Assessing Readiness of IP Networks to Support Desktop Videoconferencing Using OPNET

K. Salah[1]   P. Calyam[2]   M.I. Buhari[3]

Abstract

OPNET is a powerful network design and simulation tool that has gained popularity in industry and academia. However, there exists no known simulation approach on how to deploy a popular real-time network service such as videoconferencing. This paper demonstrates how OPNET can be leveraged to assess the readiness of existing IP networks to support desktop videoconference. To date, OPNET does not have built-in features to support videoconferencing or its deployment. The paper offers remarkable details on how to model and configure OPNET for such a purpose. The paper considers two types of video traffic (viz. fixed and empirical video packet sizes). Empirical video packet sizes are collected from well-known Internet traffic traces. The paper presents in-depth analysis and interpretation of simulation results and shows how to draw proper engineering conclusions.

Keywords: Network Design, Videoconferencing, Simulation, OPNET

1   Introduction

1.1   Background

The deployment of videoconferencing over IP network in both industry and academia has been increasing rapidly. Desktop videoconferencing applications range from internal company communications, educating and training remote employees, to telecommuting. It can eliminate certain travel requirements, thereby cutting costs. Desktop videoconferencing takes advantage of a key workplace tool that is the PC. In the past few years, an H.323 standard was introduced by the ITU, and thus paved the way to the fast growth and deployment of videoconferencing. H.323 is a full suite of protocols developed by ITU to define how real-time multimedia communications, such as videoconferencing, can be exchanged over packet-switched networks [1].

[1] Corresponding Author. Information and Computer Science Department, KFUPM, Dhahran 31261, Saudi Arabia, Email: salah@kfupm.edu.sa

[2] OARnet/The Ohio State University, 1224 Kinnear Road, Columbus OH 43212, USA, Email: pcalyam@oar.net

[3] Information and Computer Science Department, KFUPM, Dhahran 31261,Saudi Arabia, Email: mibuhari@kfupm.edu.sa

It is very advantageous and cost effective to deploy desktop videoconferencing over their existing IP networks. It is easier to run, manage, and maintain. However, one has to keep in mind that IP networks are best-effort networks that were designed for non-real time applications. On the other hand, videoconferencing requires timely packet delivery with low latency, jitter, packet loss, and sufficient bandwidth. To achieve this goal, an efficient deployment of videoconferencing must ensure these real-time traffic requirements can be guaranteed over new or existing IP networks.

Videoconferencing places a high demand on network resources. When deploying such a network service, network architects, managers, planners, designers, and engineers are faced with common strategic, and sometimes challenging, questions. What are the QoS requirements for videoconferencing? How will the new videoconferencing load impact the QoS for currently running network services and applications? Will my existing network support videoconferencing and satisfy the standardized QoS requirements? If so, how many videoconferencing sessions can the network support before upgrading prematurely any part of the existing network hardware?

A number of commercial tools have been developed to address issues related to videoconferencing deployment over data networks. We summarize most popular commercial tools. EURESOM Jupitor II [2] has a provision to test end-to-end Quality of Service (QoS) for Network-QoS-aware applications over IP networks. It considers the relationship between users' perception and network performance. NetIQ's Vivinet Assessor [3] generates RTP streams to mimic VoIP traffic between pairs of hosts and assesses the quality of these synthetic calls. BMC PATROL DashBoard [4] analyzes the impact of multimedia services on the existing network. This tool can quickly identify specific problems on the network that impact application performance. Spirent's IPTV [5] system is a product that includes various features like video infrastructure testing, IPTV video quality testing, firewall and video server load testing. RADVISION [6] offers tightly integrated infrastructure processing components called viaIP, for desktop and meeting room conferencing. Also other companies that provide VVoIP testing are Omegon [7], Lucent VitalSuite[8], and **ViDeNet** [9]. "H.323 Beacon" tool [10] is a open-source tool for assessing performance of desktop videoconferencing sessions using H.323 traffic emulation.

For the most part, these tools use two common approaches in assessing the deployment of videoconferencing into the existing network. One approach is based on first performing network measurements and then predicting the network readiness for supporting videoconferencing. The prediction of the network readiness is based on assessing the health of network elements. The second approach is based on injecting real videoconferencing traffic into existing network and measuring the resulting delay, jitter, and loss.

Other than the cost associated with the commercial tools, none of the commercial tools offers a comprehensive approach for successful VoIP deployment. In particular, none gives any prediction for the total number of calls that can be supported by the network taking into account important design and engineering factors. These factors include VoIP flow and call distribution, future growth capacity, performance thresholds, and impact background traffic. This paper attempts to address those important factors utilizing OPNET simulation.

In this paper, we demonstrate how the popular OPNET simulation tool can be leveraged to assess the readiness of existing data networks to support videoconferencing. To the best of authors' knowledge and to date, OPNET modeler does not have built-in features to support videoconferencing or deployment of real-time services. In the literature, there exists no known simulation approach on how to deploy a popular real-time network service such as videoconferencing. OPNET has gained considerable popularity in academia as it is being offered free of charge to academic institutions. That has given OPNET an edge over DES NS2 in both market place and academia. Another reason to choose OPNET is the fact that OPNET contains a vast amount of models of commercially available network elements, and has various real-life network configuration capabilities. This makes the simulation of real-life network environment close to reality. Other features of OPNET include GUI interface, comprehensive library of network protocols and models, source code for all models, graphical results and statistics, etc.

In previously related work [11], an analytic approach based on the principles of queuing networks was presented to determine approximately the number of video sessions an existing data network can support. In sharp contrast to [11], this paper primarily focuses on showing how to deploy successfully videoconferencing using OPNET modeling and simulation. This paper considers two types of traffic (viz. traffic of fixed video packet sizes as that used for the analytic approach describe in [11] and also traffic of variable video packet sizes measured from well-known traffic traces). This paper discusses in great detail the simulation configuration, setup, and generation of traffic for videoconferencing. Such information can be extremely useful for network researchers and engineers who are interested in deploying videoconferencing. The paper also gives in-depth analysis and interpretations of OPNET simulation results.

1.2 Videoconferencing-Enabled IP Network

Figure 1 illustrates a typical network infrastructure of a small- to medium-sized company residing in a high-rise building with the minimal added videoconferencing components of a H.323 gatekeeper and H.323 workstations [1,12,13]. The *gatekeeper* node handles signaling for establishing, terminating, and authorizing connections of video sessions, as well as imposing maximum bandwidth for each session. H.323 workstations or multimedia PCs have H.323 voice and video software and are equipped with a camera and a microphone. The network is Ethernet-based and has Layer-2 Ethernet switches connected by a router. The router is Cisco 2621, and the switches are 3Com Superstack 3300. All the links are switched Ethernet 100Mbps full duplex. Shared links are never suitable for real-time applications. The network shown is realistic and used as a case study only; however, our work presented in this paper can be adopted easily for larger and general networks by following the same principles, guidelines, and concepts laid out in this paper.
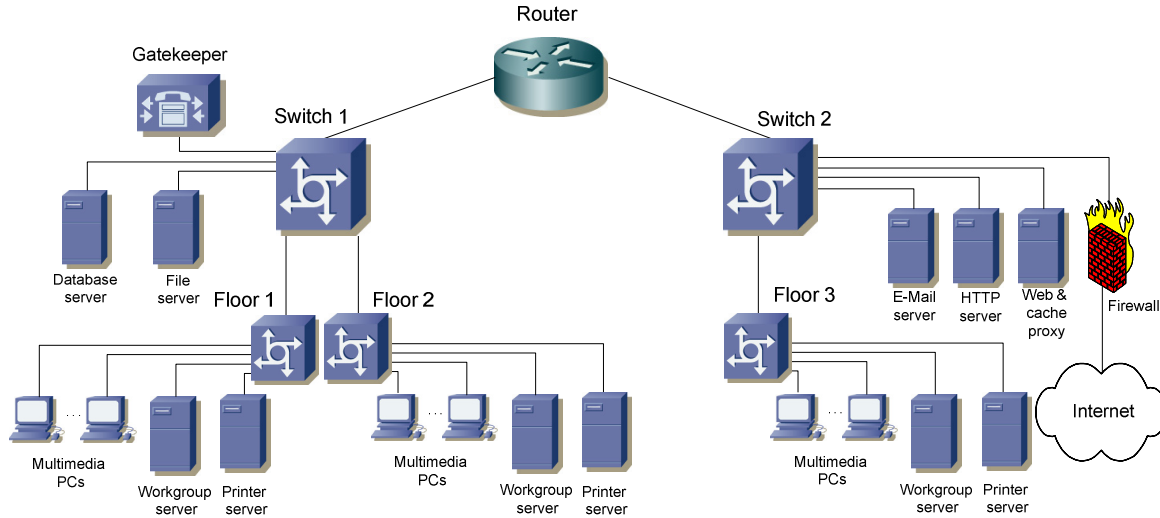
3

