CISCO 1750 router, Linux-based H.323-ISDN gateway, Microsoft NetMeeting 3.0.1, eStara SoftPhone, ISDN and PSTN terminals. An initial version of this work was presented at the First RoEduNet Conference 2002 "Networking for Education and Research" [11].

Keywords: *H.323, IntServ, RTP Control Protocol, Session Initiation Protocol, Voice over IP*

I. INTRODUCTION TO THE INTERNET MULTIMEDIA

The Internet multimedia protocol stack presented in Figure 1 is related to the general well-known four-layer TCP/IP model and includes the last achievements in the field. Host-to-Network Layer could be a SDH/Sonet or a telephone line (V.34/V.90 modem) running PPP (Point-to-Point Protocol). Other possible technology is a digital subscriber line (xDSL) running ATM (Asynchronous Transfer Mode) and AAL (ATM Adaptation Layer) on top of it. Recent deployment of Gigabit Ethernet may replace the existing Ethernet/Fast Ethernet solutions for local area networks, whilst the new coming MPLS (Multi-Protocol Label Switching) is a modern choice too. On top of the Internet Layer (represented by IPv4 and IPv6) and the Transport Layer (TCP or UDP) several protocols are mixed together. The Application Layer may include signalling, QoS (Quality of Service), media transport protocols or utilities. H.323 is a major set of ITU-T (International Telecommunications Union) specifications, approved in 1996. Versions 2 and 3 were released in 1998 and 1999 respectively. It is considered as an "umbrella" standard for both standalone devices and embedded personal computer technology, as well as point-to-point and multi-point conferences. H.323 also addresses call control, multimedia management and bandwidth management (see Table 1).

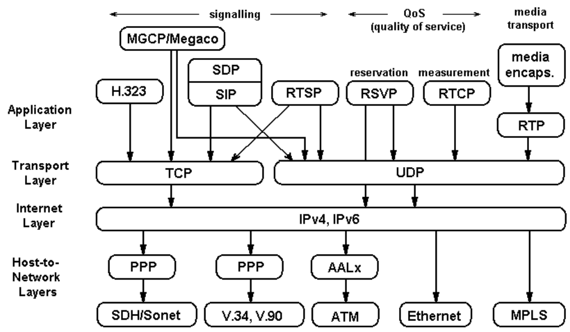

Figure 1. The Internet Multimedia Protocol Stack

H.323 Standard	Examples	Man-datory
signalling	H.225/RAS, Q.931, H.245	yes
media transport	RTP	yes
QoS	RTCP	yes
audio codec	G.711 (PCM, 64 kbps) G.722 (ADPCM, 32 kbps) G.723 (LPAS, 5.3/6.4 kbps) G.728 (LD-CELP, 16 kbps) G.729 (LD-CELP, 8 kbps)	yes no no no no
video codec	H.261 (64 kbps…2 Mbps) H.263 (>28.8 kbps) MPEG-4 (4.8…64 kbps)	no no no
data	T.120 (T.122, T.123, T.124, T.125, T.126, T.127)	no

Table 1. H.323 Standards Stack

According to IETF's principle "One Problem, One Protocol", SIP (Session Initiation Protocol), version 2.0 was firstly published as RFC 2543 in April 1999, and it was reviewed in July 2000 and February 2002, as RFC 2543 bis [7],[8]. From the beginning SIP was designed as a pure end-to-end signalling protocol, employing other protocols for transport, media transport and media description. Several analysts are expecting SIP to act as a SS7's equivalent for the future telephone communications [2]. Based on ABNF (Augmented Backus Naur Format) for representation and using a text-based encoding scheme, it borrowed some features from other Application Layer protocols. For instance the client-server architecture and the use of URLs (Uniform Resource Locators) are similar to those of the well-known HTTP (Hyper Text Transport Protocol). On the other hand plain text messages and some headers (such as `To`, `From`, `Date`, `Subject`) look like in SMTP (Simple Message Transfer Protocol). To conclude, SIP is a light weight protocol, handling call signalling, user location and basic registration. The actual description of the session in terms of time and media capabilities is performed by SDP (Session Description Protocol), as in RFC 2327 [9].

```
INVITEsip:064@172.27.208.100 SIP/2.0
Via: SIP/2.0/UDP 172.27.208.100:52692
From: "064" <sip:064@172.27.208.100>
To: <sip:012@172.27.208.161;phone-
context=unknown;user=phone>;tag=2060f
e60
Call-ID: DA7AB50A-E8190124-0-
B0343B8@172.27.208.100
CSeq: 101 INVITE
Content-Length: 173
Contact: sip:student@172.27.208.161
Content-Type: application/sdp
User-Agent: eStara SoftPHONE

v=0
o=eStara 665838 665838 IN IP4
172.27.208.161
s=eStara
c=IN IP4 172.27.208.161
t=0 0
m=audio 8010 RTP/AVP 0 101
a=rtpmap:101 telephone-event/8000
a=fmtp:101 0-15
```

