SIP over IP VPN: Performance Analysis

A. Asadi Eskandar¹, Mahbubur R Syed¹, and M.B. Zarei
¹ Computer Information Science Department, Minnesota State University, Mankato, MN, U.S.A

Abstract – With rapid growth in use of multimedia applications, including IP Telephony (also known as Voice over IP), the demand for security and privacy of communications has significantly increased. Given the fact that IP Telephony utilizes public IP infrastructure, deployment of IP VPN is one approach to protect traffic of interest. However, VPN is presumed to have a negative impact on VOIP performance due to additional packet overheads, authentication, integrity check, encryption, and extra CPU processing involved in the process. In this paper we simulate behavior of a SIP-based VOIP connection running over an IP VPN tunnel. OPNET IT GURU is used as the simulation tool. Few studies examined role of Quality of Service (QoS) parameters such as call signaling and employed VPN protocols, but the role of parameters’ arising from network environment, including design issues, were neglected. Bearing this in mind, this research considers several possible scenarios. It compares performance of VOIP calls based on end-to-end delay, delay variation and call setup time for different network configuration (with and without VPN). VPN itself didn’t show any negative impact on end-to-end delay. However, combination of network design and VPN show an impact on end-to-end delay.

Keywords: Session Initiation Protocol (SIP), Voice over IP (VOIP), Virtual Private Network (VPN)

1 Introduction

Voice over IP (VOIP) uses public IP network and its use has been growing rapidly in the past decade. Demand for security and privacy in recent years has prompted various solutions by security researchers and industry role players to assure secure communications over public networks, e.g., Internet. Confidentiality, integrity and authentication of a connection have been considered as a vital security requirement. According to Symantec [1] “registration hijacking” and “eavesdropping” are the two main attacks against SIP-based VOIP communications. Registration hijacking, as its name suggests, involves stealing a VOIP user’s subscription and ensuing packets. Eavesdropping, involves intercepting the voice communication streams (RTP packets). In brief, SIP-based VOIP communications are vulnerable to above mentioned attacks because SIP signaling messages are sent in clear text and SIP implementation does not support message integrity check, which makes it easy to modify and replay SIP signaling messages. Thus, in the case of VOIP the session encryption (to avoid issues such as eavesdropping, and man in middle attack), session integrity (to assure VOIP packets are not altered intentionally or unintentionally as they transit over Internet), and protection of related data to a VOIP connection such as packets pertaining to billing system, are the main concerns that needs to be guaranteed.

There exist several well-known and widely used protocols to provide different security and protection features for IP based communications. IPsec (implemented at IP layer), SSL and TLS (implemented at transport layer) are examples of such protocols and mechanisms. One elegant approach towards protecting IP packets is to implement a VPN (Virtual Private Network). Either a site to site VPN or a remote access VPN mode can be designed and implemented to establish an IP tunnel over public IP network. Further details on this are covered in the next section.

Because of the general purpose design of the above mentioned protocols and mechanisms, these may be used for protecting VOIP (and multimedia) packets as well as data traffic. However, VPN may have a negative impact on VOIP performance. It may degrade the quality of VOIP communications due to additional packet overheads, authentication, integrity check, encryption, and extra processing involved in the protection process. Delay (end-to-end delay), jitter (delay variation), packet loss, and call setup time are the main factors that affect the perceived quality of VoIP traffic by the end-user [2, 3]. VPN is presumed to increase the end-to-end delay [3, 4]. The exact amount of delay is dependent on the choice of VOIP signaling protocols (such as H323, and SIP etc.), audio codecs (such as G.711, G.729, G.728, G.726, etc.), related encoding algorithms (such as PCM - Pulse Code Modulation(PCM), Adaptive Differential PCM, etc.), VPN protocols employed (such as PPTP, L2TP, IPsec, etc.), and also parameters arising from network environment [5, 6, 7]. In general, VOIP signalling protocols, network Quality of Service (QoS) parameters, and employed security protocols are the broad aspects identified...
to impact VOIP Quality of Service [5]. In the case of VOIP over VPN, types of services offered to the VPN clients will depend on the choice of VPN protocols (e.g. PPTP, L2TP, GRE, IPSec), having an effect on quality of VOIP traffic. Several researches have addressed VOIP QoS and in particular impact of VPN on VOIP quality performance [3, 5, 8, 9, 10, and 11]. VOIP signaling protocols and choice of audio codec have been examined. However, role of (various types of) VPN protocols and parameters related to the network environment have been neglected or at least have received less attention. Considering this gap, this paper intends to simulate, using OPNET IT GURU network simulation tool, the behavior of a SIP-based VOIP connection, while running over IP VPN tunnel. In particular we set out to compare performance, in terms of end-to-end delay, delay variation and call setup time of a SIP-based VOIP connection running over IP with that of a SIP connection running over IP VPN tunnel. Several possible scenarios, entailing different network designs (in terms of placement of SIP components used in the network) is also created and a comparative analysis is done in order to consider impact of network design on the afore-mentioned issue.