Figure 1. Network topology with necessary videoconferencing components

1.3 Traffic Flow and Call Distribution

An important step that plays a factor in determining the number of sessions to be supported is the flow of sessions (or calls) and their distribution. Traffic flow has to do with the path that session travels through. Session distribution has to do with the percentage of sessions to be established within and outside of a floor, building, or department. For our example, we will assume that the generation of sessions is symmetric for all three floors. The intra-floor traffic will constitute 20% of over all traffic, and the other 80% will constitute inter-floor traffic. Such a distribution can be described in a simple probability tree shown in Figure 2.
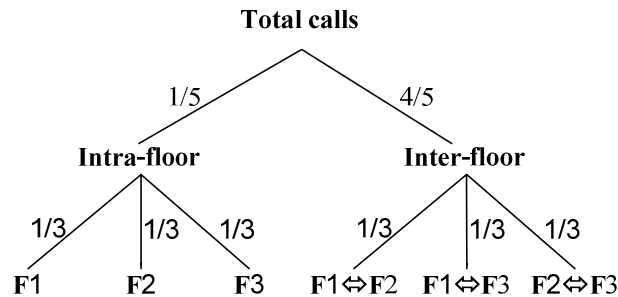

Figure 2. Probability tree describing session distribution

1.4 Additional Considerations

Throughout our work, we assume voice and video calls are symmetric. We also ignore the signaling traffic generated by the *gatekeeper*. We consider the worst-case scenario for videoconferencing traffic. The signaling traffic involving the *gatekeeper* is only generated prior to the establishment of the session and when the session is finished. This traffic is relatively limited and small compared to the actual voice call traffic. In general, the *gatekeeper* generates no signaling traffic throughout the duration of the videoconferencing session for an already established on-going session [14]. In order to allow for future growth, we will consider a 25% growth

factor for all network elements including router, switches, and links. This factor will be taken into account in our simulation study.

The rest of the paper is organized as follows. Section 2 describes important key design issues and requirements for videoconferencing that play a role in assessing the network readiness. These are primarily the bandwidth and delay bounds. Section 3 is the detailed simulation work of OPNET. The section describes in detail the simulation model, configuration, setup, and generation of traffic for videoconferencing. The section also interprets simulation results for both fixed and empirical video packet sizes. Section 4 concludes the study and identifies future work.

2   Delay and Bandwidth Requirements

For deploying a new network service such as desktop videoconferencing, one has to characterize first two important metrics. First is the available bandwidth. Second is the end-to-end delay. The actual number of videoconferencing sessions that the network can sustain and support is bounded by those two metrics. Depending on the network under study, either the available bandwidth or delay can be the key dominant factor in determining the number of sessions that can be supported. In this paper, we assume a point-to-point desktop videoconferencing. Streaming stored video and broadcast video [15] is not considered in this paper and is left for future work.

2.1.1   Bandwidth

A videoconference session consists of two independent bidirectional streams: voice and video [16]. The required bandwidth for a voice call on any one direction, is 50 pps or 90.4 kbps with packet overhead. G.711 codec samples 20ms of voice per packet. Therefore, 50 such packets need to be transmitted per second. Each packet contains 160 voice samples in order to give 8000 samples per second. Each packet is sent in one Ethernet frame. With every packet of size 160 bytes, headers of additional protocol layers are added. These headers include RTP + UDP + IP + Ethernet with preamble of sizes 12 + 8 + 20 + 26, respectively. Therefore, a total of 226 bytes, or 1808 bits, needs to be transmitted 50 times per second, or 90.4 kbps, in one direction. For both directions, the required bandwidth for a single call is 100 pps or 180.8 kbps assuming a symmetric flow.

As opposed to the fixed packet sizes of voice traffic, the packet sizes for video traffic are variable. The variability in the packet sizes of video traffic is dependent on the actual temporal and spatial nature of the video content being encoded. In Ethernet, one video frame is packetized in one Ethernet frame with sizes ranging from 65-1518 bytes. Figure 3 shows a histogram and the corresponding CDF of video packet (or Ethernet frame) size characteristics collected from well-known Internet testbed and reported in [17]. The video packet sizes characteristics correspond to an aggregated representation of video traffic from H.261, H.262 and H.263 video codec streams arising from desktop videoconferencing end-points. The point of interest in the historgram shown in Figure 3 is the proportions of large (>512 bytes) and small (<512 bytes) packet sizes in the video streams. Video codec traffic modeling studies such as [18] and [19] have also demonstrated comparable

proportions of large and small packet sizes in the video streams generated by video codecs such as H.261, H.262, H.263 and H.264 found in popular desktop videoconferencing end-points.
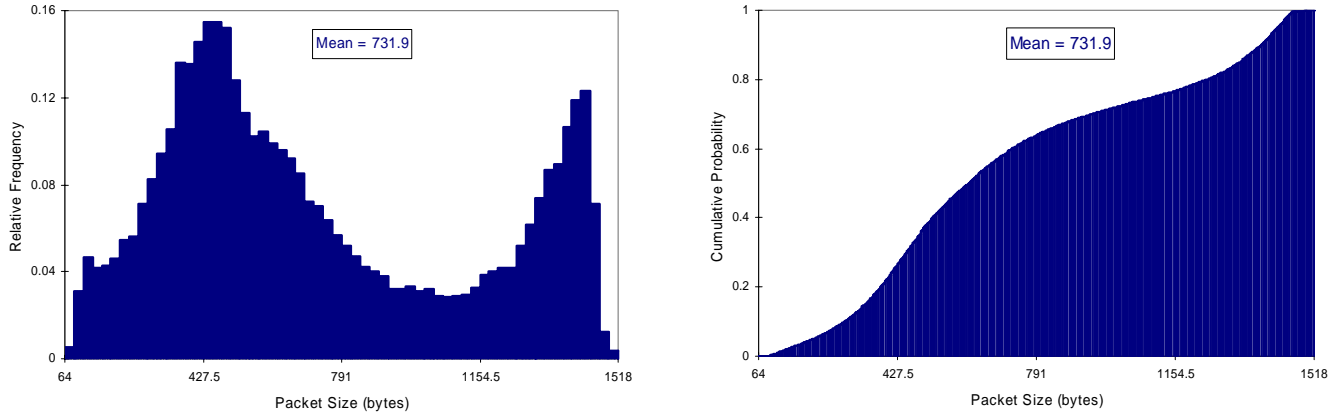

Figure 3. Histogram of empirical video packet sizes and their corresponding CDF

As seen from the CDF in Figure 3, the range of packet sizes above 512 bytes constitute close to 65% of all packet sizes. This percentage was also confirmed in [13]. To be conservative, we will assume (for our simulation scenario of fixed packet sizes) a video frame size of 1344 with a rate of 30 fps. This gives approximately a rate of 320 kbps for pure video traffic. A bandwidth of 320kbps is a multiple of the basic 64kbps communication channel and is an acceptable bandwidth for business desktop videoconferencing with default recommendations of H.261 video codec, CIF video resolution, and H.323 frame rate of 30. [12,15,20,21].

