Determining the Issues to Consider When Deploying VoIP onto a Small Enterprise Network

Quazi Abidur Rahman, Adam Abubakar
Department of Information and Computer Science
King Fahd University of Petroleum and Minerals, Dhahran, Saudi Arabia
E-mail: {qabidn, abubakar}@ccse.kfupm.edu.sa

Shaik Sirajuddin, Syed Shamsul Islam, Mohammad Abdur Razzaque
Department of Computer Engineering
King Fahd University of Petroleum and Minerals, Dhahran, Saudi Arabia
E-Mail: {siraj, sislam123}@ccse.kfupm.edu.sa, sk_raz@yahoo.com

Abstract

In this paper, we report some lessons learnt when deploying VoIP onto a small enterprise network. While these lessons resulted from a small network environment, some of the issues could have a more general applicability. Some theoretical analyses were first conducted, followed by some experimental work. An experiment was conducted for about 24 hours in order to capture the network measurement of a small LAN. The results obtained from the measurement were used in analyzing the whole network based on queuing theory. This provides ground for making judicious modifications to the network in order to support the deployment of VoIP. The modified LAN was simulated using OPNET for a period of 3 minutes after which the router began to drop some packets. With this, the simulation was terminated and results of the simulation were collected. The results obtained from queuing analysis agree with those obtained at the end of the simulation.

1. Introduction

The VoIP concept is simply the ability to send voice To have data and voice travel over the same network is a wonderful idea and it is a technology that has been around for sometime. VoIP offers quite a number of benefits which include cost saving and simplification. However, some customers are wary of this technology partly because they are not very conversant with the whole concept. Moreover, they have a switch network that works fine for voice transmission. In this paper we discussed the modifications that are required and things to consider when deploying VoIP onto a data network.

Like all other voice communications, VoIP needs two types of protocols:

- Protocol for sending the conversation data in the IP medium and
- Protocol for the signaling

For sending the conversation data in the IP medium RTP/RTCP (Real Time Protocol/Real Time Control Protocol) protocol is used over UDP. RTP is responsible to control the voice packet and voice quality. And RTCP is used for exchanging messages between session users regarding the quality of session like lost RTP packets, delay etc. Signaling protocol is needed for Call setup, Monitoring call progress and Call release. The protocols available for this purpose are:

1) IETF (Internet Eng. Task Force): SIP and S/MGCP
2) ITU-T (International Telecom. Union): H.323
3) MEGACO/H.248 has developed jointly by IETF and ITU.

The following are some of the assumptions we made in this research.

a) All the voice session will be point to point. This means, we considered only unicast traffic.
b) G. 711 CODEC is used for voice coding and decoding. It is chosen as it supports [1][5]:
   a. high speed and high bandwidth (64kbps).
   b. the best voice quality, since it does no compression, introduces the least delay
   c. less sensitive than other CODECs to packet loss.
c) Echo cancellers are built-in to CODECs.
d) There is little or no packet loss in the network.

The remaining paper is organized as follows. In section 2, the network under consideration is described. In section 3, the queuing analysis is described along with the assumptions made. Section 4, discusses the throughput analysis of the VoIP network. Section 5 presents the results of the network simulation using OPNET. And finally, we discuss our results and conclude in sections 6 and 7 respectively.

2. Experimental Network Setup and Traffic Measurement

A small enterprise network consisting of a router, two switches, couple of servers and three LANs distributed over three different floors of a building is used for the experiment. The logical diagram of the network along with the type and name of the switches, server and other components are shown in Fig. 1. Three VLANs on switch 1 and another two are employed on switch 2 to isolate broadcast and multicast traffic. VLAN11 includes port P1, P2 and P12. VLAN12 includes P3 and P12. VLAN3 includes P4 and P12. VLAN21 includes P7 and P11 and VLAN22 includes P5, P6, P23, P24 and P11.

The traffic measurements of the network were taken for a period of 24 hours, with an interval of 10 minutes using Getif and SNMP Traffic Grapher (STG). As summarized in[4] from the measurement we found that over utilized interfaces are the port1&2 of router, Port 12 of switch 1 and port 11 of switch2 and average packet size being 1450 bytes.
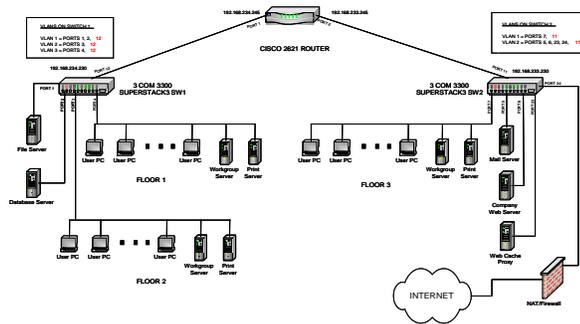

Figure 1. Logical diagram of the existing network.

