ENSC 427: Communication Networks

ANALYSIS OF LONG DISTANCE 3-WAY CONFERENCE CALLING WITH VOIP

Spring 2010

Final Project

Group #6:

Gurpal Singh Sandhu (gss7@sfu.ca)

Sasan Naderi (sna14@sfu.ca)

Claret Ramos (ckr@sfu.ca)

www.sfu.ca/~sna14

Table of Contents

Abstract

VoIP is an alternative to circuit-switched networks. It provides a means of communication using IP networks over local and long distances. In recent times, VoIP has been responsible for drastically cutting the cost of long-distance calls, and as such, it has seen an increase in popularity. This project will analyze the performance of long distance 3-way voice conference calling using VoIP. Performance will be evaluated by examining packet loss, end-to-end packet delay, delay jitter, and speech quality during a conference call. We will attempt to test the network by varying voice codecs, background and link loads, which might affect the overall user experience. A voice call can be modeled using the traffic importer tool in OPNET. Then through testing, we can determine the call quality through subjective and quantitative means.

List of Figures

List of Tables

List of Acronyms

IP	Internet Protocols
MOS	Mean Opinion Score
RTP	Real Time Transport Protocol
SIP	Session Initiation Protocol
UDP	User Datagram Protocol
VoIP	Voice over Internet Protocol
PPP DS-1	Point-to-Point Digital Signal-1

1.0 Introduction

It's an understatement to say that VoIP is popular since it has a commanding dominance in the telecommunications industry. The large spike in popularity is due to its cost-effectiveness and flexibility in enterprise and consumer markets. VoIP only requires a broadband connection to complete local and long distance communication over the IP network. It also supports many of the same features included with traditional telephony but at much lower cost to the consumer. Particularly in an enterprise setting, 3-way conferencing over VoIP is very practical since it provides a platform for team meetings without the costs and delays of travelling. In terms of quality, IP telephony includes many voice compression algorithms known as voice codecs that can be used in diverse networks. The varying voice codecs can yield voice quality that is lower, the same or greater than typical circuit switched telephony depending on the network's delay and bandwidth availability.

It is therefore our goal to determine the overall performance of VoIP in a simple 3-way long-distance conferencing call for two popular and contrasting voice codecs. We will also be examining the performance with and without the presence of background traffic over our network. The outcome will determine which type of voice codec is better suited for different networks.

1.1 VoIP Overview

VoIP provides an alternative to a circuit-switched telephone network by allowing telephone calls to be made over the Internet from a packet-switched network. Thus, voice communication is established between two parties using the Session Initiation Protocol (SIP). SIP instructs the necessary hand-shaking signals between the two parties that are involved in the call process. As such, SIP can provide the conventional call features such as dial, answer, hold, reject, call forward, and call transfer. SIP hand-shaking signals provide VoIP the ability to establish a three-way conference call as seen in figure 1.

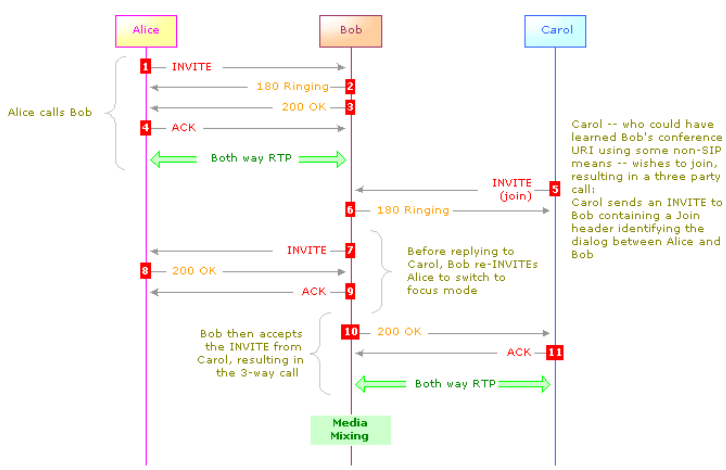

Figure 1: SIP Conference Call Flow

When a call is made, voice gets sampled and digitized for transmission. Voice packets are encoded using a standardized voice codec such as G.711, G.723.1, G.726, G.728, G.729, and more [7]. These codec differ in terms of algorithm, bandwidth, modulation, and payload, but they are all design to compress digital audio signal containing speech data. Once encoded, voice packets are transmitted on RTP/UDP/IP as shown in table 1.

Table 1: Voice Data Packet Format

Voice Data (20-150 Bytes)	RTP (8 Bytes)	UDP (12 Bytes)	IP (20 Bytes)

When the receiver gets these packets, they are de-packetized and data is reconstructed and delivered to the client [7].

Various factors affect the voice quality when using VoIP such as delay, packet loss, and jitter. When analyzing and evaluating the performance of conference calling using VoIP, these factors must be taken into account. Furthermore, the voice codec G.711 is widely used as the industry standard for digital telephony. As a result, the G.711 voice codec will be used as the base codec when testing the VoIP performance on a network during a conference call.

2.0 Discussion

2.1 VoIP Performance Evaluation

The performance of a VoIP call is affected by factors involving jitter, packet loss, and end-to-end delay, which all consequently affect the voice quality. These factors vary in terms of the burden VoIP has on the network traffic, and additional burden that other network applications have on the traffic. To investigate the performance of a conference call using VoIP, it is essential to analyze the parameters that affect the voice quality and efficiency of VoIP.

2.1.1 Jitter

Jitter describes the variation in delay caused by some deviation or displacement in a periodic signal during a transmission [3]. Jitter is caused by poor quality links between nodes and/or traffic congestion on the network. However, a receiver will have a jitter buffer to compensate or minimize this delay variation. The jitter buffer queues in arriving packets, which allows for a continuous stream of data to be transmitted over a network [3]. Unfortunately, if the packet's arrival rate is longer than the jitter buffer length, then packets get discarded[3]. Thus, jitter not only affects the periodic signal, but it subsequently affects packet loss, end-to-end delay and therefore voice quality. As a result, only an acceptable amount of jitter is allowed on a network, and this is typically less than 60 ms [3]. To validate the performance of a VoIP conference call, jitter needs to be taken into account for evaluating voice quality and network burden.