When considering the additional 66 bytes of layer headers, as specified in [22] and happens to be similar to byte overhead for VoIP, the required bandwidth for a video call would be 338.4 kbps. For both directions, the required bandwidth for a single video call is 60 pps or 676.8 kbps assuming a symmetric flow. Hence, for a bidirectional videoconferencing session the required bandwidth is 160 pps or 857.6 kbps.

2.1.2 End-to-End Delay

In order to achieve a natural interactive videoconferencing session, the end-to-end upper bound delay (sometimes termed latency) for a video or voice packet should be kept to minimal. Essentially, such a delay can be broken into at least three contributing components, which are as follows (i) voice sampling or frame grabbing, encoding, compression, and packetization delay at the sender (ii) propagation, transmission and queuing delay in the network and (iii) buffering, decompression, depacketization, decoding, and playback delay at the receiver.

According to recommendations by ITU [23], when delays are less than 150 ms, most interactive applications, both speech and non-speech, will experience essentially transparent interactivity. For voice, the end-to-end delay is sometimes referred to by M2E or Mouth-to-Ear delay [24]. [25] imposed a more conservative delay of

100 ms for both voice and video streams in order to enable natural human interaction. In videoconferencing, there is no separate delay for voice and video streams as both voice and video are synchronized in what is commonly known as "lip-sync". According to real experimental work by [26] the delay difference (termed also skew) between voice and video should be less than 80 ms to allow for natural human interaction and impression.

For our upper bound end-to-end one-way delay of a video or voice packet, we will use 100 ms. This can be broken into 80 ms for the network delay and 20 ms delays for both sender and receiver workstations. This is in accordance with work and recommendations presented by [25] and [26].

3   Simulation Study

For our simulation study, we use MIL3's OPNET Modeler simulation package[4], Release 8.0.C [27]. This section describes in detail simulation model, traffic model, various simulation configurations, as well as the simulation results.
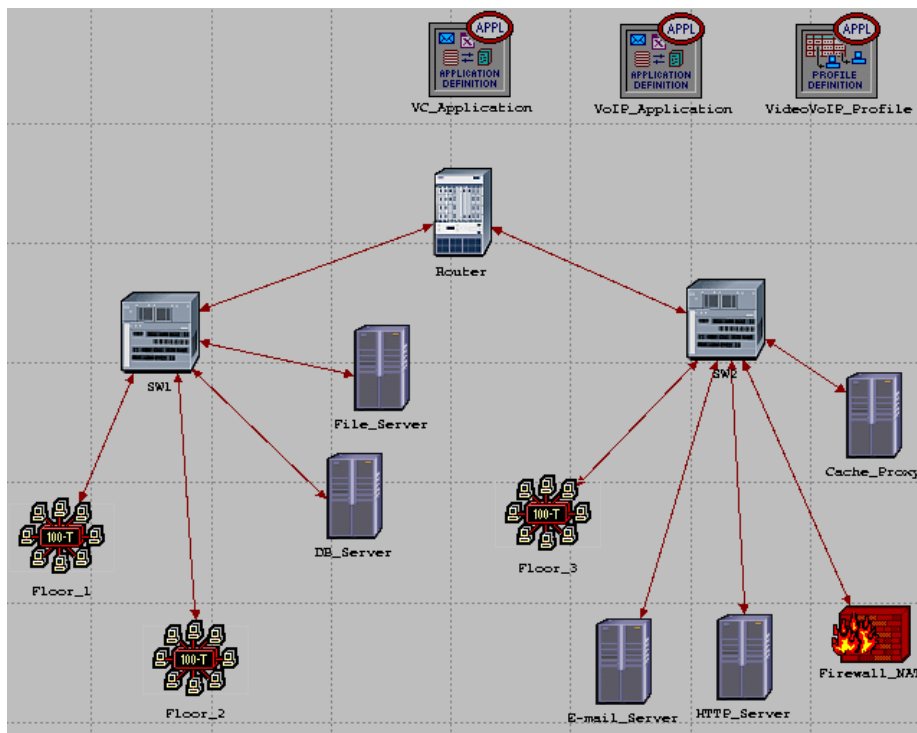

Figure 4. OPNET model of organization network with voice and video

---

[4] OPNET Modeler was provided under the OPNET University Programs
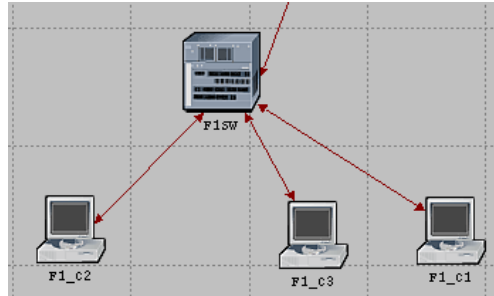
Figure 5. The floor subnet model

3.1.1   Modeling the Network

Figure 4 shows a simulation model for the existing network. The simulation model of the organization network, for the most part, is an exact replica of the real network. Figure 4 shows the described topology.   As discussed in Section 1.4, the *gatekeeper* signaling traffic is ignored, and hence modeling such an element and its traffic is not taken into account as we base our study on the worst-case situation.

Floor LANs have been modeled as subnets that enclose an Ethernet switch and three designated Ethernet workstations used to model the activities of the LAN users, as shown in Figure 5. For example, the Ethernet workstations for Floor 1 are labeled as `F1_C1`, `F1_C2`, and `F1_C3`. `F1_C1` is a source for sending voice and video calls. `F1_C2` is a sink for receiving voice and video calls. `F1_C3` is a sink and source of background traffic. It is to be noted that the model of floor LANs does not represent precisely the floor multimedia PCs. However, building a model with such exact floor network configurations will make simulation very tedious and almost manual.   This is so because it requires for each time a new voice or video call (or a group of calls) is added to perform two tasks: first adding individual PCs with individual profiles and settings, and then running the simulation. This has to be repeated manually and results have to be examined after each simulation run. As will be noted later, the simulation run can take up to 15 hours. Our simulation approach is an automated one, as the simulation is configured to automatically keep generating three sessions every two seconds. Our approach and model has no impact on the performance of internal nodes and links inside the core network. The only impact is on the utilization and queueing delay at the edge links connecting the floor subnet workstations to their intra-floor switch. The health of these edge links must be checked individually at the end of the simulation run.

In [28], an OPNET simulation model was presented for a similar organizational network for deploying *only* VoIP. In sharp contrast to the simulation work presented in [28], the simulation configuration and setup for videoconferencing are considerably different when considering the deployment of both voice and video calls simultaneously. The reader is encouraged to review the work of [28] especially when it comes to the following issues: (1) VLAN setup, (2) default configurations for the router, switches, and links in order to address the issue of future growth factor, and (3) generating background traffic.