3. Queuing Analysis:

Total end-to-end delay in the network comprises of some fixed delay and the delays in the queues involved in the network components. Among the fixed delays propagation delay, fixed component delay, CODEC delay (packetization and jitter buffer delay)[3] were considered and serialization delay is ignored. As discussed in [4] we

assumed total fixed delay to be 85ms. Since according to the ITU-T's recommendation G.114, there is a constraint of 150ms of end-to-end delay for optimal voice quality, our tolerable variable (queuing) delay is 65ms [1].

The detail of the queuing analysis is presented in our previous work 0[4]. At first an analytical model for the worst case scenario is developed and then that was generalized over the whole network. There are three different scenarios of transmitting voice from one part of the network to another: (i) intra floor calls (i.e. calls within a floor), (ii) calls from floor 1 to floor 2 or vice versa and (iii) calls from floor 1 or 2 to floor 3 and vice versa. These traffic traversal paths are discussed in detail in 0[4] and it is found that the worst-case scenario is the traffic flown from floor 1 to floor 2 or vice versa. To establish a call between floor 1 and floor 2 traffic will pass through switch11, switch1, router, switch1 and swich12. Traffic has to pass through the router because floor 1 and floor 2 are on two different VLANs. This makes the interface between the router and switch 1 to be utilized twice for the same traffic.

For analytical modeling, the calls were distributed in the following way:(F1-F1): (F2-F2): (F3-F3): (F1-F2): (F1-F3 or F2-F3) = 4: 4: 4: 2: 1 .This means that, for the delay calculation seven calls were initially considered: 4 of the calls within a floor, 2 between floor 1 and 2 and only one call between either floor 1 or floor 2 and floor 3. Thus seven calls were added for each delay calculation to see how it affects the traffic.

We considered 80% utilization of the devices and extra 20% of the capacity is left for the future growth of the network. Accordingly service rate of the router, switches and interfaces are assumed to be 20kbps, 1.04Mpps and 80Mbps respectively. It is also calculated in 0[4] that the voice packet size is 226 bytes requiring bandwidth per call to be 90 Kbps (Approx) for Cisco Router G.711[2].

All the queues (of the interface, switch and router) were considered as M/M/1 queues assuming packets arrive according to a Poisson process, buffer sizes are infinite and the different queues are independent. M/M/1 is preferred since it gives worst case i.e. an analysis based on this assumption gives conservative results. This is nice because tables are available for the M/M/1 case and values can be looked up quickly [3].

Considering a single call from floor1 to floor 2 and background traffic between components equations are derived in [4] and found the total queuing delay for the call to be 58.51μsec making the total end-to-end delay to be 85.0585ms. Therefore the remaining permissible delay of about 65ms can be utilized for adding simultaneous sessions. Now assuming x call within a floor, y between Floor 1 and Floor 2 and z call between either floor 1 or floor 2 and floor 3 out of (x + y + z) simultaneous calls,

equations derived in worst case scenario are generalized over the whole network as shown below:

$$T_{sw11} = \frac{1}{\mu_{ps} - (x+y+z)2\lambda_p} + \frac{1}{\mu_0 - (y+z)\lambda} \qquad (1)$$

$$T_{sw12} = \frac{1}{\mu_{ps} - (x+y)2\lambda_p} + \frac{1}{\mu_0 - y*\lambda} \qquad (2)$$

$$T_{router} = \frac{1}{\mu_{pr} - ((y+z)2\lambda_p + \sum \lambda_{rtbgi})} + \frac{1}{\mu_0 - (y\lambda + \sum \lambda_{swlbgo})} \qquad (3)$$

$$T_{sw1} = \frac{1}{\mu_{ps} - ((y+z)2\lambda_p + \sum \lambda_{swlbgi})} + \frac{1}{\mu_0 - ((y+z)\lambda + \sum \lambda_{swlbgo})} \qquad (4)$$

- m = Service rate of the interfaces in bps
- μ_{ps} m = Service rate of the switch in pps
- μ_{pr} m = Service rate of the router in pps
- λ = Arrival rate of voice traffic in bps
- λ_p l = Arrival rate of voice traffic in pps
- $\sum \lambda_{bgi\ sw1}$ = incoming background traffic to switch1 in bps
- $\sum \lambda_{bgo\ sw1}$ = Total outgoing background traffic from switch1 to the router in bps
- $\sum \lambda_{rtbgi}$ = Total incoming background traffic to router in pps
- $\sum \lambda_{rtbgo}$ = Total outgoing background traffic from router to switch1 in pps.