2.1.2 End-to-End Delay

The end-to-end delay is determined by the time it takes for a packet to arrive from the source node to the destination node. Factors such as jitter, packet generation, and the path taken can affect the end-to-end delay of a network [9]. As previously discussed, jitter adds to the delay of the packets arrival time, because of the deviation that is encountered in the periodic signal. The time it takes for the transmitter to generate packets and the receiver to re-assemble the data also contributes to the end-to-end delay of the network. More importantly, the path taken when transmitting packets ultimately defines the time it takes for packets to travel across a network. End-to-end delay is therefore calculated by applying the following formula, which is already done in OPNET[9]:

$$d_{end\text{-}end} = N[\ dtrans+dprop+dproc]\ where$$

$$d_{end\text{-}end}= \text{end-to-end delay}$$

$$d_{trans}= \text{transmission delay}$$

$$d_{prop}= \text{propagation delay}$$

$$d_{proc}= \text{processing delay}$$

$$N = path$$

2.1.3 Packet Loss

The occurrence of packet loss is the result of packets failing to arrive at their destination. Along with jitter, other causes of packet loss can be attributed from signal degradation, corrupted packets, and channel congestion [10]. As a result, it is essential to monitor the quality of service at the cost of packet loss, especially when there is added burden on the network traffic. In terms of VoIP conference calling, an acceptable amount of packet loss can be justified by the voice quality of codec when evaluating the change in the Mean Opinion Score (MOS).

2.1.4 MOS

The Mean Opinion Score (MOS) provides a numerical indication of the perceived quality from a voice codec during and after the transmission and compression of voice data [11]. Factors that can affect MOS include packet loss, jitter, and end-to-end delay. Table 2 below describes the MOS scale that is used to rate the overall speech quality of a voice codec:

Table 2: Mean Opinion Score (MOS) Quality Values

MOS	Quality	Impairment
5	Excellent	Imperceptible
4	Good	Perceptible but not annoying
3	Fair	Slightly annoying
2	Poor	Annoying
1	Bad	Very Annoying

2.2 Network Topology

The VoIP network topology consists of three residential subnets located in Vancouver, Saskatoon, and Ottawa. Each Subnet is connected to an IP cloud using PPP DS1 duplex links. The PPP DS1 duplex links is common for residential homes, because it is a cost-effective T-1 carrier solution with an acceptable bit-rate of 1.544 Mbps. The IP cloud represents the Internet, which is a packet-switch network that will allow for 3-way VoIP conference calling. Figure 2 below shows the VoIP network topology that is implemented to establish 3-way conference calls.

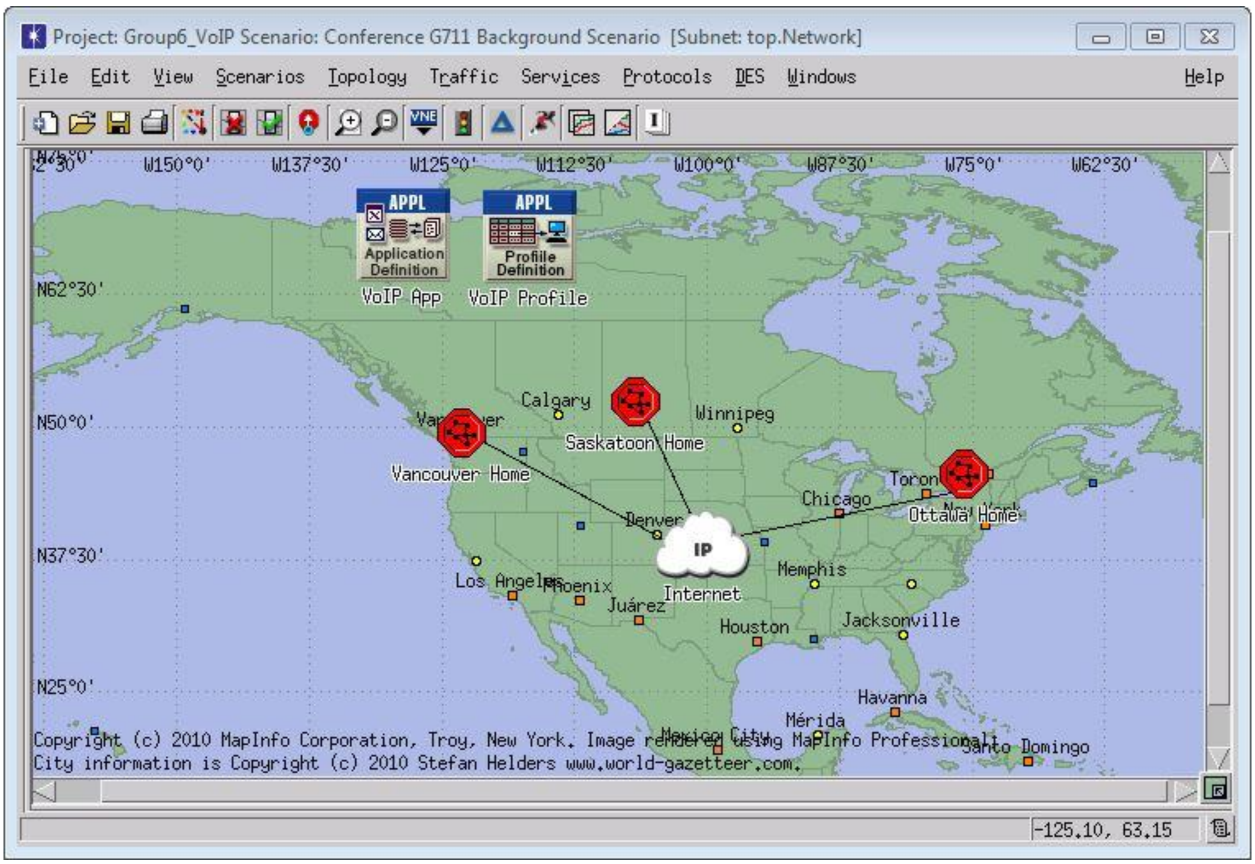

Figure 2: VoIP 3-Way Conference Call Network Topology

Inside each subnet, there exist multiple Ethernet workstations that are connected to a Cisco C4000 router. The Cisco C4000 router represents the common household router used in residential homes. Each Ethernet workstation is defined to be a typical residential network connected device such as a laptop, desktop, game console, and more importantly a VoIP phone. Each subnet may vary in terms of network connected devices, but they all have VoIP connected phone as illustrated in figures 3, 4 and 5.