8

3.1.2 Generating Videoconferencing Traffic

Application and Profile Settings. OPNET has two built-in or predefined applications for voice and video, that is VoIP_APPLICATION and VC_APPLICATION, respectively, as shown in Figure 4. Figure 6 shows the configurable parameters for voice and video applications. Figure 6(a) shows configurable parameters for voice application. One important parameter is the "*Encoder Scheme*", which is set to G.711. Another important attribute is the Voice Frames per Packet. In OPNET terminology, a voice frame is a collection of 32 voice samples of which each sample is 8 bits, i.e. each voice frame is 32 bytes in size. However, in the standard we are adopting, each VoIP packet has a payload of 160 bytes. Hence we set Voice Frames per Packet attribute to 5. Figure 6((b) and (c)) show the configurable parameters for video application. We set the attributes to have each video stream consisting of 30 frames/sec, with each frame containing 1344 bytes worth of payload.
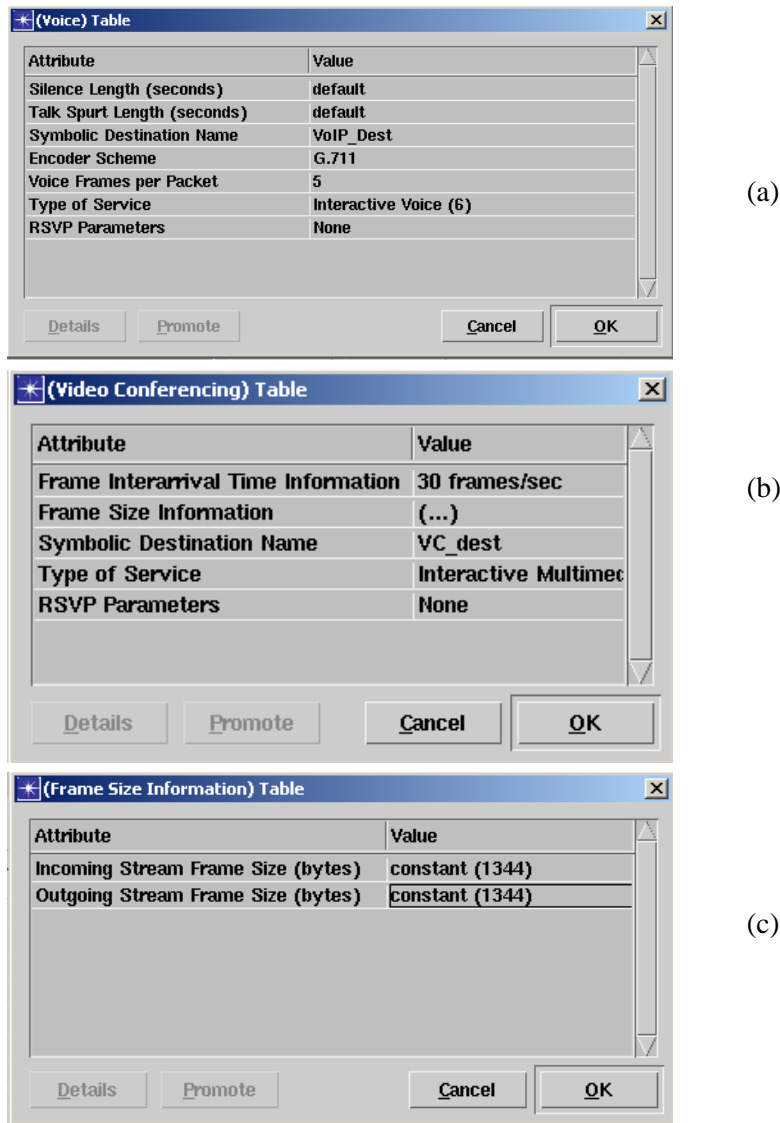

Figure 6. Voice and video application definitions

After defining and configuring the voice and video applications, it is required to configure the way in which workstations will be implementing this application. In general, the behavior of a network workstation is defined through its profile which is basically a collection of applications that can be configured to control their start and end times, in addition to their repeatability. Afterwards, the floor subnet workstations, designated for generating voice and video calls, will be configured to support this profile. The designated workstations include `F1_C1`, `F2_C1`, and `F3_C1`.

While configuring the profile, it should be kept in mind that our goal is to determine the network capacity for supporting both voice and video calls. We accomplish this by adding these calls incrementally (i.e., increasing the videoconferencing load slowly) to the network while monitoring the thresholds or bounds for delay and network capacity or bandwidth. When any of these bounds has been reached, the maximum number of calls can then be known. Note that the bound for network bandwidth is reflected in seeing packets being lost or observing a mismatch between traffic being sent and received.
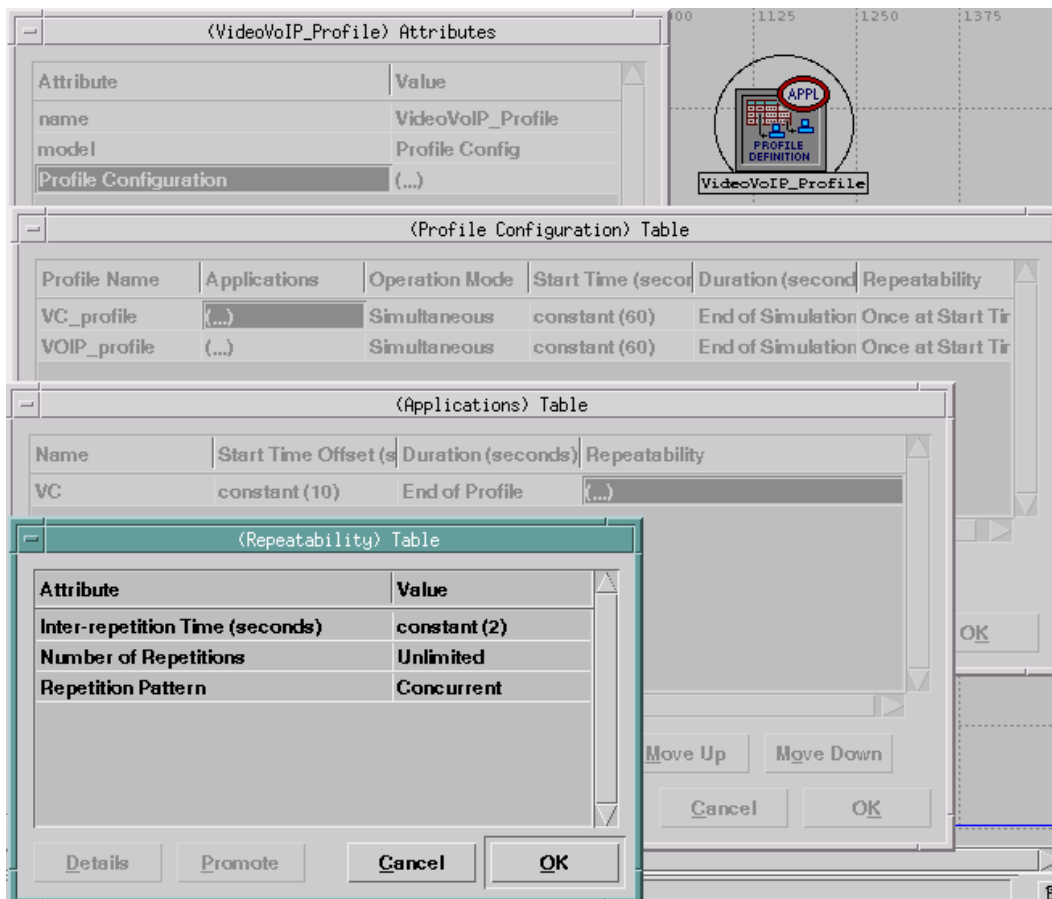

Figure 7. Voice and video profile settings

In OPNET terminology, we need to configure the profile to add calls repeatedly at a fixed rate. To achieve this, the application profile must be first configured to establish one call at the beginning and then add a call every 2 seconds. Hence, we set the `VideoVoIP_Profile` attributes as shown in Figure 7. It is to be noted that the first voice and video traffic will be produced traffic after 70 seconds (60 seconds with another offset of 10 seconds) from the start of the simulation run. By default in OPNET, the background traffic starts after 40 seconds. These start times are tunable and can be helpful in debugging the simulation by making sure simulated traffic is being properly generated at certain times. To keep generating calls, the repeatability attribute is set to *Unlimited* with an inter-repetition time of 2 seconds, as shown in Figure 7.

Generating Traffic Based on Call Distribution. The floor subnet workstations of `F1_C1`, `F2_C1`, and `F3_C1`, designated for generating videoconferencing traffic, are configured to support the `VideoVoIP_Profile`. This is simply done by adding this profile to each workstation's lists of supported profiles. Also, `VC_profile` and `VoIP_profile` are defined as the supported application profiles for the designated workstations responsible for generating voice and video traffic. It is essential to set the destination preferences of each workstation so that the call distribution ratios can be implemented. For a workstation to qualify as a valid destination, its client address attribute should be set to a unique value, which can be used to reference this workstation. Next, we add the symbolic name `VoIP_Dest` and `VC_Dest`, shown in Figure 6((a) and (b)), in the `Application: Destination Preferences` table of each sending workstation, as shown in Figure 8(a). As explained before, the symbolic name defines the workstation to which the call will be made. To define the actual destinations and the call distribution ratios, a set of actual names is mapped to each symbolic name. For example, Figure 8(b) shows the list of actual names that map the symbolic names of voice and video for `F1_C1`. Remember that `F1_C1` is the sending workstation of Floor 1. Figure 8(b) also shows the selection weights that are used to control the distribution of calls. In this case, the selection weights are computed based on the distribution of the probability tree shown in Figure 2.