Figure 2. SIP+SDP message

The request message INVITE is an example of a complete Setup Message (SIP+SDP), as in Figure 2. Let us suppose that the telephone number 064 (IP source address = 172.27.208.100) is calling 012 (IP destination address = 172.27.208.161). A 173-byte SDP message is enclosed by the user agent eStara SoftPHONE within SIP message. The fields are the following: version = 0; origin = IPv4 172.27.208.161; subject = eStara phone call; connection to 172.27.208.161; time = 0; media format = audio; media transport = RTP (Real Time Protocol) at port number 8010; some other attributes. For more details, see [2].

RTSP (Real Time Streaming Protocol) is an Application Layer protocol which establishes and controls either a single or several time-synchronized streams of continuous media (audio, video). It can aggregate multiple streams, supports unicast and multicast and is also a text-based, with a syntax similar to HTTP [5].

II. QUALITY OF SERVICE FOR VOICE OVER IP

Voice transmission over Internet is generally considered to be acceptable if there are only few lost messages, the average transfer delay is reasonable and the delay variation (jitter) is not excessive. The complete definitions of these parameters are very similar to those applied to fixed-length packets, i.e. ATM cells. Note that the jitter is considered as an additional delay, which determines the receiver to become unable to recover the speech stream. In this case, it will interpolate or generate a silent period.

The QoS measurement should be performed between two network access points, separately for each communication direction. Several profiles may be defined, based on the default values for packet loss ratio, return transfer delay and maximum jitter.

QoS	Packet loss ratio [%]	Return transfer delay [ms]	Max. jitter [ms]	Listening voice quality
Min.	100	10000	1000	Unusable
Profile1	5	600	75	Bad
Profile2	3	400	50	Average
Profile3	2	150	20	Good
Max.	0	0	0	Excellent

Table 2. Default values for QoS in VoIP at Alcatel OmniPCX 4400

Packet loss is important because it tracks persistent congestion, whilst the jitter tracks a transient congestion before it leads to packet loss.

Note that the lowest category encountered at any of these three parameters is determining the general QoS. For example, if the packet loss ratio is according to "Profile2", the return transfer delay is close to "Profile3", but the jitter corresponds to "Profile1", the link is considered of "Profile1".

A QoS Requirement is established by the management system for each terminal. The adopted principle is stating that all communications initiated by a terminal will have the same QoS requirement. This observation is also valid for a group of trunks. QoS requirement could be used as follows [3]:

- *Never Voice over IP*: the communication will never use the IP network, whatever value QoS has.

- *Always Voice over IP*: the communication will always use the IP network, whatever value QoS has.
- *Profile3*: the communication will use the IP network, for QoS of Profile3.
- *Profile2*: the communication will use the IP network, for QoS of Profile2 or Profile3.
- *Profile1*: the communication will use the IP network, for QoS of Profile1, Profile 2 or Profile3.

Supposing that the quality of service requirement cannot be fulfilled, a new IP route should be chosen. In case of path determination failure, the communication will be canceled. If the QoS is unknown and the QoS requirement is different to *Never VoIP*, the link will be established and some statistics may be used for future calls. Usually the system will not check the QoS requirement during the call. If the QoS decreases, the listening voice quality is decreasing for all participants. In case the link is becoming unusable, the only one choice available for users is to release the call. Next time they are trying to establish a connection, due to the fact that QoS has been updated, the routing system will choose another path. An unknown QoS requirement-based call will never be suspended . Note that usually QoS is valid few minutes after the last call only, but the system's management could adjust it. As an example, let us see the VoIP parameters of Alcatel OmniPCX 4400 by typing:

```
> compvisu sys
=================================
C O M P V I S U
=================================
VAD (Voice Activity Detection)no
ECE (Echo Canceller)......... yes
PFE (Post Filter)............ no
Volume ...................... 8
VRE ......................... no
Law ..................... A law
Global compression type .. G723
IP version................ IPv4
IP QoS Data Life Time..... 10 min

profile packet_loss jitter  delay
QoS inferior(#1) 20% 200ms  800ms
QoS medium  (#2) 10% 100ms  400ms
QoS superior(#3)  5%  20ms  150ms
=================================
```

Figure 3. QoS measurement at Alcatel 4400

III. RTP CONTROL PROTOCOL (RTCP)

One major step towards the evaluation of QoS is RTCP (RTP Control Protocol), which is based on the periodic transmission of control packets. They are using the same paths as data packets, but the services are offered at different UDP ports.

