

A CASE STUDY OF VoIP PERFORMANCE ACROSS DIFFERENT NETWORKS

Mohammad Sabbir Alam,¹ Ashir Ahmed,² and Michael Cohen¹

¹Spatial Media Group; University of Aizu; Aizu-Wakamatsu, Fukushima-ken 965-8580; Japan

²NTT Communications; 3-20-2 Nishi-Shinjuku, Shinjuku-ku; Tokyo 163-1421; Japan

E-mail: {d8062102, mcohen}@u-aizu.ac.jp, a.ahmed@ntt.com

ABSTRACT

SIP, the Session Initiation Protocol, is a signaling protocol for internet conferencing, telephony, presence applications, event notification, and instant messaging. This article reports observed VoIP traffic when clients are in a LAN, a WAN, and in a hybrid network with an ADSL link. The deployed architecture consists of a SIP server (with registrar, proxy, and presence functions) and two SIP clients. Using a traffic analyzer, we monitored the SIP messages as well as the real-time quality, measured in the form of delay, packet loss, and jitter.

1. INTRODUCTION

WWW, e-mail, and text-chat have been the most popular internet applications. Voice Over Internet Protocol (VoIP) has been marked as the next “killer disruptive technology.” Several protocols have been proposed for VoIP session control; H.323 and SIP [1] are the most significant. SIP is an application-layer control (signaling) protocol for creating, modifying, and terminating sessions with one or more participants. These sessions include internet telephone calls, multimedia distribution, and multimedia conferences [2]. SIP normally runs over TCP (Transmission Control Protocol) [3] or UDP (User Datagram Protocol) [4]. SIP messages, monitored by the diagnostic tool Ethereal [5], are transmitted via UDP in this research. Two protocols that are most often used along with SIP are RTP (Real Time Protocol) [6], used to carry real-time multimedia data (including audio, video, and text), and SDP (Session Description Protocol) [7], used to describe and encode capabilities of session participants.

In this research, we study SIP: the protocol and its applications, and also observe VoIP characteristics across different networks— namely in a LAN, a WAN, and in a hybrid network with an ADSL link.

The SIP environment that we deployed in our lab consists of a SIP server (with registrar, proxy, and presence functions) and two SIP clients. The case study results characterize two sets of parameters: (1) signaling performance— the sequences of SIP messages, and (2) application performance— the VoIP performance (delay, jitter, and packet loss) characterizing the network. To characterize the performance of the different network arrangements, we performed an evaluation on captured RTP packets.

2. SIP AND ITS PERFORMANCE

2.1 SIP Components

SIP is a peer-to-peer protocol. The peers in a session are called user agents. A user agent can function in one of the following roles:

- User-agent client (UAC)— A client application that initiates a SIP request.
- User-agent server (UAS)— A server application that contacts the user when a SIP request is received and that returns a response on behalf of the user.

A SIP end point is capable of functioning as both a UAC and a UAS, but typically functions only as one or the other per session, depending on the user agent that initiated the request.

2.2 How SIP Works

As shown in Fig. 1, a caller sends an INVITE request through a proxy. The proxy immediately returns a 100/Trying reply to the caller and forwards the request towards the callee. A 180/Ringing indicates that the callee is being alerted. A 200/OK message is generated once the callee’s user agent accepts the INVITE. Upon receiving a 200/OK from the callee,

the caller sends an ACK, usually triggered by a human user. The session is established at this point. Therefore, there is an out-of-signal delay between the 180/Ringing and 200/OK messages, depending on when (after how many rings) the user accepts the call. Parties wishing to tear down a session send a BYE request within the session dialog.

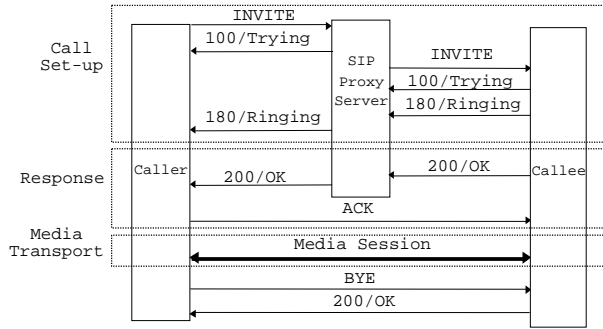


Fig. 1 Call flow of a SIP session Scenario

2.3 SIP Performance

Considering the delay we have split the signaling performance into two: (1) call set-up time (INVITE to Ringing) and (2) response time (200/OK to ACK). Fig. 1 shows the two steps of signaling before starting the media session. Human interaction occurs between these two steps. SIP communication includes audio, video, and text messages. We observed the real-time quality, based on the following parameters:

- Transmission delay, round-trip latency
- Jitter, variability of delay
- Packet loss, dropped data

The International Telecommunication Union (ITU) has determined recommendations (Table 1) for the maximum round trip delay, jitter, and packet loss [8,9].