The voice and video traffic will be generated and received by workstations within the floor. Since we are only interested in determining the network delay and capacity, we rule out any delays to be introduced by the workstations of subnet floors. Also, we rule out having the processing of workstations being the bottleneck for bandwidth. Known the processing delay at the workstation is of no interest to us, since a delay of 20 ms was accounted for the receiver and sender as discussed in Section 2.1.2. In other words, we are interested in only the network delay which has a bound 80 ms. Hence we set the processing rate for each work station to *infinity*. We accomplish this, as shown in Figure 8(c), by changing the workstation's attribute of `Datagram Forwarding Rate`. This attribute is one of the attributes for the `IP Processing Information` table of the workstations.
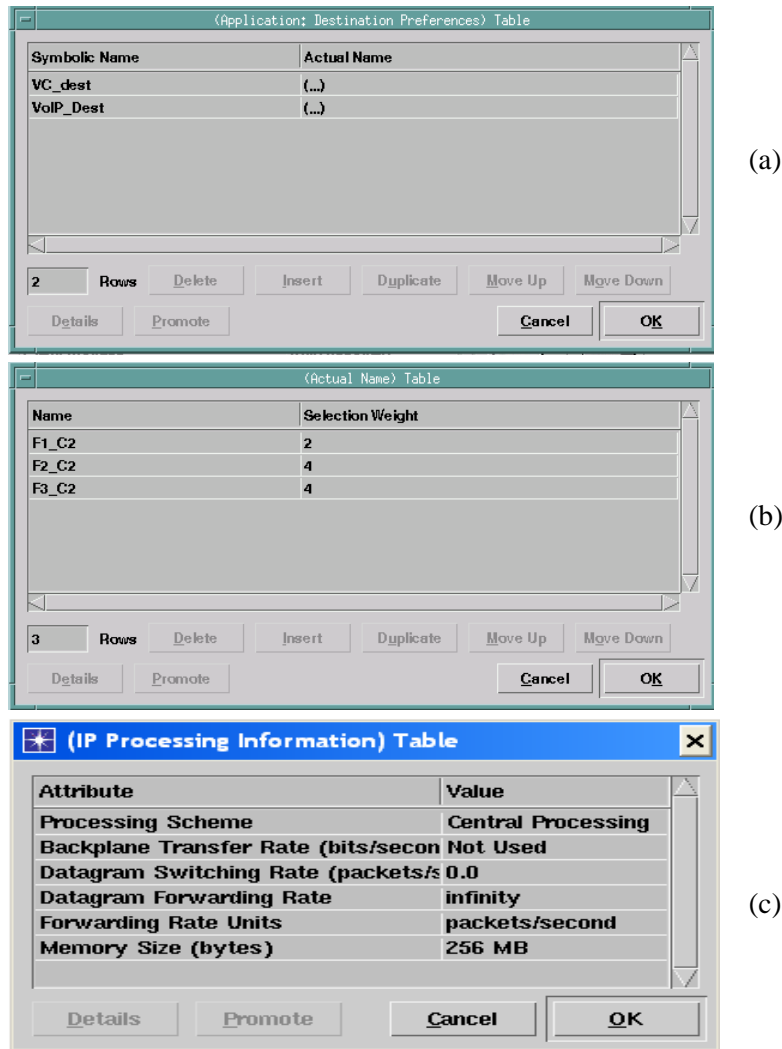
Figure 8. Settings of multimedia workstations

3.1.3   Simulation Results

OPNET has to be configured to obtain graphed results for numerous network components which include voice and video traffic, router, switches, and links. In this section, graphed results for some of the most important components are shown. We configured the duration of the OPNET simulation run for 4 minutes. As per our configurations, the generation of background traffic, by default in OPNET, started at 40 seconds from the start time of the simulation run. The videoconferencing traffic of both voice and video started at 70 seconds at which a total of 6 calls are initially added (3 for voice and 3 for video). Then, every 2 seconds another 6 calls are added (as shown in   Figure 9(a) verifies that at 70 seconds of simulation time 3 voice calls and 3 video calls are added, and thus producing traffic load of 300 packets for voice and another 180 packets for video. This is expected and is in line with analysis since the bandwidth for a single bidirectional voice call is 100 pps and for a single bidirectional video call is 60 pps. Figure 9(b) verifies the corresponding bandwidth being

generated in bps (bytes per seconds). At 70 seconds, 254 kbps is generated for 3 video calls and 68 kbps is generated for 3 voice calls. These rates are also in line with bandwidth analysis of Section 2.1.1.
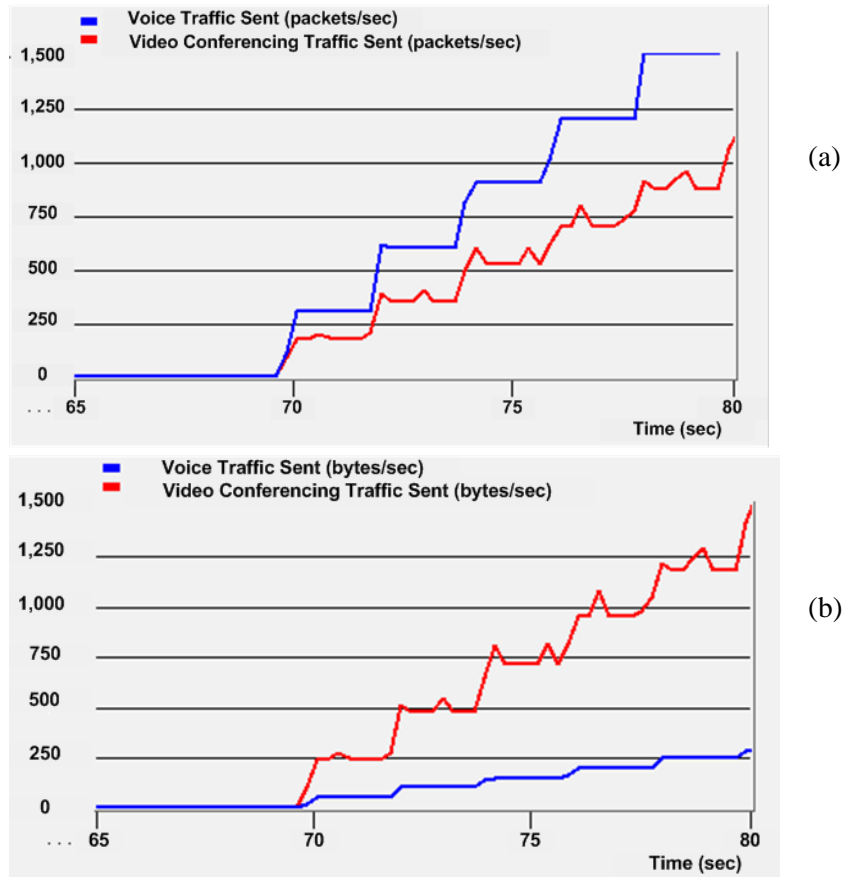

(a)

(b)