RFC 1889 defines several RTCP packet types such as:

1. *SDES* (Source Description) is including the canonical end-point identifier CNAME, which is unique among all participants within a RTP session. This identifier could be an e-mail address, for example `root@p2.el.obs.utcluj.ro`, as in Figure 4. Other relevant information is TOOL, e.g. ISDN-H.323 gateway [6].

```
⊟ Real-time Transport Control Protocol
    Version: RFC 1889 Version (2)
    Padding: False
    Source count: 1
    Packet type: Source description (202)
    Length: 11
  ⊟ Chunk 1, SSRC/CSRC 2129287575
      Identifier: 2129287575
    ⊟ SDES items
        Type: CNAME (user and domain) (1)
        Length: 7
        Text: root@p2
        Type: TOOL (name/version of source app) (6)
        Length: 20
        Text: ISDN - H.323 gateway
```

Figure 4. Example of RTCP - Source Description

2. *SR* (Sender Report) is issued by an active sender as often as bandwidth constraints allow (normally less than 5% of the total traffic). The session bandwidth is independent with respect to the media encoding, but the encoder should take care of bandwidth.

```
⊟ Real-time Transport Control Protocol
    Version: RFC 1889 Version (2)
    Padding: False
    Reception report count: 1
    Packet type: Sender Report (200)
    Length: 12
    Sender SSRC: 2129287575
    Timestamp, MSW: 3224588906
    Timestamp, LSW: 3198720832
    RTP timestamp: 144000
    Sender's packet count: 601
    Sender's octet count: 144240
  ⊟ Source 1
      Identifier: 793046859
    ⊟ SSRC contents
        Fraction lost: 0 / 256
        Cumulative number of packets lost: 0
    ⊟ Extended highest sequence number received: 0
        Sequence number cycles count: 0
        Highest sequence number received: 0
      Interarrival jitter: 49
      Last SR timestamp: 0
      Delay since last SR timestamp: 0
```

Figure 5. Example of RTCP - Sender Report

3. *RR* (Receiver Report) is issued by a non-active sender and includes from zero up to 31 reception blocks. Actually it is similar to SR, except the 20-byte sender information section. The active senders could sent also RR if the site has sent no data packets during the interval since the last report transmission.

Other RTCP packet types are BYE, which indicates the end of participation and APP, the application specific functions.

```
⊟ Real-time Transport Control Protocol
    Version: RFC 1889 Version (2)
    Padding: False
    Reception report count: 1
    Packet type: Receiver Report (201)
    Length: 7
    Sender SSRC: 793046859
  ⊟ Source 1
      Identifier: 2129287575
    ⊟ SSRC contents
        Fraction lost: 0 / 256
        Cumulative number of packets lost: 0
    ⊟ Extended highest sequence number received: 50132
        Sequence number cycles count: 0
        Highest sequence number received: 50132
      Interarrival jitter: 80
      Last SR timestamp: 1382923944
      Delay since last SR timestamp: 52608
```

Figure 6. Example of RTCP - Receiver Report

Interarrival jitter J is calculated continuously according to the following equation:

$$J = J + (|D(i-1,i)| - J)/16 \qquad (1)$$

where the mean deviation $D(i-1,i)$ of the RTP timestamps for two consecutive packets may be expressed as:

$$D(i-1,i) = [R(i) - R(i-1)] - [S(i) - S(i-1)] \qquad (2)$$

$R(i)$ and $S(i)$ represent the reception, respectively the transmission RTP timestamps. The gain parameter 1/16 gives an acceptable noise reduction ratio, while maintains a reasonable rate of convergence.