Table 1: ITU Network Performance Recommendations

Round trip delay (ms)	Max jitter (ms)	Packet loss (%)	Voice quality
> 400	> 50	> 3	Poor
400	50	3	Average
150	20	1	Good
< 150	< 20	0	Excellent

The above parameters motivate characterization of different network environments, especially for real-time traffic. A high-performance network is characterized by high bandwidth, small delay, and low packet loss.

3. CASE STUDY: RESULTS AND ANALYSIS

After establishing a SIP session, the media session is started. In the media session, RTP packets transmit voice data between the clients.

3.1 Experimental Set-up

The set-up comprises one SIP server and two clients. The server software SER (SIP Express Router) [10] runs on a Linux-based PC. The client has Windows Messenger [11] installed on Windows XP. One of the clients (Client_1 in the following figures) has the packet-capturing tool Ethereal (v. 0.10.4) installed. In a typical scenario, one client initiates a call and terminates the call after some voice communication. Three configurations follow:

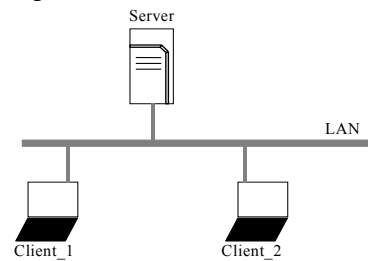


Fig. 2 LAN Scenario

- In the LAN scenario, shown in Fig. 2, two clients reside on the same LAN.

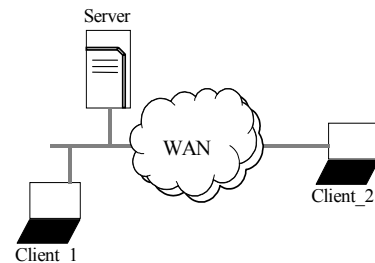


Fig. 3 WAN Scenario

- In the WAN scenario, shown in Fig. 3, Client_2 is at another university, connected via WAN link.

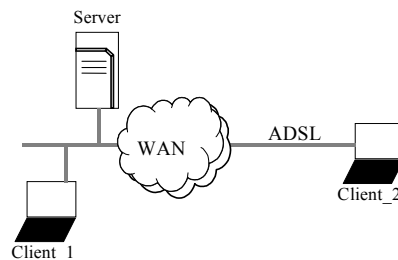


Fig. 4 Hybrid Scenario

- In the Hybrid network scenario, shown in Fig. 4, Client_2 is connected to the Internet through an ADSL link.

3.2 Observation and Analysis

3.2.1 Signalling Performance

Using Ethereal, we captured and analyzed a sequence of SIP messages. This confirms a SIP session containing an INVITE request (to initiate the call), ringing status, and BYE request to terminate the SIP session. We captured SIP messages at different times of day and on different days of the week. In Figs. 6–11, the X-axis shows the trials and the Y-axis shows the signaling performance.

On the average, it took around 245 ms to alert the other party (step 1: from INVITE to Ringing) in the LAN. However, the set-up time varied from 110 to 278 ms (Fig. 6). Fig. 7 shows the response time (step 2: from 200/OK to ACK). The 200/OK message goes through the Proxy, while the ACK goes directly to the peer. The average performance observed in this case was 4.3 ms.