Figure 9. Generating videoconferencing load in pps

The Simulation stops at 4 minutes, reaching maximum calls at actually 3 minutes and 58 seconds. At 3.58 minutes, the maximum number of calls is $6 + ((3 \times 60 + 58 - 70)/2) \times 6 = 510$. Ideally, this would translate into 255 calls for voice or (255x100 = 25,500) voice packets, and another 255 calls for video or (255x60=15,300) video packets or frames. It is clear from Figure 9 that the maximum packet rate for voice and video that were sent at the end of simulation run was close to these numbers. Figure 9 shows the behavior of global videoconferencing traffic as voice and video calls are added every two seconds to the network by the designated workstations. Figure 9 shows the traffic in pps that was sent, received, and dropped.
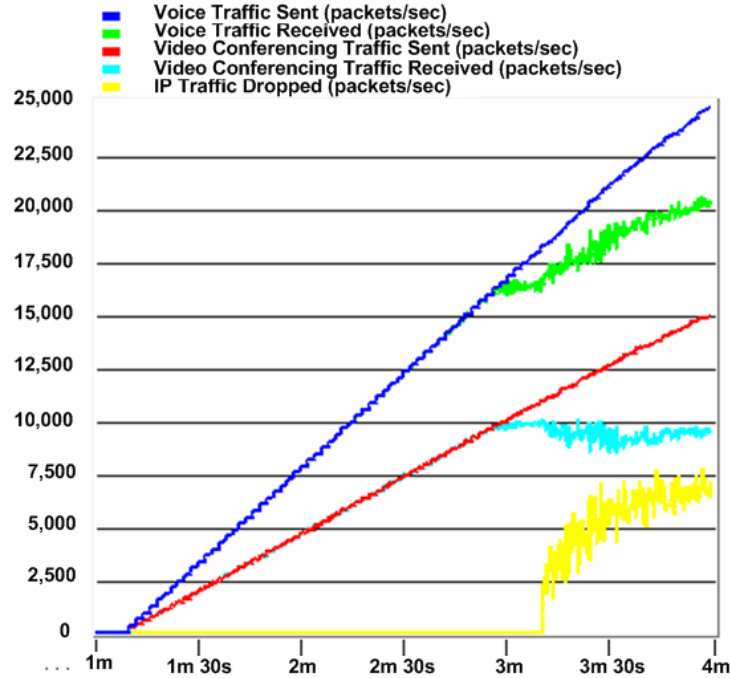
Figure 10. Global videoconferencing traffic in pps

One can determine the total number of calls that the network can sustain by examining network bandwidth or delay bounds. We first investigate the bandwidth bound. Figure 10 shows clearly that not all of the packets being sent get received. There is a mismatch between traffic sent and received. We can determine the number of calls that can be supported by examining the X and Y axes. When zooming in and examining the X axis of the simulation run time, it is clear that the last successful addition of three calls was at exactly 2 minutes and 56 seconds. The next addition, as shown, was at 2 minutes and 58 seconds and resulted in a mismatch. This mismatch point happens to be true for both voice and video traffic. One can determine the number of calls to be supported by the network for voice and video simply by calculating how many calls have been added until the last successful addition of three calls (i.e., at 2 minutes and 56 seconds.) Since the last successful addition point was the same for voice and video, this yields to $3 + ((2 \times 60 + 56 - 70)/2) \times 3 = 162$ videoconferencing calls. Figure 10 shows also the dropped traffic of IP packets. These dropped packets occur mainly due to overflow in the router memory buffer. Simulation results of the Router shows that dropping of packets happens after the mismatch points of voice and video traffic, as it takes some time for the buffer to completely fill up.
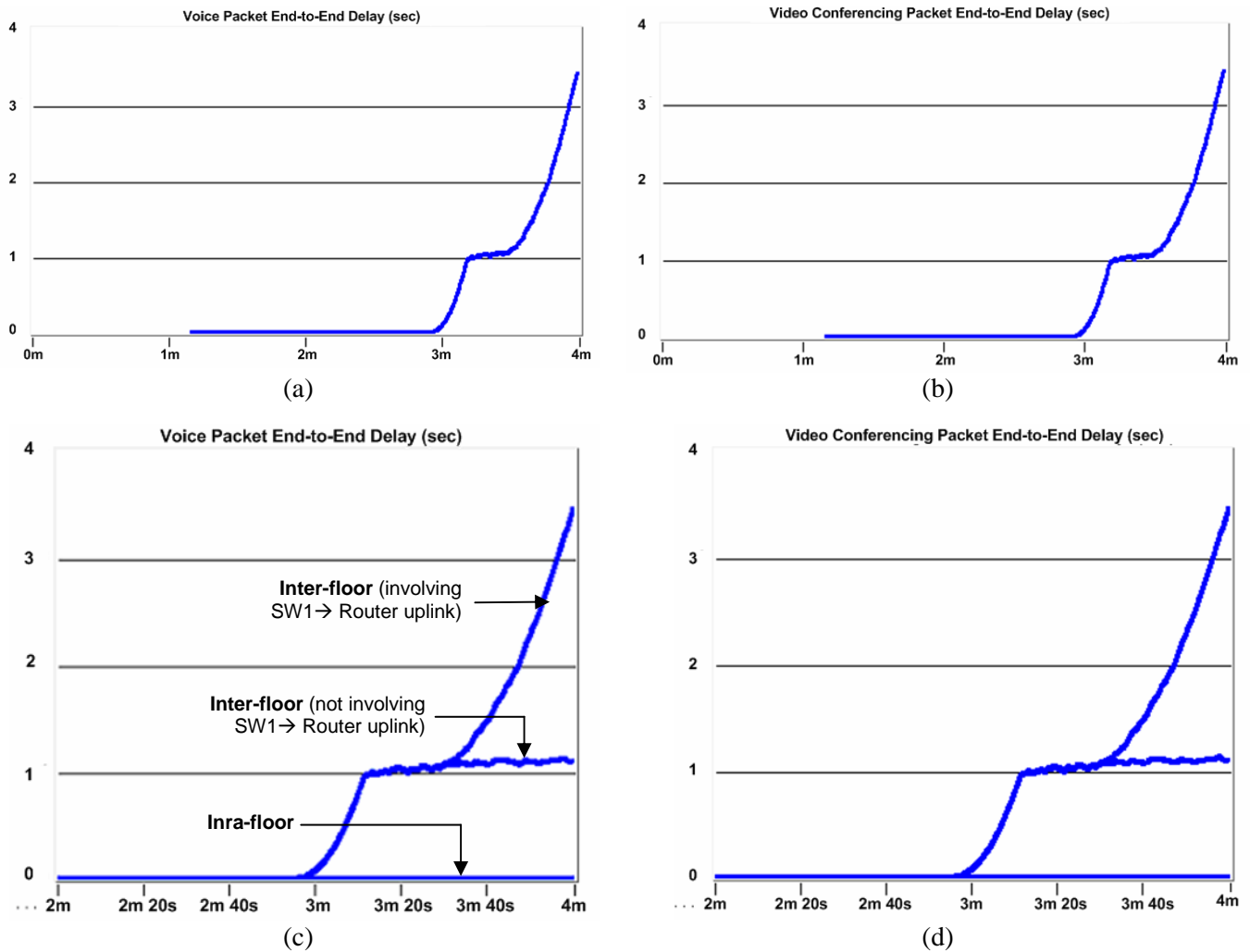
14

Figure 11. Global videoconferencing end-to-end delay

Figure 11 shows the corresponding voice and video end-to-end delay of Figure 11. For Figures 11((a) and (b)), the delay is reported as the maximum values of a bucket of 100 collected values. The OPNET default reported delay configuration is the sample mean of a bucket of 100 collected values. Figure 11((c) and (d)) depict a different collection mode, in which "all values" are collected and plotted. There are three types of delays. First, intra-floor delay of packets associated with calls within the floor and not passing through router, SW1, or SW2. These are the smaller delays, in which the majority of these values stay close to 2.5 ms. Second, inter-floor delay of packets associated with calls involving the router and passing through SW1→Router uplink. And finally, inter-floor delay of packets associated with calls involving the router but not passing through SW1→Router uplink. The simulation graphed results of links show that SW1→Router uplink gets saturated to 100% as calls are added. This results in huge delays as the arrival rate exceeds the link capacity. In OPNET, links are modeled with infinite buffer, and thus the delay starts shooting to infinity, and thus resembling the delay behavior of M/M/1 or M/D/1 queue. On the other hand, the router is modeled in OPNET with finite buffer (i.e., similar to an M/M/1/B queue). For this case as the router gets saturated, the delay flattens off (as expected from the known behavior of M/M/1/B queueing latency).