2 Literature Review

Telecommunication standardization bodies such as International Telecommunications Union (ITU-T) have identified a number of factors contributing to QoS for a voice connection. The identified factors include ITU-U voice codecs and algorithm, end-to-end delay, jitter, packet loss, and network design. According to ITU-U guidelines, a voice call facing a delay greater than 150 milliseconds (note: some authors refer to 200ms) and/or a jitter of greater than 20ms is not considered to be of a good quality, and accordingly any voice call facing delay of greater than 300ms and/or a jitter of greater than 50ms is considered to be of a poor quality [12, 13, 14]. Table1 (below) outlines the accepted voice quality measures.

<table>
<thead>
<tr>
<th>Network parameter</th>
<th>Good</th>
<th>Acceptable</th>
<th>Poor</th>
</tr>
</thead>
<tbody>
<tr>
<td>Delay (ms)</td>
<td>0-150</td>
<td>150-300</td>
<td>&gt; 300</td>
</tr>
<tr>
<td>Jitter (ms)</td>
<td>0-20</td>
<td>20-50</td>
<td>&gt; 50</td>
</tr>
</tbody>
</table>

End-to-end delay measures the total amount of time taken for a voice packet to traverse the network path between caller and called entity. Conditions such as poor network capacity (e.g. low bandwidth link, or a low ingress/outbound traffic rate policy associated with a connection), and congestion, will result into voice packet getting delayed, leading into a delay greater than 300ms for voice streams.

Jitter measures the variation from the regular delay time between consecutive voice packets received by the receiver. For example if sender sends consecutive voice packets every certain milliseconds, the receiver is supposed to receive these consecutive packets with roughly similar interval. As mentioned above, 50ms is the maximum tolerable jitter (Note: some authors refer to 75ms).

IP VPN though is an elegant approach towards protecting VOIP packets passing through IP networks, is known to lead to higher delays and jitter in VOIP calls. Although packet loss, ITU-T codecs & algorithms, and parameters arising from the network environment are factors that play a role in voice QoS, but delay and jitter are considered critical factors that have been examined by most of researches measuring quality of voice calls.

Gouda I. Salama et al. [3] examined “impact of IPsec on the quality of transmitting voice over communication links using OPNET simulator”. Result of their research showed that IPsec results in an increase in packet loss, end to end delay, call setup time, and jitter.

To address QoS for IPsec transmitted packets over IP network, R. Barbieri et al. [9] proposed to down size IPsec encapsulated packets (i.e. the actual IP packet inside IPsec header) by almost 4 bytes using compression. The proposed approach is criticized by Gouda I. Salama et al., [2] for neglecting actual compression time, which in turn will lead to a processing delay that might be even larger than encryption delay.

A.H. Muhamad Amin [5] presented an analysis of VOIP performance measurements using QoS parameters. Congestion was found to negatively impact voice quality parameters such as delay and jitter. VPN showed to have a similar effect on the voice traffic. In a non-ideal network environment, the voice quality parameters showed even worse results compared with an ideal network environment.

Ibrahim S. I. Alsukayti et al. [15] used OPNET modeler to investigate performance of VOIP over VPN running over a BGP/MPLS (Multi-Protocol Label Switching) network. Results suggested that not only a VPN over BGP/MPLS has not a negative impact on VOIP quality but also “positively improves performance of VOIP as compared with its performance over an MPLS network. He also showed that G729A (bit rate = 8 kb/s) is the best choice of voice codec for such scenario (i.e. BGP/MPLS VPN) due to bringing a balance between end-to-end delay and bandwidth efficiency. (Note: SIP was the VOIP signal protocol under examination in the study).