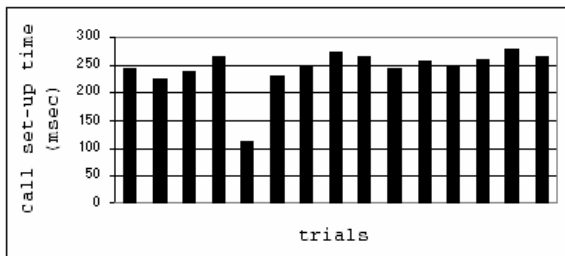


Fig. 6 Call set-up time in LAN environment

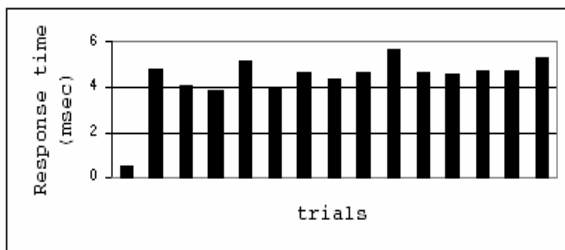


Fig. 7 Response time in LAN environment

Interestingly, call set-up time observed in the WAN environment was less than that observed in the LAN. The call set-up times are around 150 ms on the average. There is an unusual 960 ms spike also, which we speculate, occurred when the network was highly congested. The average response time is around 4.4 ms, as shown in Fig. 9.

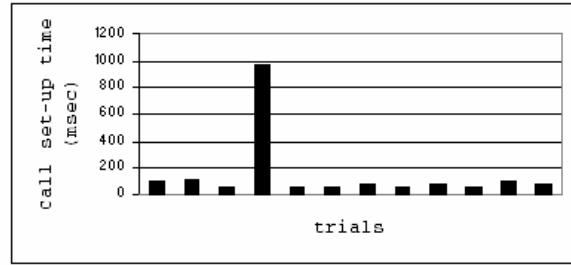


Fig. 8 Call set-up time in WAN environment

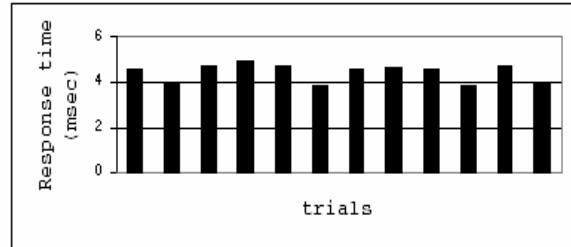


Fig. 9 Response time in WAN environment

Fig. 10 shows call set-up times of 12 observed trials in our hybrid environment, in the range of 185 to 835 ms. A bigger delay is expected in this scenario because there are more hops compared to the LAN and WAN. We see the expected variety of the performance in this scenario because the WAN and ADSL links have more dynamic characteristics, i.e. the congestion pattern changes. All the 12 trials results show less than 12 ms response time, as shown in Fig. 11.

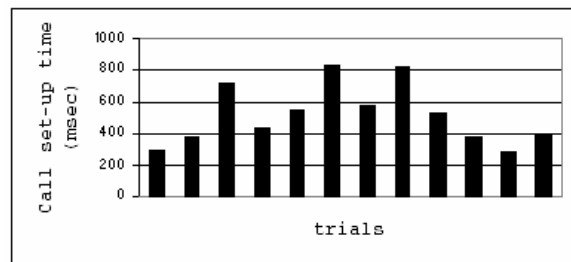


Fig. 10 Call set-up time in Hybrid environment

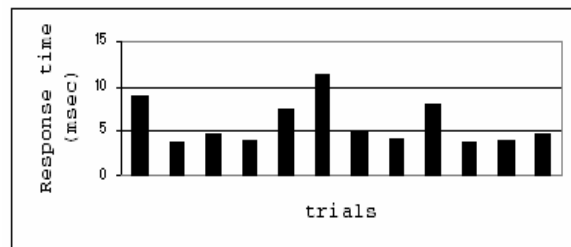


Fig. 11 Response time in Hybrid environment

3.2.2 VoIP Performance

The measurements are based on monitoring of RTP packets transmitted through different network scenarios. The figures below show the results. The X-axis of these graphs is the time in seconds. The left Y-axis shows the delay (gray) and the right Y-axis indicates jitter (black), both in milliseconds.

As we have used the normal internet, obtained results are affected by other network traffic at the time of observation. The results displayed here were selected as typical of each network scenario.

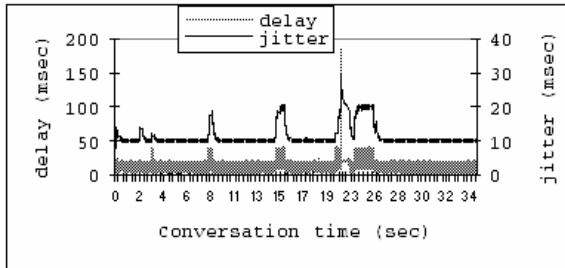


Fig. 12 Observed in LAN

In the LAN (Fig. 12), delay ranges from 3 to 40 ms. However some spikes observed during the session.