15

As discussed in Section 2.1.2, the end-to-end network delay should not exceed 80 ms for voice or video packets. The delay for voice and video, as shown in Figure 11, behaves similarly for the most part because both packets follow the same paths. When zooming in on the area of 3 minutes of Figure 11, it is depicted that the delay stays less than 80 ms until a simulation time of 3 minutes and 2 seconds at which the delay increases sharply. One can then determine the number of calls that the network can support to satisfy the 80 ms time constraint. The number of videoconferencing calls can be computed as $3 + ((3 \times 60 + 2 - 70)/2) \times 3 = 171$. Therefore one can conclude that, based on these simulation results, the number of videoconferencing calls to be supported by the network is bounded more by the network bandwidth than the delay. Hence, the number of videoconferencing sessions that the network can support based on simulation is 162 calls.

Router. The router is a vital component for the network under study. It is imperative to investigate and understand its health, performance, and behavior. Graphed results of a number of important key health indicators are shown in Figure 12 and Figure 13. The figures show key performance indicators for IP traffic, CPU utilization, processing and queuing delays, and link utilization.
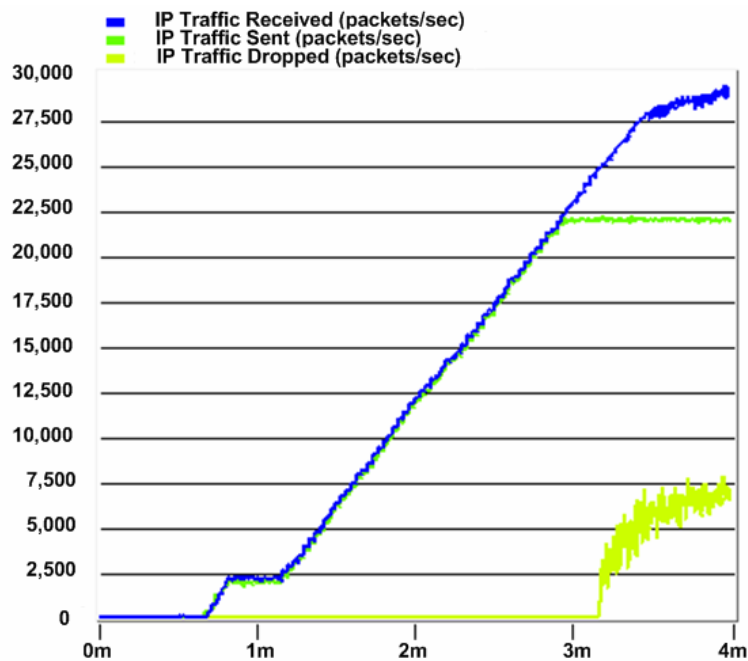

Figure 12. IP traffic handling at the Router

Figure 12 confirms that the background traffic starts at 40 seconds and stays flat for 30 seconds with a total of 812 and 869 packets coming from SW1 and SW2, respectively. At 70 seconds, the videoconferencing traffic gets generated with a rate of 6 sessions every 2 seconds. Figure 12 shows that the IP traffic being processed by, i.e. sent out from, the router stays flat at 2 minutes and 56 seconds. Figure 13 shows that at this time the corresponding CPU utilization has reached 100%. This is the saturation processing point for the router and at which the mismatch between traffic sent and received occurs. After this point, the router buffer starts being

filled very quickly by new calls, resulting in dropping of packets, as shown in Figure 13. The dropping of packets occurs after 14 seconds, precisely at 3 minutes and 10 seconds. At this point the memory of 32 Mbytes of the router's forwarding queue is completely full. Also, the corresponding processing (that includes processing) delay sharply increases and it stays relatively flat with a delay of approximately 1.1 seconds for a duration of 18 seconds or until 3 minutes and 28 seconds. After that, the delay increases rapidly. The reason for that is that saturation of SW1 → Router uplink at 3.35 minutes. The resulting global end-to-end delay shape, shown in Figure 11 and also in Figure 13, stems from the way OPNET models the router and link.
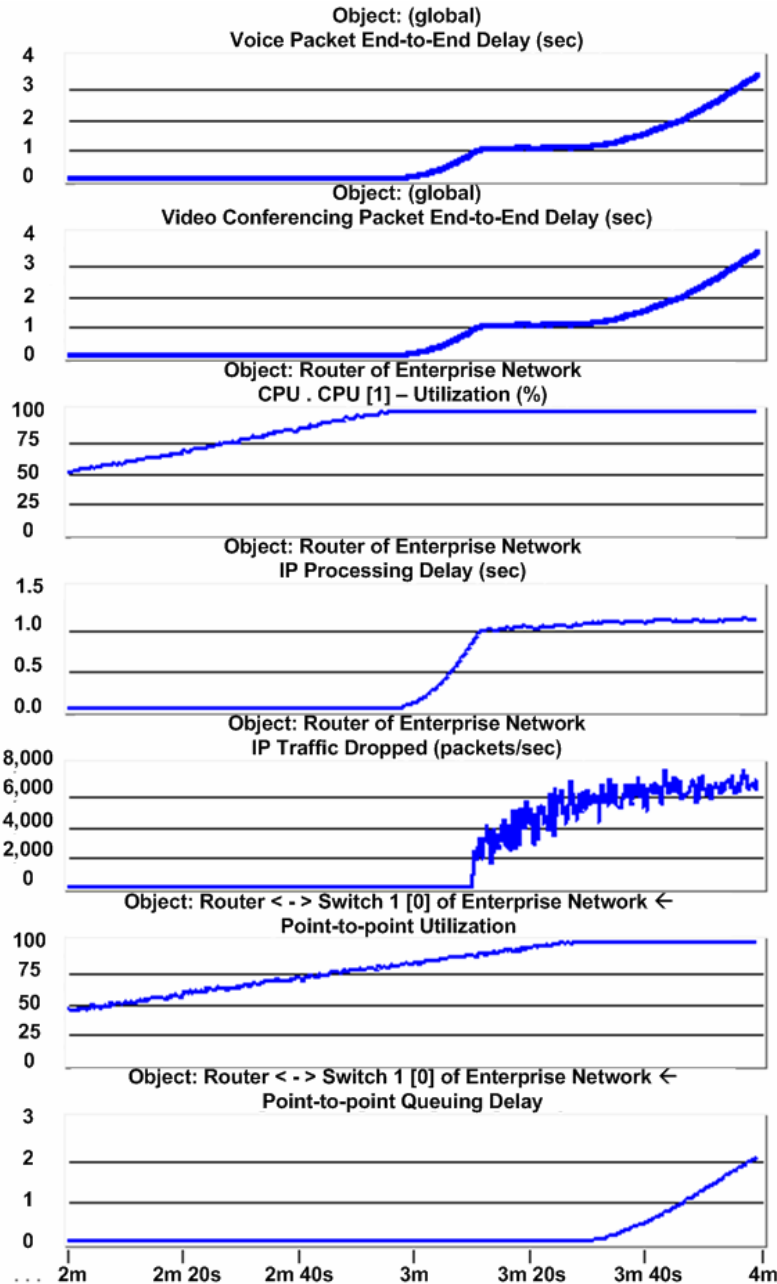

Figure 13. Other key health indicators for the Router

17

Links. Checking the health of key links in the network is essential. The uplinks and downlinks connecting the inter-floor switches to the router are of great significance. Also, it is imperative to check the health of uplinks and downlinks connecting the intra-floor switch to the designated workstations responsible for generating videoconferencing traffic. Out of these links, the uplinks and downlinks of the router were the most critical. All other links showed acceptable utilization and small to negligible queueing delays. Thus we focus our attention on analyzing and investigating the state of bidirectional links for SW1⇔ Router and SW2 ⇔ Router.