Above equations were solved by a program using C Programming language and sessions of calls (as assumed in call distribution) were added until one of the two conditions were satisfied: (i) the delay becomes more than 65 ms (ii) the traffic coming in to the router exceeds the service rate of the router. The output of the program shows that the total number of sessions that can be supported by the network can be distributed as follows.

Intra floor   = 244
Floor 1 to Floor 2 = 122
Floor 1 to Floor 3 = 61

The interesting point here is that the delay of the voice traffic from floor 1 to floor 2 is 8.7 ms. This proves that the bottleneck of our enterprise network is the router as it has an effective service rate of only 20Kpps (Kilo packets per seconds), whereas the links and switches in the network are good enough to support excellent quality of voice over the network. From Fig. 2, it is observed that delay shoots out after 122 sessions. It indicates that maximum number of simultaneous sessions that can be supported maintaining high QoS is 122.
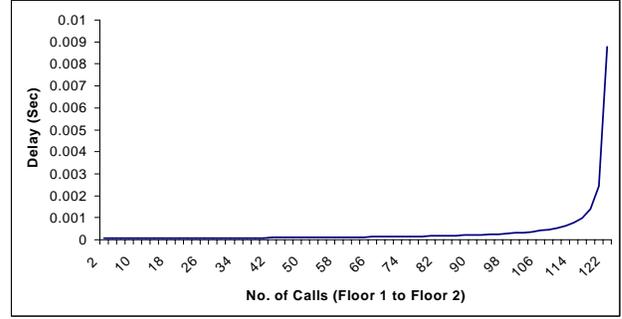

Figure 2. Graph showing increase in delay with an increase in the number of sessions

Increasing number of session requires adding more IP telephone sets to the network. Whenever the ports of the switches at a particular level are filled, another level of local switches can be added to the model.

4.   Throughput Analysis

Throughput is one of the most important parameters for measuring network performance. Throughput of a network is defined as the measurement of processing or handling ability, which measures the amount of data, accepted as input and processed as output by the network. Two parameters associated with throughput are capacity and available bandwidth. Capacity is the maximum throughput that a network path can provide to a traffic flow, when there is no competing traffic load(cross traffic). On the other hand, available bandwidth is the maximum throughput that the path can provide to a flow, given the current cross traffic load in the path.

The link/device with the minimum transmission rate commonly called 'bottleneck' determines the capacity of the path while the link with the minimum spare capacity limits the available bandwidth. This is evident from the following example of packet pair technique for measuring capacity of a three-link path, using the fluid analogy:
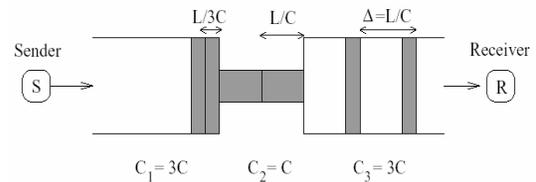

Figure 3. Bottleneck of the network

The packet pair experiment shown above consists of two packets sent back-to-back (i.e. with a spacing that is as short as possible) form the source to the sink. Without any cross traffic in the path, the packet pair will reach the receiver with a spacing or dispersion, that is the transmission delay in the narrow link $\tau_n = L/C$. The receiver can calculate the capacity C from the measured dispersionΔ, as $C = L/\Delta$.

Specifically, if H is the number of the hops(links/device) in the path, C_i is the transmission rate or capacity of link I, C_0 is the transmission rate of the source, then the capacity of the path is

$$C = \min_{i=0...H} C_i \qquad (5)$$

Additionally, if μ_i is the utilization of link i (with $0<=\mu_i<1$ and $\mu_0=0$) over a certain averaging interval, the spare capacity in link I is $C_i(1-\mu_i)$, and so the available bandwidth of the path in the same interval is given by,

$$A = \min_{i=0...H}[C_i(1-\mu_i)] \qquad (6)$$

In our case, Capacity of each Ethernet link, each switch and router is respectively 100Mbps, 1.3Mpps, 25 Kpps. So, obviously the bottleneck of our network is the router. Considering the future growth we are taking the capacity of router as 20kpps. Again from our background traffic measurement it was found that highest utilization of router was 9.5%, which is 8.1 % if the available capacity is considered as 20000 pps. Therefore, the total available bandwidth of our bottleneck can be calculated using above equations as,

   A=20*(1-0.0815)=18.3kpps

For IP telephony voice transmission rate is 50pps (which is 100pps, when voice is interactive). So, maximum number of sessions or simultaneous calls that can be supported 18300/100 or 183.