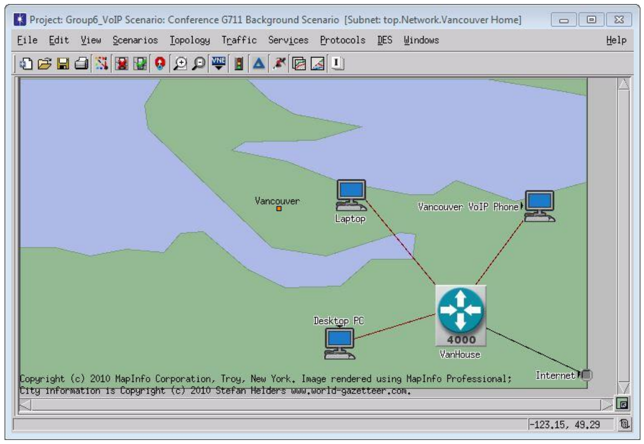

Figure 3: Vancouver Residential Home Subnet

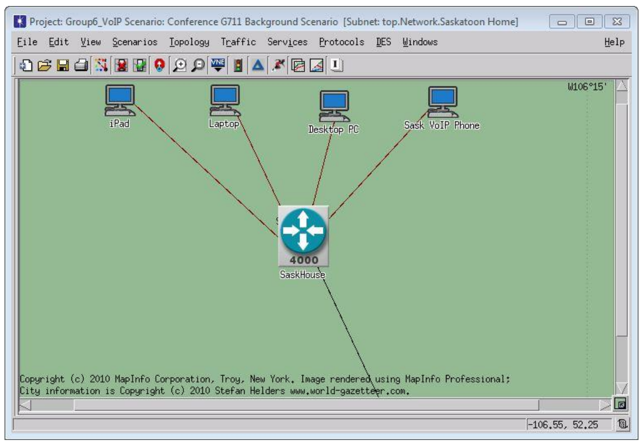

Figure 4: Saskatoon Residential Home Subnet

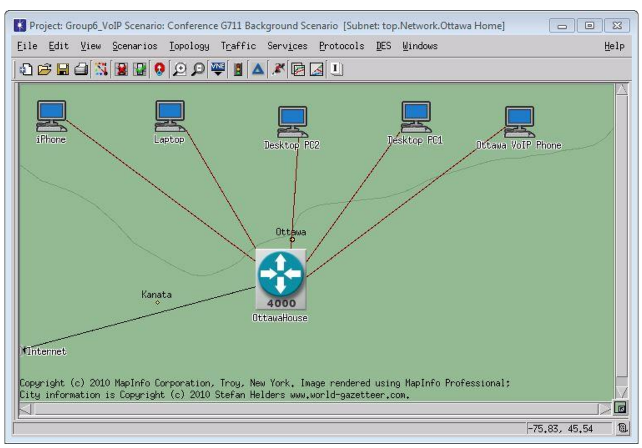

Figure 5: Ottawa Residential Home Subnet

To successfully simulate a 3-Way VoIP conference call, we need to define the following: Application Configuration, Profile Definition, and Background Traffic Definition.

2.3 Application Configuration

The application definition was configured to support the predefined VoIP application. In figure 6, the user can customize the VoIP application by manipulating attributes such as types of services and encoder scheme to fit their application.