2.1 Session Initiation Protocol

Session Initiation Protocol (SIP) is one of the peer to peer VOIP standard protocols defined in RFC 2543 and standardized by the IETF MMUSIC Working Group. This protocol contains initiation, termination and also modification standards for user sessions, which consist of video or audio elements, online games, instant messaging, virtual reality or generally multimedia elements [14]. Table 2 shows SIP massages for a call establishment.
### Table 2. SIP Messages [16]

<table>
<thead>
<tr>
<th>SIP Request Type</th>
<th>Function</th>
</tr>
</thead>
<tbody>
<tr>
<td>INVITE</td>
<td>Session establishment request</td>
</tr>
<tr>
<td>ACK</td>
<td>Acknowledgement of receiving INVITE message</td>
</tr>
<tr>
<td>OPTIONS</td>
<td>Capabilities of the server is being queried</td>
</tr>
<tr>
<td>BYE</td>
<td>Client asks server to terminate the call</td>
</tr>
<tr>
<td>CANCEL</td>
<td>Cancel the request</td>
</tr>
<tr>
<td>REGISTER</td>
<td>Registering address with SIP server</td>
</tr>
</tbody>
</table>

Four logical entities defined in SIP are: User Agent (which include: UAC - User Agent Client and UAS - User Agent Server), Registrar, Redirect Server and Proxy Server. User Agent Client (UAC) creates the SIP requests. User Agent Server (UAS) receives a SIP request and responds to it by accepting, rejecting or redirecting the request [17]. Redirect Server’s job is to receive requests and reply them with a response message indicating the next place request should be sent (i.e. determining the address of the called device) [17]. For instance, the redirect server would keep track of the user’s location and reply a response indicating the location. A SIP Proxy Server acts primarily as a Router by forwarding the SIP request to the next hop. Proxy Severs also play the role of both clients and servers by accepting the users’ request and creating a requests on behalf of the users. Proxy Servers are of two types: stateful and stateless. In stateless proxies the reliability of request is not guaranteed and they have the simple function of just forwarding incoming requests to another server. In contrast stateful proxies maintain each transaction’s state and include the request and response of that transaction.

The transaction model used in SIP protocol is a request/response model. SIP is considered a text based protocol. A SIP message contains start line, header and body [4]. As mentioned above, the requests are routed to the user’s current location using proxy servers. Other responsibilities of proxy servers include user authentication and authorization for services, and call routing policy implementation. Another interesting feature provided by SIP is registration functionality to allow users upload their locations to be later used by proxy servers. SIP runs on top of some transport protocol [3]. The handshake model of the SIP protocol is well illustrated in [2, 19] and also included in figure 1.

TCP and UDP can both be employed for SIP transmission. UDP is more commonly used as the choice of Transport protocol, mainly because of its lower overhead as compared to TCP. Since unreliable UDP transport protocol can be used for SIP massage transmitting, SIP can be responsible for reliability on its own [18].

#### 2.2 Virtual Private Network (VPN)

VPNs are responsible for providing a secure connection through a public network. To meet this goal VPN establishes an IP Tunnel- a virtual point to point link between end-nodes, which are separated by a random number of networks in between [9]. VPNs are capable of establishing a secure virtual link between different branch offices. Tunneling makes a virtual lease line (by adding extra headers), and protects private information from being accessed and modified by unauthorized entities throughout the public network [10]. So, tunnel in a VPN is actually a virtual pipe, and responsible for making physical network seems transparent to the packets as they move on their way across the internet. Tunneling can also be referred as the encapsulation process of IP packets into another outer IP header. Figure 2 illustrates a VPN tunnel. VPN uses a dedicated virtual link (i.e. tunnel), for transferring data from source to destination. So, the chance for proxy interruption is minimized. There exist two types of tunnel: permanent and temporary. Permanent tunnels (or static tunnels) are highly resource intensive. This kind of tunnels can be considered too wasteful because they use huge amounts of bandwidth for transmitting not very much data. The disadvantage could be more obvious in a business environment where they are not used 24 hours a day. So in practice a VPN does not utilize a static pipe, instead it uses the more efficient alternative which is dynamic or temporary pipes. These types of tunnel are more flexible considering the fact that they can be established and removed any time needed [12].
There are two types of VPNs: remote-access and site-to-site. The remote-access VPN establishes a virtual connection between the remote user and the company’s internal network. In site-to-site VPN, a virtual connection is established among more than one fixed site via a public network. This type of VPN needs exclusive equipments, because outbound and inbound traffic traverse through peer VPN gateways located in each site. VPN gateways are responsible for encryption/decryption and encapsulation of packets. [11].