5. Simulation

The modified LAN was simulated using OPNET for a period of 3 minutes (which took more than a day to complete) after which the router began to drop some packets. With this, the simulation was terminated and results of the simulation were collected. In this section, we discuss how the simulation was conducted.

The queuing analysis conducted earlier shows that the throughput of the network is determined by the router. For this reason we monitor, in the simulation, the traffic sent and traffic received by the router and also the packet end-to-end delay. The statistics of the traffic sent and received by the router is shown in Fig. 4 (the unit of measurement is packet per second).
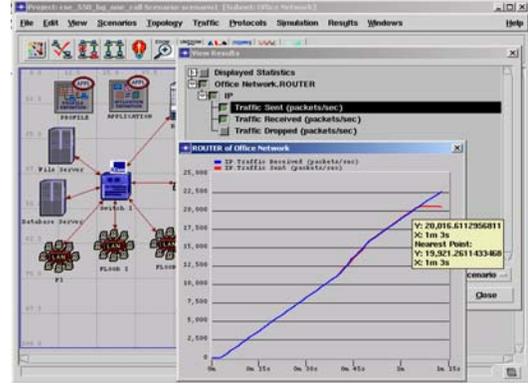

Figure 4. Traffic sent and received by router

By critically analyzing the results of the simulation, we have the following conclusions to make.

The router was saturated after 63 seconds
Since one set of calls is added after every second, the total number of sets of calls = 63

The selection weight of each set is:
Floor1 to Floor 2 = 2
Floor1 to Floor 3 = 1
Intra-floor calls were not considered since they don't pass through the router. Hence, the total number of calls passing through the router is
- Floor 1 to Floor 2 = 126
- Floor 1 to Floor 3 = 63

This is almost the same to the result obtained from the queuing analysis, which is
- Floor 1 to Floor 2 = 122
- Floor 1 to Floor 3 = 61

The end to end delay for voice packet is shown in Fig. 5. From the Fig. 5, we can see that the delay is 0.00877 seconds which is equal to 8.77 ms. This is approximately equals the result obtained from the queuing analysis i.e. 8.7 ms.
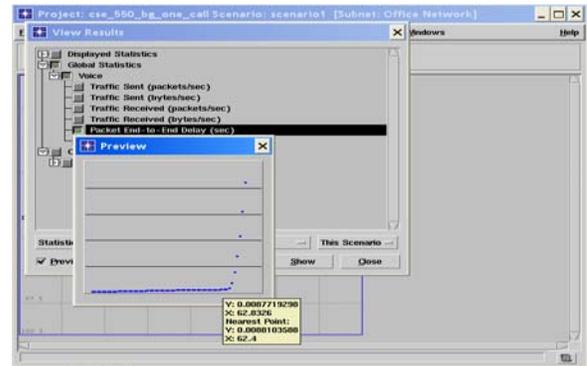

Figure 5. End to end delay

6. Discussion

For our simple network we have found that Delay is not the main determining factor for the maximum number

of sessions obtainable and it is the throughput of the network which decides the maximum sessions possible. We further found that the bottle neck of the network is the router.

We found little difference between the two results obtained from queuing analysis and simulation which might be for our selection of the traffic unit as Mbit/sec in background traffic setup and for the assumption of the average packet size being same all through the background traffic setup.

7.   Conclusion

1. To ensure QoS, the throughput of the network should be carefully taken into consideration together with the delay.
2. Queuing analysis is a good approach in determining the bottleneck of a network by considering the worst case scenario of the network traffic.
3. Most often, one of the nodes in a network may be a bottleneck; if such node is identified then replacing it may be the right decision to take if doing so is cost-effective.

8.   References

[1]   Brans, T., Keyser, T.D., Pollin, S. and Peirs, C., "Voice over IP", 2001.
http://www.esat.kuleuven.ac.be/~h239/reports/2001/voip/verslagvoip.pdf

[2]   Cisco- Voice over IP –Per call Bandwidth Consumption,
http://www.cisilion.com/pdfs/Tech_VOIP_Bandwidth.pdf

[3]   Doing a VoIP Assessment with Vivinet TM Assessor
http://www.netiq.com/products/va/whitepapers.asp

[4]   Quazi A. Rahman. et al, "Deploying VoIP over a Small Enterprise Network," Proceedings of the 8th World Multi Conference on Systemics, Cybernetics and Informatics, July 18-21, 2004 - Orlando, Florida, USA.

[5]   Tanenbaum, A.S., "Computer Networks", 4th edition, Prentice Hall, New Delhi, 2002