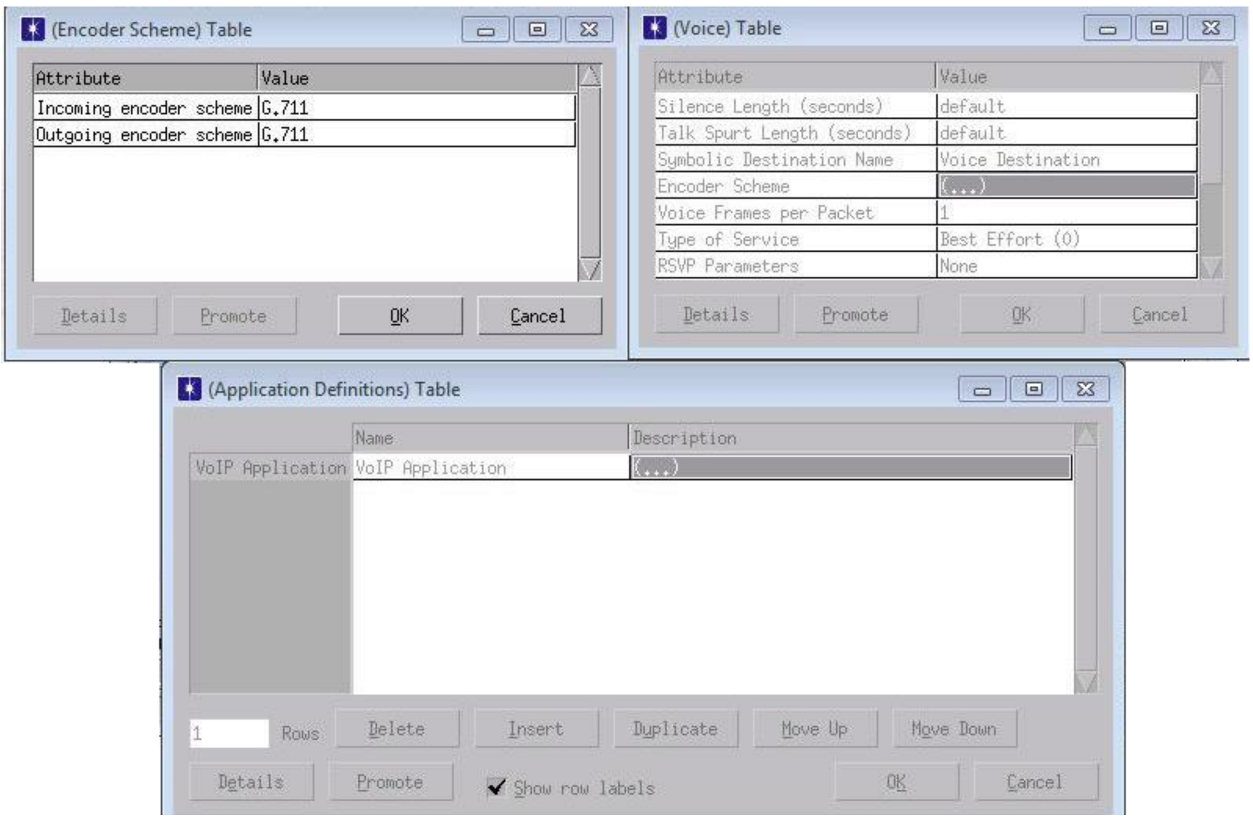

Figure 6: VoIP Application Definition

In the case of conferencing calling, the best effort service was implemented to simulate IP telephony. Additionally, the encoder scheme alternated between G.711 and G723.1 to compare and contrast the performance of VoIP conference calling with and without background load. The voice frame per packet was defined as “1” to simulate the typical audio sample of 32 byte payload. The length of talk and silence time used in a call was left with the default exponential distribution since it already replicated a typical conversational scenario. Once the application definition has been defined, the profile definition can be used to apply the service to the respective network device.

2.4 Profile Definition

The profile definition is built on top of the VoIP application where it specifies which Ethernet workstation will support VoIP services. Since OPNET only supports P2P or Client-Server relationship for VoIP application, three profile definitions were created as shown in figure 7.

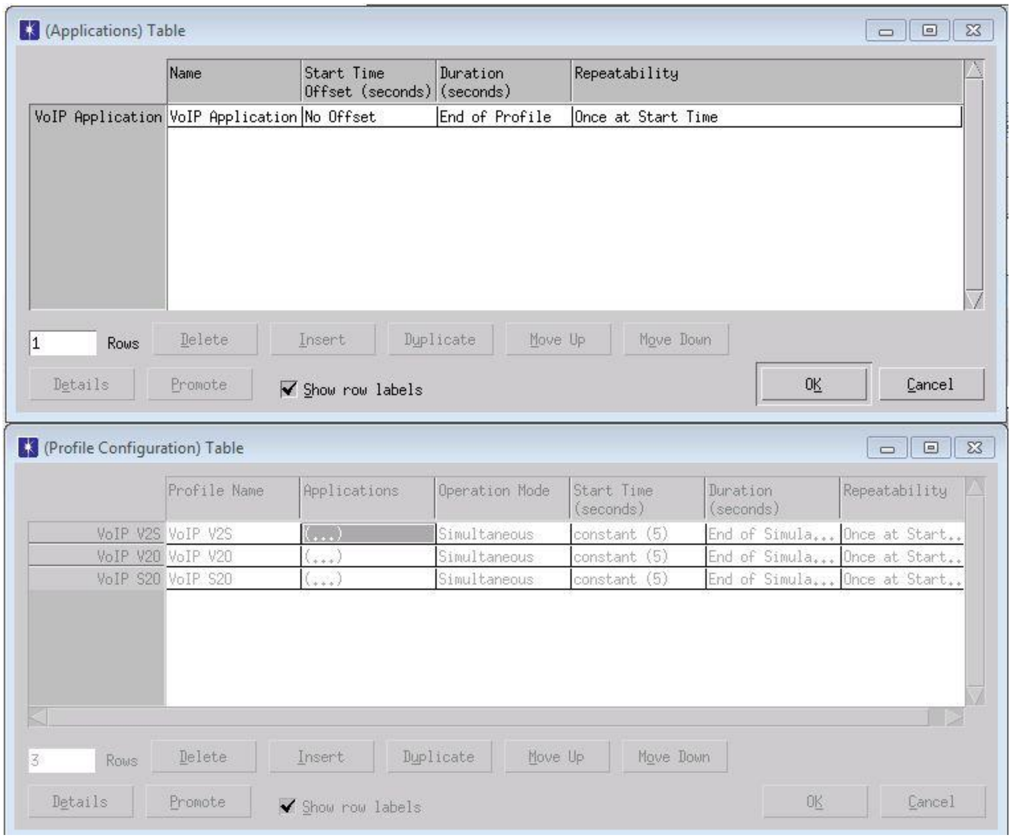

Figure 7: Profile Definition

Each profile represents a connection from one VoIP phone to another, and therefore each VoIP phone supported 2 profile services to simulate the conference call. The profile services of each VoIP phone will start simultaneously after 5 seconds into the simulation in order to establish a conference call. This setup is illustrated in figure 8.

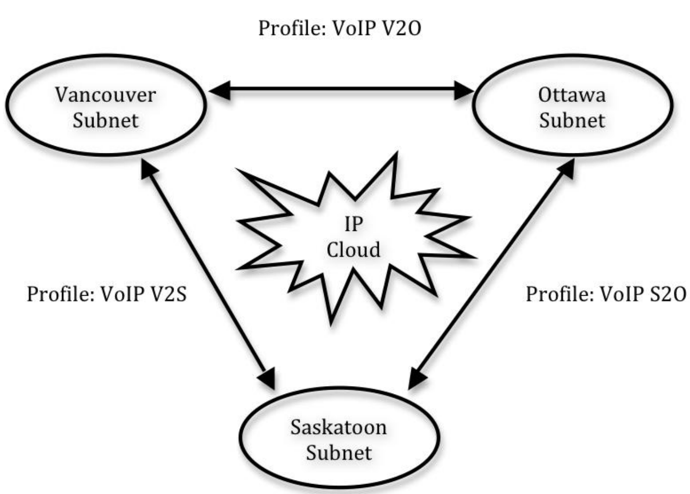

Figure 8: VoIP Profile Services for Conference Calling

2.5 Background Traffic Definition

To simulate the effect that background load will have on VoIP performance, we will need to model it as an implicitly-defined traffic load on the PPP DS1 links. Again, these links are used in connecting each residential subnet to the 32-bit IP Internet cloud. The traffic load profile was simply named “background_load” and was defined in terms of link utilization of DS1's total link capacity of 1.544 Mbps. For example, the first 25 seconds of the profile has 20% of the link utilized as background load which is exactly 308,800 bps. Figure 9 shows the background load profile in its entirety.