The round-trip time computation at the sender's site is based on the time A when the reception report block is received from a given destination. It calculates the last SR timestamp (LSR) as the middle 32 bits out of 64 in the NTP (Network Time Protocol) timestamp received within the most recent RTCP sender report packet. The return transfer delay is:

$$RTT = A - LSR - DLSR \qquad (3)$$

where DLSR represents the delay since last SR, as in Figure 5 or Figure 6.

IV. EXPERIMENTAL RESULTS

The VoIP testbed demonstrator was designed according to the specific voice/data communications needs within CAMAN (Cluj-Napoca ATM Metropolitan Academic Network). The experiments were carried out mainly for TUCN buildings, which are geographically distributed within the city. The telecommunications node Dorobantilor is based on Alcatel OmniPCX 4400, which is an IP private branch exchange, with ISDN/ PSTN access. PC-based terminals running Microsoft NetMeeting 3.0.1, will act as H.323 terminals, whilst eStara SoftPhone will transform them into SIP User Agents. According to Figure 7, two types of gateways were included: a CISCO 1750 router, acting as H.323 or SIP to PSTN gateway, and a Linux-based H.323 - ISDN gateway. Obviously the work is under progress, so other implementations might be envisaged for evaluation. There were the following groups of experiments performed:

PC-to-Phone:
- Experiment 1: H.323 - Cisco GW H.323 – PSTN terminal
- Experiment 2: H.323 - Linux GW H.323 - ISDN terminal
- Experiment 5: SIP User Agent - Cisco GW SIP - PSTN

PC-to-PC:
- Experiment 3: H.323 terminal – H.323 terminal
- Experiment 6: SIP User Agent - SIP User Agent
- Experiment 7: SIP User Agent - Cisco GW SIP - Linux GW H.323 - H.323 terminal

Phone-to-Phone:
- Experiment 4: PSTN terminal - Cisco GW H.323 or SIP - Linux GW H.323 or SIP - ISDN terminal

The measurement methods were described within section III. Experiment 1 conditions were the following: about 560 seconds of real traffic (11 millions of bytes), i.e. 51728 packets with an average throughput of 19 kbps.

Packet No.	Time [s]	Cumulative packet lost	Inter-arrival jitter	DLSR [s]
197	17.244797	2	2	0.000000
593	23.183336	3	7	2.552734
881	27.559629	5	0	0.690430
…	…	…	…	…
10371	125.570562	7	0	3.214844
…	…	…	…	…
20235	229.399861	16	0	1.071289
…	…	…	…	…
30332	332.768498	24	0	2.914063
…	…	…	…	…
40380	437.549165	63	0	3.524414
…	…	…	…	…
51612	558.573189	179	0	0.810547