Due to complexity of encryption and decryption process in cryptographic algorithms associated with VPN tunnels the performance would become a bottleneck. For this reason, dedicated hardware is proposed as a way to maximize the throughput while the latency gets minimized. In addition, BGP/MPLS VPN technology is being used as a novel VPN approach. In this approach the benefits of MPLS (Multiprotocol Label Switching) technology is combined with security aspects of a VPN. This feature allows service providers to offer VPN service by using their MPLS network for secure delivery of different type of traffic such as voice [8].

3 Methodology

This section describes the simulation tool and processes employed in this research. OPNET IT GURU is used as the simulator in this research. The main research objectives to achieve in this simulation are:

i. examine the impact of parameters arising from the network and network design on performance of a SIP connection running over IP in terms of end-to-end delay, delay variation and call setup time;

ii. compare the results in i) with a SIP connection running over IP VPN tunnel;

iii. examine the above for several possible scenarios entailing different network designs that include placement of SIP components.

The parameters arising from the network entail a broad range and number of variables and factors, such as network throughput, link capacity and bandwidth, congestion issues, traffic intensity, queuing issue, choice of routing protocols, type of the service provided by the network (e.g. Best effort vs. Guaranteed service), diversity and nature of applications running in the network (Multimedia vs. data, etc.). This research is limited to achieving the simulation objectives stated earlier. Also, to make sure that the primary issue under examination is simulated and experimented with as much control as possible, we kept the possible design scenarios as simple and neat as possible. This would eliminate other possible considerations tied to more complex network design scenarios. Note: EIGRP was employed as the routing protocol in the study.

We created several possible scenarios resembling a simple SIP-based connection over IP backbone with and without VPN. Two IP-phones (using SIP as the signaling protocol) and SIP Proxy servers (one or two - depending on the possible design scenarios) are configured. Proxy server(s) served as the primary source to forward SIP requests and responses. Each scenario was simulated once without VPN and once with VPN in place. For VPN connection, the packets sent from one IP phone to the other one were sent through a VPN tunnel that was established between the routers connected to each IP phone (i.e. ROUTERA and ROUTERB). Figure3 illustrates one of the scenario (named as scenarios 2) used during the experiment.
Figure 4 illustrates the traffic flow of a ping echo-request and each-response from PHONEA to SIPPROXY. Note: “Compulsory” VPN mode was employed during the experiments, meaning that packets destined to the IP phones would always pass through to the VPN tunnel, no matter if the routing information knew a shorted path to the destination.

Figure 5 shows another scenario (Scenario 3) used in the experiment.

Figure 5. Scenario 3

G.711 is used as the voice codec and SIP as signaling protocol during the experiment. Figure 6 and 7 details the voice codec and voice application definition used in this study.

In order to generate the VOIP traffic, a VOIP Profile is defined and assigned to PHONEA, which would make constant 2-minute phone calls to PHONEB during the simulation, which ran for 1 hour. Figure 8 shows the configured profile on PHONEA. PhoneB is configured to support the pre-defined VOIP application as presented in Figure 7, but it is not configured to generate any phone call. This configuration made the controlling and track of VOIP call flows and simulation simpler for later analysis.

Figure 6. Voice Codec definition

Figure 7. Voice Application Definition

Figure 8. VOIP Profile

4 Results

Results of the experiments to achieve the objective, as discussed in the previous section, is summarized in this section.