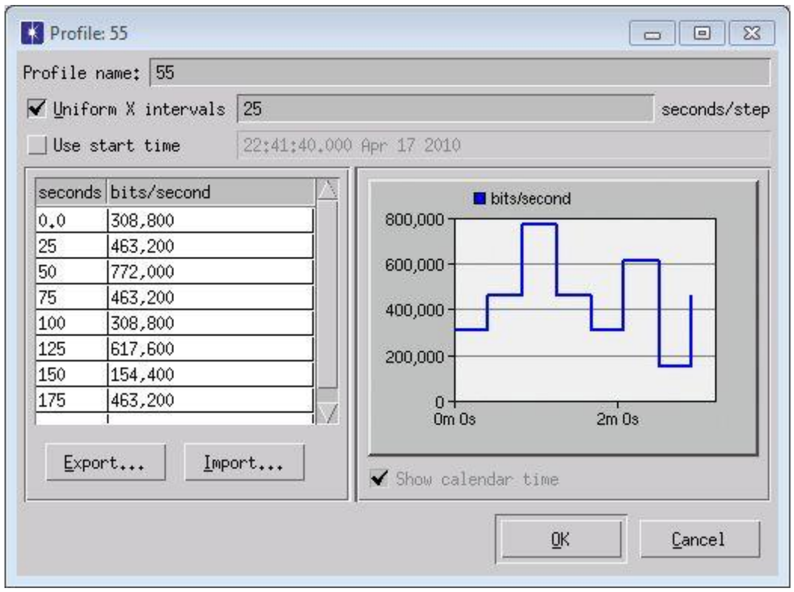

Figure 9: Background Load Profile Definition

After defining and saving the background load profile, all of the PPP DS1 links had their attributes configured to support it in both the ingress and egress directions. Although it may be obvious, only the scenarios, which are simulating a background load had their PPP DS1 links configured for the background load profile.

3.0 Scenarios

There were four scenarios created to compare and contrast the performance of VoIP conference calling in terms of jitter, packet loss, end-to-end delay, and voice quality. The following table describes the scenarios that were implemented to evaluate and analyze the performance of VoIP conference calling:

Table 3: VoIP Conference Call Scenarios

Scenarios	G.711	G.723.1	No Load	Load
Scenario 1	✓		✓	
Scenario 2		✓	✓	
Scenario 3	✓			✓

Scenario 4		✓		✓

4.0 Results

The following section describes the results of simulating all of the scenarios. It also explains why our simulations arrived at those results.

4.1 End-to-End Delay

From figure 10 it is clear that the end-to-end delay of the G.723.1 codec is much greater than the G.711 codec when there is no background traffic present. The same thing can also be said for when background traffic is present in figure 11. This result reconfirms the fact that the G.711 codec greatly outperforms the G.723.1 codec in a situation with heavy background traffic.

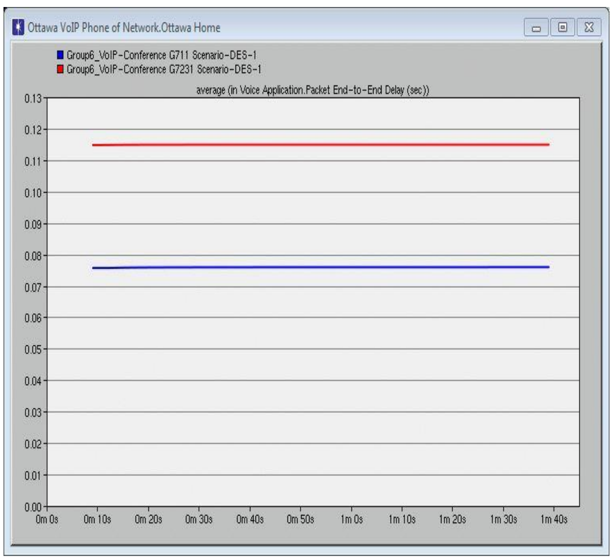

Figure 10: End-to-End Delay of Both Codecs Without Background Load

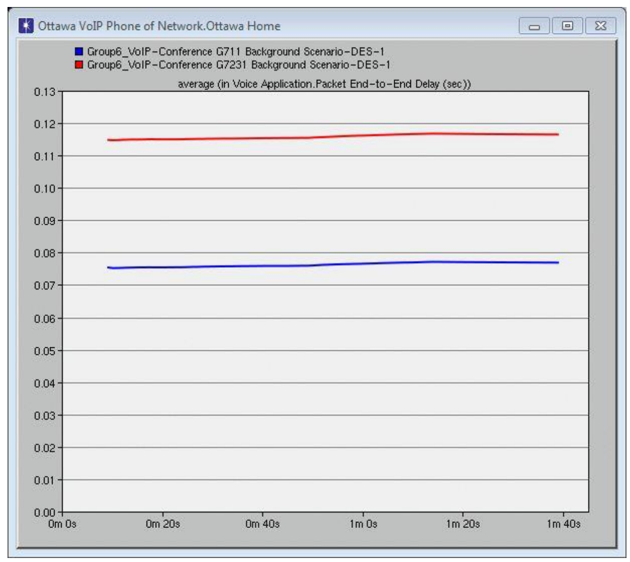

Figure 11: End-to-End Delay of Both Codecs With Background Loads

The following comparison determines if there is any difference between the G723.1 and G.711 codecs in a situation where there is a considerable amount of background traffic. As it is evident in both figures 12 and 13, the codecs only marginally perform better when there is no background traffic present in the network. However that performance difference is negligible.