Table 3. Experiment 1: H.323 terminal -> Cisco GW H.323 -> PSTN

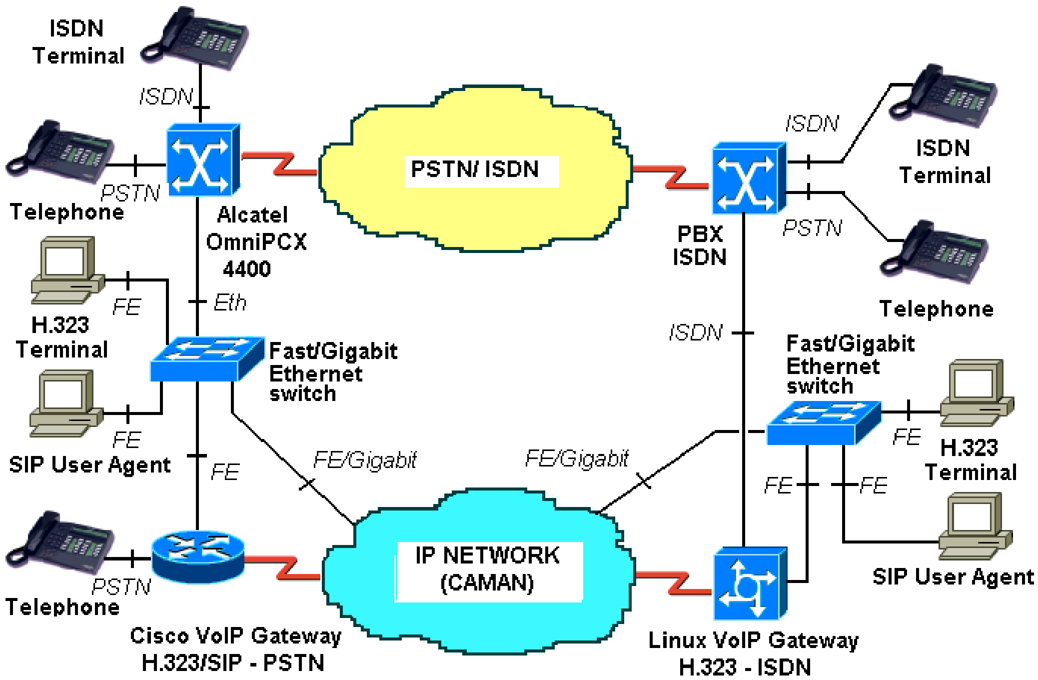

Figure 7. VoIP Testbed Demonstrator

Packet No.	Time [s]	Cumulative packet lost	Inter-arrival jitter	DLSR [s]
209	17.354956	0	512	0.107986
391	20.629664	0	344	3.383987
832	26.868636	0	112	3.683990
…	…	…	…	…
10065	122.355940	0	80	3.215988
…	…	…	…	…
20128	228.328320	4	80	3.303986
…	…	…	…	…
30040	329.854308	9	112	1.287994
…	…	…	…	…
40025	434.024096	20	112	1.519989
…	…	…	…	…
51725	559.784932	27	208	1.219986

Table 4. Experiment 1: PSTN -> Cisco GW H.323 -> H.323 terminal

According to Table 3 and Table 4, the packet number 197, having the arrival time 17.244797 is correlated to packet number 209 (arrival time 17.354956). As the DLSR is 0.107986, the return transfer delay from H.323 terminal to CISCO GW is 2.173 ms.

The interarrival jitter and cumulative packet lost graphics are presented in Figure 8 and Figure 9 for both communication directions.

By the time this paper was submitted, several measurements were under progress. The preliminary results are presented in Table 5.

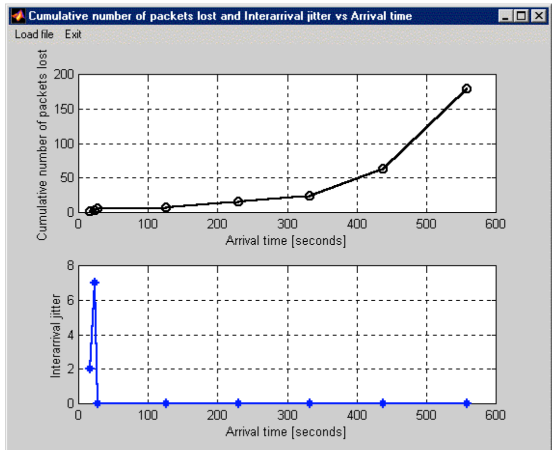

Figure 8. Cumulative packet lost and interarrival jitter for H.323 terminal -> Cisco GW H.323 -> PSTN

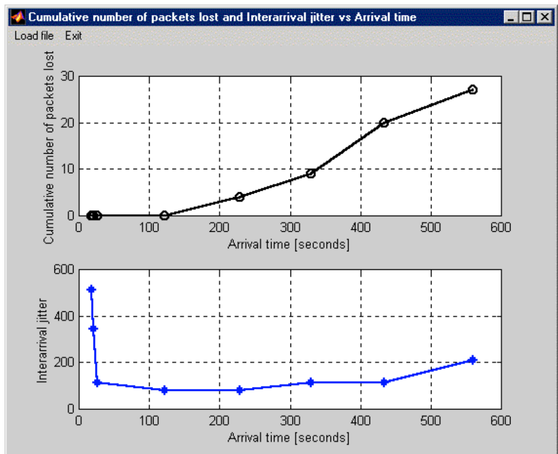

Figure 9. Cumulative packet lost and interarrival jitter for PSTN -> Cisco GW H.323 -> H.323 terminal

Exp	Equipment	RTCP	RSVP	Notes
1	H.323 terminal	All	Yes	-
1	CISCO GW H.323-PSTN	All	No	RSVP must be enabled
1,5	PSTN terminal	N.A.	N.A.	Speech quality [16]
2	H.323 terminal	All	No	-
2	Linux GW H.323-ISDN	Jitter	No	No useful RR
2	ISDN terminal	N.A.	N.A.	Work under progress
3	H.323 terminal	All	Yes	No RR for one site
5, 6	SIP User Agent	None	No	No RTCP or RSVP packets
5	CISCO GW SIP-PSTN	Jitter, Packet lost	No	No RTT calculated