4.1 Call Setup Time

Findings of this study suggest that VPN will not always lead to higher call setup time, which is different from the findings reported in prior research (e.g. [3]). This research assumed parameters arising from network and network design, will play a role in call setup time of a SIP-based communication. In scenario 2 (fig 3) the IP phones are configured to send the call setup requests to a SIP Proxy server, which is located in the same network segment as PhoneB. In scenario 3 (fig 5) the IP phones are configured to send the call setup requests to their designated SIP Proxy servers, which reside in the same network segment as the IP phones are located. For these two scenarios the VPN show no negative impact on the call setup time. In scenario 1 (fig 9) each of the IP phone is configured to send call setup requests to a SIP Proxy server, which are located in the Internet. In scenario 4 (fig 10) each of the IP phone is configured to send call setup requests to their designated SIP Proxy server, which are located in the Internet. For these two scenarios, the VPN resulted in an increase in the call setup time.
Figure 9. Scenario 1

Figure 10. Scenario 4

Figure 11. Call Setup time for Scenario 1 and 4

Figure 12. Call Setup Time for All Scenarios

4.2 End-to-End Delay

One significant performance metric for VOIP quality of service is the end-to-end delay issue. Analysis of the end to end delay result, as illustrated in figure 13 and figure 14, suggest that VPN had no negative impact at all on end to end delay. For scenario 1, where each IP phone is configured to send call setup requests to a SIP Proxy server on the Internet, VPN show reduced end-to-end delay compared to when voice streams are forced to go through VPN Tunnel between PhoneA and PhoneB.

Figure 13. End-to-End Delay for Scenario 1
Comparative analysis of end-to-end delay for all scenarios, as shown in figure 14, suggest that for scenario 3 has the lowest end-to-end delay. It is interesting to note that VPN do not lead to degradation in VOIP quality in terms of end-to-end delay. Also, VPN reduces the end-to-end delay in conjunction with network design issue. With VPN present in Scenario 2, VOIP end-to-end delay is lower than that of Scenario 4 (with or without VPN). If VPN is not present in Scenario 1, end-to-end delay is higher than that of Scenario 4. A quick reference to Voice Quality Measures as suggested by ITU-T (see table 1), suggests that the end-to-end delay results obtained in the experiment are actually good. Therefore, a comparative analysis of end-to-end delay and call setup time indicate that VPN in conjunction with design can actually improve the VOIP performance.

4.3 Delay Variation (Jitter)

This section is allotted to provide a brief review of the Delay Variation results obtained in the simulation experiment. Based on results of figure 15, VPN do not degrade the delay variation.

According to Voice Quality Measures as suggested by ITU-T (see table 1), delay variation results obtained in the experiment are good.

5 Conclusions and Future Works

This paper is intended to simulate, using OPNET IT GURU, the behavior of a SIP-based VOIP connection over IP VPN tunnel. Findings of this study suggest that VPN do not always lead to higher call setup time, as opposed to the findings reported in the prior research (e.g. [3]). Analysis of call setup time suggests that network design plays an evident role in call setup time and that with a proper design (in this case, proper placement of SIP Proxy server) VPN will not bring a negative impact on the call setup time. Results of figure 12 (call setup time) also suggest that using a single SIP Proxy server in the network segment in one of the VPN tunnel end-points is preferred than to having the IP phones send their call setup requests to a single designated SIP Proxy server, which is located in the Internet. Comparative analysis of end-to-end delay for all scenarios also suggest that Scenario 3 show the lowest end to end delay.

Figure 15. Delay Variation

Analysis and findings of this research are different from few related findings e.g. a study by Gouda I. Salama et al., [3] and another study by Muhamad Amin [5] suggested VPN as a factor that increase end-to-end delay. However, findings of this study are aligned with those of Ibrahim S. I et al., [8], where he showed a BGP/MPLS VPN to improve VOIP quality by reducing the end-to-end delay.

Future works may include setting up a network environment scenario in which factors such as network throughput, link capacity and bandwidth, congestion issues, traffic intensity, queuing issue, choice of routing protocols, type of the service provided by the network (e.g. Best effort vs. Guaranteed service), diversity and nature of applications running in the network (Multimedia vs. data) are considered. For instance traffic intensity, which in this study is not set to exceed (or at
least closely reach) the link capacities, can be set to a ratio to cover network congestion and throughput issues. A potentially more realistic scenario could also examine end-to-end delay in a network environment in which data packets are sent alongside voice streams.

6 References