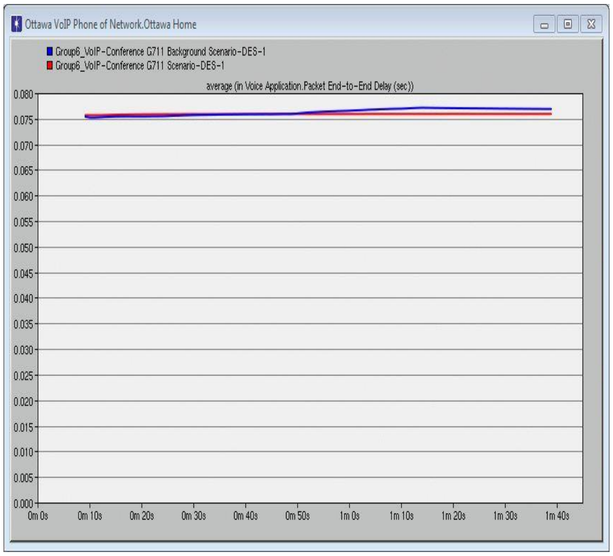

Figure 12: End-to-End Delay of G.711 Codec

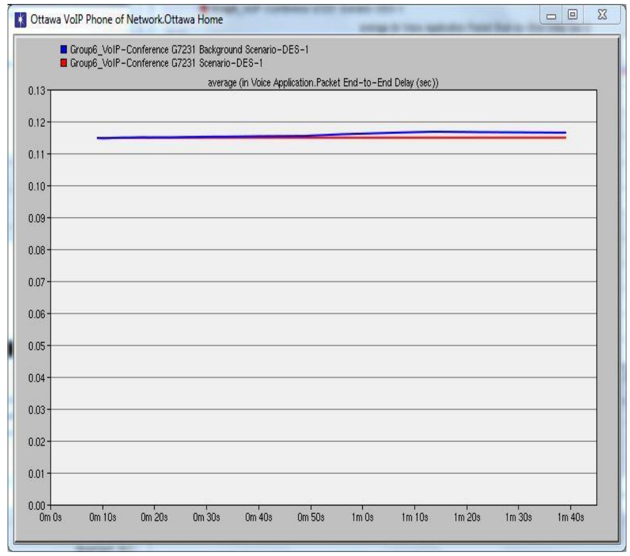

Figure 13: End-to-End Delay of G.723.1 Codec

4.2 Packet Loss

The packet loss in our network simulation displays that the G.711 codec has a great packet discard ratio than the inferior G.723.1 codec. Both of the scenarios were modeled with a background load. Therefore in a situation where a large amount of background traffic is present, the G.723.1 codec outperforms the G.711 codec in terms of packet loss as seen in figure 14.

Intuitively, the packet loss for both codecs in a simulation without background traffic yields very little packet loss.

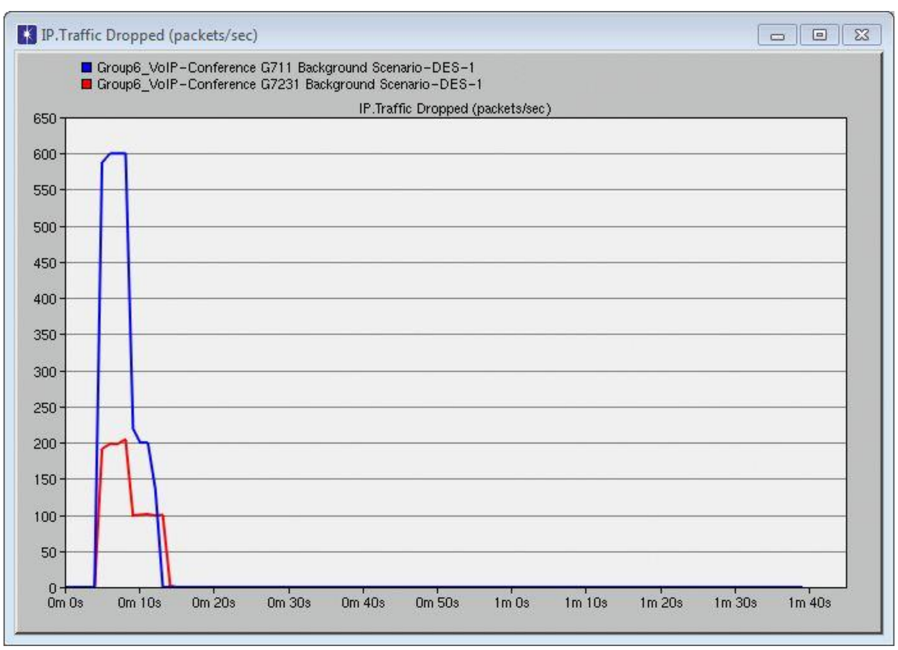

Figure 14: Packet Loss

4.3 Jitter

As evident from figure 15, the jitter between the G.711 and G.723.1 codecs without background traffic is minimal and almost non-existent. However, in figure 16, the jitter is clearly simulated and dominated by the G.711 codec which yields a much lower jitter value. This reaffirms what we know to be true about the superior G.711 codec and proves that in a scenario with background traffic, the G.711 codec is desirable.