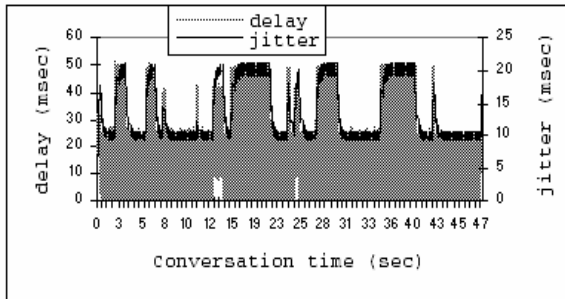


Fig. 13 Observed in WAN

In the WAN (Fig. 13), delay fluctuated from 5 to 40 ms. Compared to the LAN, delay and jitter were observed more frequently. Although both jitter and delay remained in the tolerable range, fluctuated more frequently than LAN scenario.

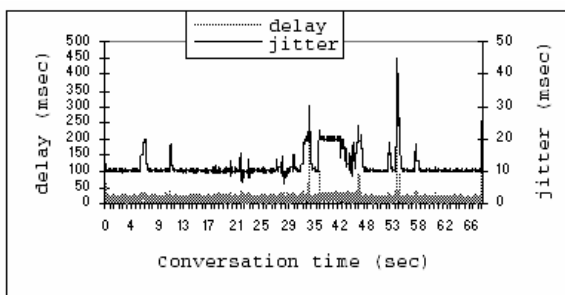


Fig. 14 Observed in Hybrid

Frequent high spikes observed in hybrid scenario (Fig. 14) went into the intolerable range. Generally, the delay stays below 25 ms, but the spikes seen exceed 400 ms, affecting jitter as well. Human users might have heard irritating noise at such times.

The delay affects the jitter since jitter is cumulatively calculated from current delay and the previous jitter values. The derived jitter metrics use the IETF-suggested expression [6] $j[n] = j[n-1] + (|\Delta d| - j[n-1])/16$, where $|\Delta d|$ is the absolute value of the difference between the last two observed packet delays. These growth/decay patterns in derived jitter metrics can be modeled as exponentially asymptotic expressions: $j[n] \sim C_1 (1 - b^{-n/T})$ for growth and $j[n] \sim C_2 (b^{-n/T})$ for decay—where n is the packet sample index, $j[n]$ is the historically damped jitter value, C_1 and C_2 are scaling constants matching hypothetical unchanging delay values, b is an exponential base, and T is a time constant modeling sluggishness of the convergence.

4. CONCLUSION

In this article, we presented monitored SIP signaling and application performance across different network scenarios. Future research includes monitoring performance on a wireless link as well using mobile SIP terminals.

REFERENCES

- [1] Alan B. Johnston, SIP: Understanding the Session Initiation Protocol, Second Edition, Artech House, 2004.
- [2] J. Rosenberg, H. Schulzrinne, G. Camarillo, A. Johnston, J. Peterson, R. Sparks, M. Handley, E. Schooler, RFC-3261— SIP: Session Initiation Protocol, 2002, www.ietf.org/rfc/rfc3261.txt
- [3] Information Sciences Institute, University of Southern California, RFC-793— Transmission Control Protocol 1981, www.ibiblio.org/pub/docs/rfc/rfc793.txt
- [4] J. Postel, RFC-768— User Datagram Protocol, 1980, www.ietf.org/rfc/rfc0768.txt?number=768
- [5] Ethereal: A Network Protocol Analyzer, www.ethereal.com
- [6] H. Schulzrinne, GMD Fokus, S. Casner, R. Frederick, V. Jacobson, RFC-1889— RTP: A Transport Protocol for Real-Time Application, 1996, www.ietf.org/rfc/rfc1889.txt
- [7] M. Handley, V. Jacobson, RFC-2327— SDP: Session Description Protocol, 1998, www.ietf.org/rfc/rfc2327.txt
- [8] Cisco Systems, Inc., “Overview of SIP,” 2004, www.cisco.com/univercd/cc/td/doc/product/software/ios123/123cgcr/vvfax_c/callc_c/sip_c/sip1_c/chapter0.htm
- [9] Cisco Systems, Inc., “Understanding Delay in Packet Voice Networks,” Document ID 5125, 2004, www.cisco.com/warp/public/788/voip/delay-details.html
- [10] SIP Express Router, www.iptel.org/ser
- [11] Free Instant Messaging service, www.microsoft.com/windows/messenger/default.asp