Table 5. Preliminary results

V. CONCLUSIONS AND FURTHER WORK

We built our own QoS management tool, designed for both IETF's models: IntServ (Integrated Services) and DiffServ (Differentiated Services). The IntServ approach is based on capturing, analyzing and generating of RTCP (RTP Control Protocol) and RSVP (Resource Reservation Protocol) packets. The packet loss ratio, the transfer delay and the interarrival jitter are determined and compared to QoS requirement. Cumulative number of packet lost is important because it tracks persistent congestion, whilst the jitter tracks a transient congestion before it leads to packet loss. According to preliminary results, several current SIP User Agent implementations were not able to perform IntServ at all. Note that the management tool may initiate also a QoS-enabled link on behalf of terminals which are not implementing RTCP/RSVP. The end-to-end evaluation of VoIP parameters between PSTN/ISDN terminals is also under progress [10].

It is for further work to study the DiffServ approach, supposing that a client (or a node) is able to switch from QoS on a per-flow basis (as in IntServ approach) to QoS on a per-hop behavior [12]

Due to the fact that Voice over IP protocol stack uses dynamically allocated ports above 1024 for audio and data channels, the security aspects may request firewalls. On the other hand, intelligent algorithms should parse the TCP/UDP headers and leave open the ports for the duration of the call only. Another problem to be solved is the potential return of investment (ROI). Cisco Systems has developed its own financial modeling tool called CNIC (Converged Network Investment Calculator). It is for further work to investigate the optimal migration costs from TDM (Time Division Multiplexing)-based equipment to IP telephony within CAMAN.

ACKNOWLEDGMENTS

We would like to acknowledge the support from RoEduNet Cluj team (led by Professor Pusztai Kalman), Horia Toader (technical manager, Ericsson Academy Romania), Ileana Creosteanu and Dan Matei (TOP 9+ Bucharest). As the equipment involved in the testbed demonstrator was partially supported by the World Bank Project 188/1, special thanks go to Professor Monica Borda.

REFERENCES

[1] R.G. Cole and J.H. Rosenbluth, "Voice over IP Performance Monitoring", *Computer Communication Review*, ACM-SIGCOMM, Vol.31, No.2, April 2001, pp. 9-24.

[2] Johnston, Alan B., *SIP: Understanding the Session Initiation Protocol*, Artech House, Boston, London, 2001.

[3] ***, "Alcatel Omni PCX 4400 R4.1 System Documentation", *Alcatel*, France, Reference: 3BA19918ENAA, Edition March 2001

[4] ***, "Guide to Cisco Systems' VoIP Infrastructure Solution for SIP", Version 1.0, *Cisco Systems, Inc.,* San Jose, CA, USA, 2001.

[5] ***, "Voice over IP", *Telefonaktiebolaget LM Ericsson,* Second Release, December 2000, LZU 102 324 RIC

[6] H. Schultzrinne, S. Casner, R. Frederick and V. Jacobson, "RTP: A Transport Protocol for Real-Time Applications", *RFC 1889*, IETF, January 1996

[7] M. Handley, H. Schulzrinne, E. Schooler and J. Rosenberg, "SIP: Session Initiation Protocol", *RFC 2543*, IETF, March 1999

[8] J. Rosenberg, H. Schulzrinne, G.Camarillo, A. Johnston, J. Peterson, R. Sparks, M. Handley and E. Schooler, "SIP: Session Initiation Protocol", *draft-ietf-sip-rfc2543bis-09.txt*, IETF February 2002

[9] M. Handley and V. Jacobson, "SDP: Session Description Protocol", *RFC 2327*, IETF, April 1998

[10] F. Kettler, H.W. Gierlich, J. Berger, I. Kliche, K.-D. Michael, T. Scheerbarth, R. Scholl and J.-L. Freisse, "End-to-End Speech Quality Testing in a Complex Transmission Scenario", *HEAD acoustics GmbH*, 2002.

[11] V.Dobrota, D.Zinca, C.M. Vancea, G.Lazar, "Voice over IP Solutions for CAMAN: H.323 vs. SIP", *Proceedings of the 1st RoEduNet Conference*, Cluj-Napoca, 18-19 April 2002, pp.37-43.

[12] V. Fineberg, "A Practical Architecture for Implementing End-to-End QoS in an IP Network", *IEEE Communications Magazine*, Vol.40, No.1, January 2002, pp.122-130