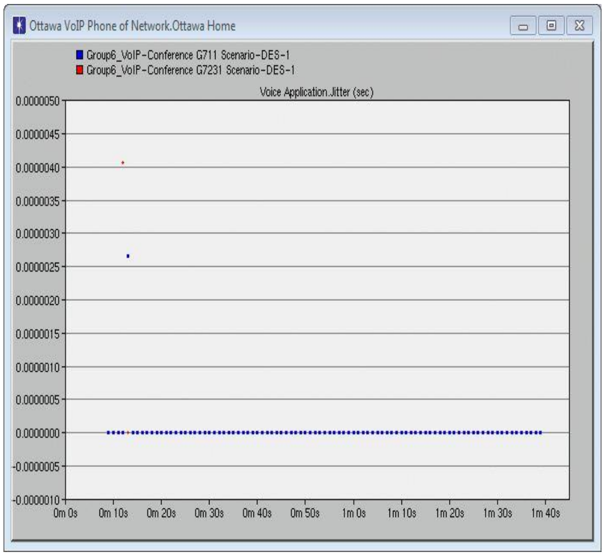

Figure 15: Jitter Without Background Load

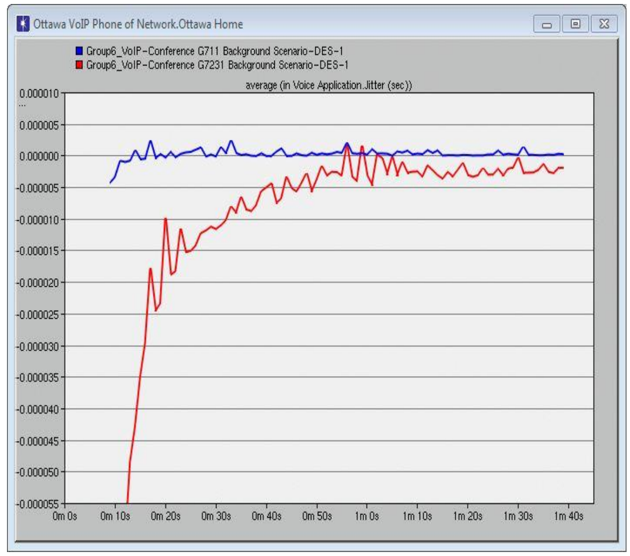

Figure 16: Jitter With Background Load

In the following comparison, the introduction of background traffic clearly causes a large variation in the jitter of both the G.711 and G.723.1 codecs. This makes sense, as the more traffic is present on the network, the more we expect to see variances in the packet arrival time. However, from the conclusion we drew from figures 17 and 18, we can see that the G.711 codec still has a clear advantage over the G.723.1 codec.

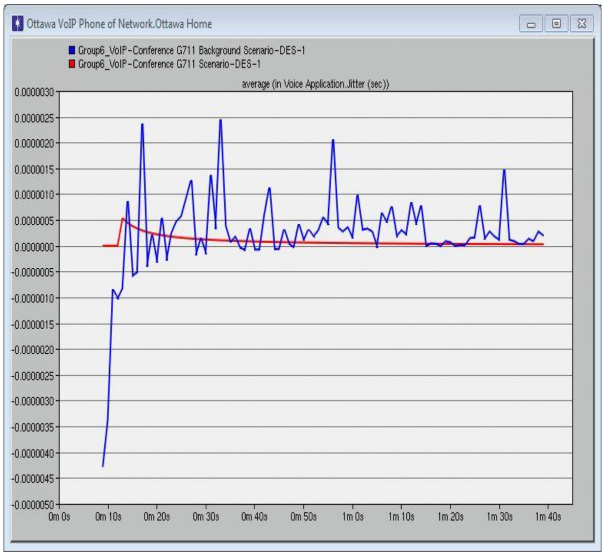

Figure 17: G.711 Jitter With and Without Background Load

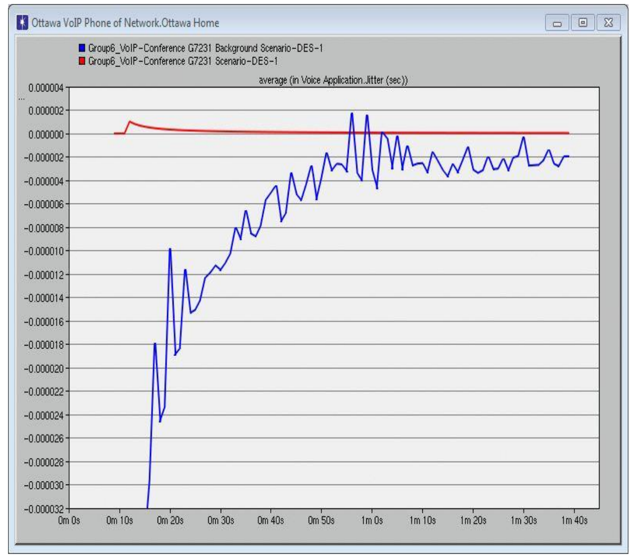

Figure 18: G.723.1 Jitter With and Without Background Load

4.4 Speech Quality (Mean Opinion Score)

In what can be seen as the most drastic evidence in support of the G.711 codec, the Mean Opinion Score (MOS) can be seen in figures 19 and 20. Clearly, the presence of background traffic does not affect the MOS value, therefore we can neglect it. However the G.711 again outperforms the G.723.1 codec with far greater speech quality. There is a difference between 3.6 and 2.5, which suggest that the user experience better quality with G.711 voice codec.

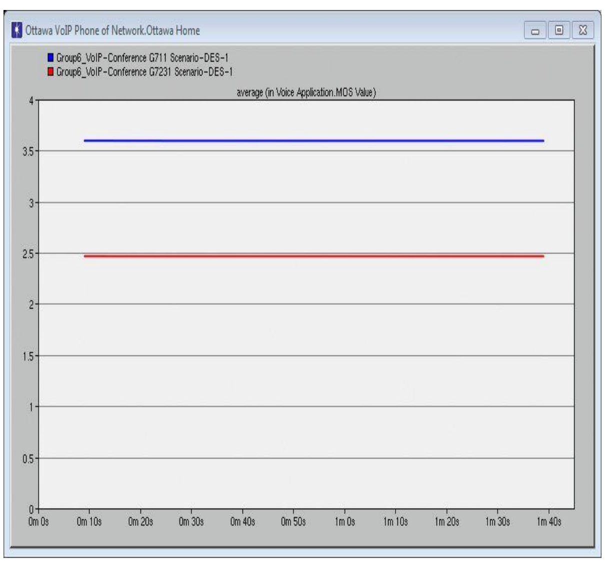

Figure 19: Codec MOS Without Background Load

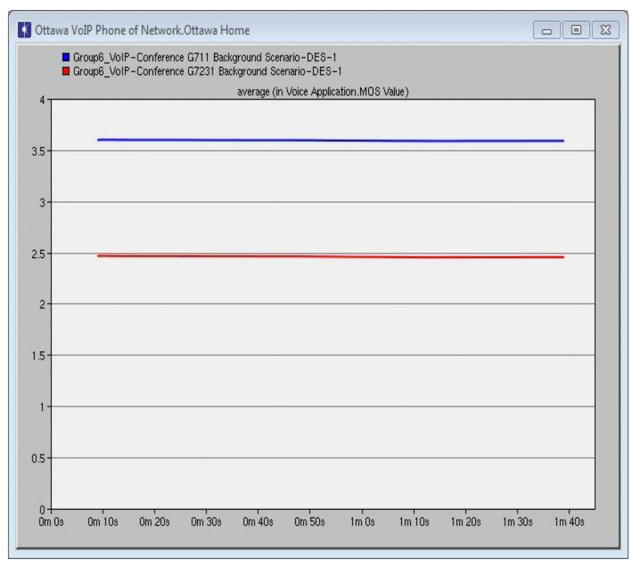

Figure 20: Codec MOS With Background Load

The following comparison considers both the G.711 and G.723.1 codecs with and without a background load. As evident in figures 21 and 22, there is very minimal impact on the speech quality of a VoIP call when background traffic is present.

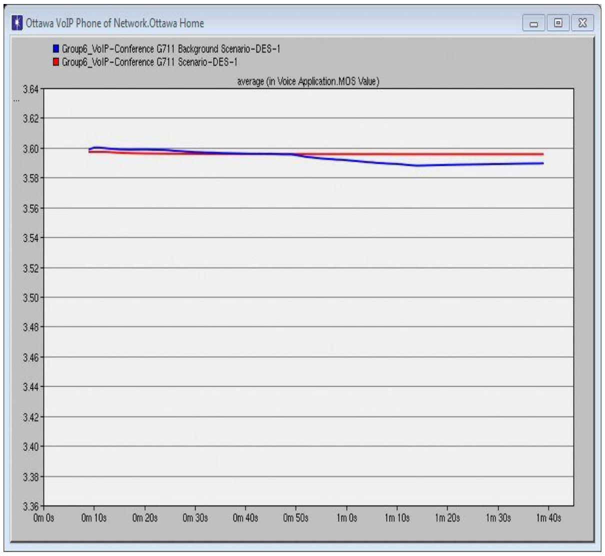

Figure 21: G.711 MOS With and Without Background Load

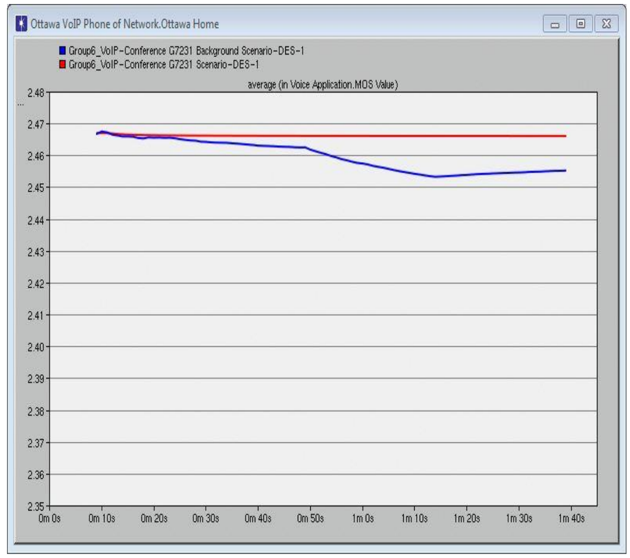

Figure 22: G.723.1 MOS With and Without Background Load

5.0 Conclusion

In our OPNET simulation, we decided on examining the overall performance of VoIP in a simple 3-way long-distance conferencing call for G.711 and G.723.1 voice codecs. We also examined the performance with and without the presence of background traffic over our network to determine which type of voice codec is better suited for different networks.

In our results we uncovered that the presence of background traffic on the network had minimal effects on the packet end-to-end delay and the voice speech quality. However, the jitter and packet loss were significantly affected by background traffic. The G.711 codec greatly outperformed the G.723.1 codec in all of the tested categories except packet loss. Although this could be seen as an issue, its strong dominance in all of the other categories suggested that the G.711 codec is far superior and hence more desirable in VoIP applications. This explains why it is one of the most widely used codecs in the VoIP industry at the moment.

6.0 References

[1] H. Schulzrinne, S. Casner, R. Frederick, and V. Jacobson, “RTP: A transport protocol for real-time applications,” IETF RFC 3550, Jul. 2003.

[2] J. Rosenberg, H. Schulzrinne, G. Camarillo, A. Johnston, J. Peterson, R. Sparks, M. Handley, and E. Schooler, “SIP: Session Initiation Protocol,” IETF RFC 3261, Jun. 2002.

[3] B. Goode, “Voice Over Internet Protocol VoIP,” Proc. IEEE, vol. 90, pp. 1495- 1516, Sept. 2002.

[4] S. Garg, and M. Kappes, “Can I add a VoIP call?” IEEE, 2003.

[5] P. Smith, P. Kabal, and R. Rabipour, “Speaker Selection For Tandem-Free Operation VoIP Conference”, Proc. IEEE Workshop Speech Coding, pp. 120- 122, Oct. 2002.

[6] C. Boutremans, G. Iannaccone, and C. Diot, “Impact of link failures on VoIP performance,” Proc. NOSSDAV, May 2002.

[7] R. G. Cole and J. H. Rosenbluth, “Voice over IP performance monitoring,” Computer Communication Review, vol. 31, no. 2, pp. 9–24, April 2001.

[8] L. Balliach, "Voice over IP," July 2003. Available: http://www.opalsoft.net/qos/VoIP.htm. [Accessed: March 23 2010]

[9] End-to-End Delay. Available: http://en.wikipedia.org/wiki/End-to-end_delay

[10] Packet Loss. Available: http://en.wikipedia.org/wiki/Packet_loss

[11] MOS. Available: http://www.rhyshaden.com/voice.htm