In Figure 14 we investigated the utilization for SW1→ Router uplink and observed its impact on the global end-to-end packet delay of both voice and video traffic. Figure 14((a) and (b)) show graphed results of the utilization for the links in both directions between the router to SW1 and SW2, respectively. Except for the uplink utilization of SW1 → Router, it is observed that the other utilizations in general are healthy, even with the addition of the growth factor of 25% utilization of each link. As depicted, at the startup of the simulation run, the links are utilized 25% and then shoots higher at 40 seconds to about 10% as the background traffic is generated. As calls get generated every 2 seconds, the link utilizations keep steadily increasing. SW1→ Router uplink has more traffic passing through it as it is not affected by the traffic processed by the router. The generated traffic by the designated workstations of floor 1 and floor 2 keeps pouring into it. The same is true for SW2→ Router uplink, but the generated traffic is generated by the one designated workstation of floor 3, and thus has less traffic pouring into it. On the other hand, the traffic passing through the downlinks of both SW1 ← Router and SW2 ← Router are affected by the packet processing of the router. As the router reaches its saturation point, the utilization starts to flatten off.
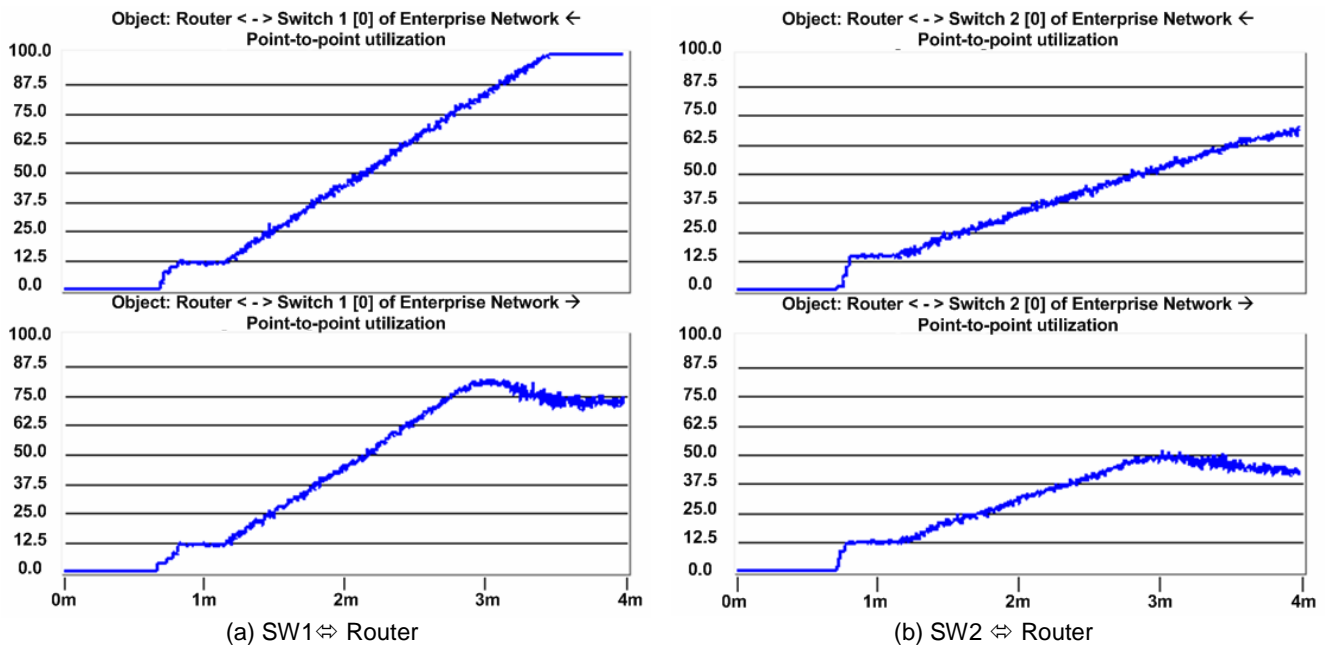

(a) SW1⇔ Router   (b) SW2 ⇔ Router

Figure 14. Utilization of Router ⇔ Switch links

3.1.4 Simulation Accuracy and Final Run

Our previous OPNET simulation was executed on a Sun E450 server. The elapsed time for the simulation run was approximately 20 hours. It is possible to cut down on the simulation run time by configuring the generation of more calls simultaneously at the start of the simulation, say 150 calls, and then add more calls slowly at a rate of 6 calls (3 voice and 3 video) every 2 seconds. In addition, one can also stop the simulation at earlier time, say at 3 minutes instead of 4. In order to gain accuracy (with a narrow confidence interval) of our simulation results, following the popular guidelines presented in [29, 30], five simulation replications were run by feeding OPNET with different initial seeds. OPNET's pseudo random number generator is based on BSD's algorithm which permits safely (i.e., with no concern of overlapping of random number streams) any integer value to be an initial seed. Five simulation replications were sufficient. Each simulation replication produced very similar graphical results, which when interpreted and analyzed led to the same total number of calls to be supported.

Based on the simulation results, the existing network can support 162 videoconferencing calls while satisfying the videoconferencing QoS of throughput, latency and packet loss. In Section 3.1.2, videoconferencing calls were added every 2 seconds and the simulation was not allowed to stabilize for a long time to reach the steady state. Our attention was focused on finding out the number of voice calls that the network can sustain. As a final check to ensure a healthy network and a normal behavior for all network elements, we perform a final simulation run in which 162 videoconferencing calls are added, all at once at the start of the simulation, say after 70 seconds. We let the simulation run execute for a prolonged amount of time, say good 5 minutes, to reach a steady state. Then we examine the health of each network element. In our example, this simulation run of supporting 162 calls was not successful. This simulation run took approximately 28 hours on Sun E450 server. The simulation run showed a mismatch between traffic sent and received and a delay of more than 80 ms. However, a successful simulation run of 144 calls showed compliant and healthy results with no packet loss, end-to-end delay of 2.15 ms, and adequate utilization of router and switch CPUs and links.
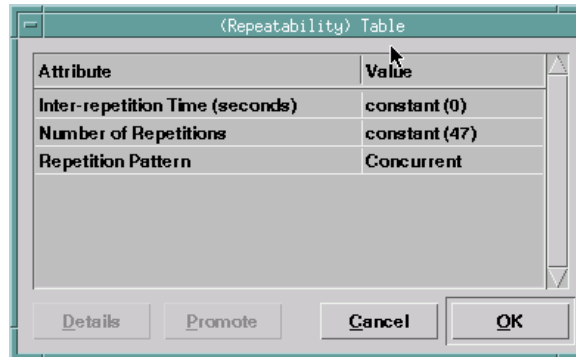

Figure 15. Configuration to generate 144 calls

In order to generate 144 concurrent calls, the configuration of Figure 7 for profile settings must be followed. The only difference is for setting the repeatability table. The table has to be modified as shown in Figure 15. After the generation of the initial 3 calls, the table shows that each floor designated workstation will generate 47 calls all at once, and hence the total number of calls that will be generated will be $3 + 47 \times 3 = 144$.

3.1.5   Simulation Results for Empirical Video Packets

Thus far we analyzed and interpreted simulation results when using fixed video packet sizes of 1344 bytes. In this section, we show results of simulation when using real measurement of video packet sizes. The empirical distribution of video packets is discussed in Section 2.1.1. OPNET can be configured to use empirical video packets by simply changing the values of incoming and outgoing stream frame size of Figure 6(c) to OPNET special "scripted" distribution in which a filename containing the packet sizes is specified.

Figure 16 shows the traffic generation in pps for both voice and video packet. The figure shows the IP traffic dropped fluctuates sharply when compared to the behavior of Figure 10, although in both figures the sent rate is the same. The sharp fluctuation is due to the size of the empirical video packets. A drop in a video packet of a large size would result in dropping multiple packets of smaller size from the router buffer. The fluctuation of packet sizes is more apparent in Figure 17 which shows the traffic sent in bps. By following the same analysis of Figure 10 on bandwidth bound by examining the point of mismatch of traffic sent and received, it can be found that 156 video conferencing sessions can be supported.
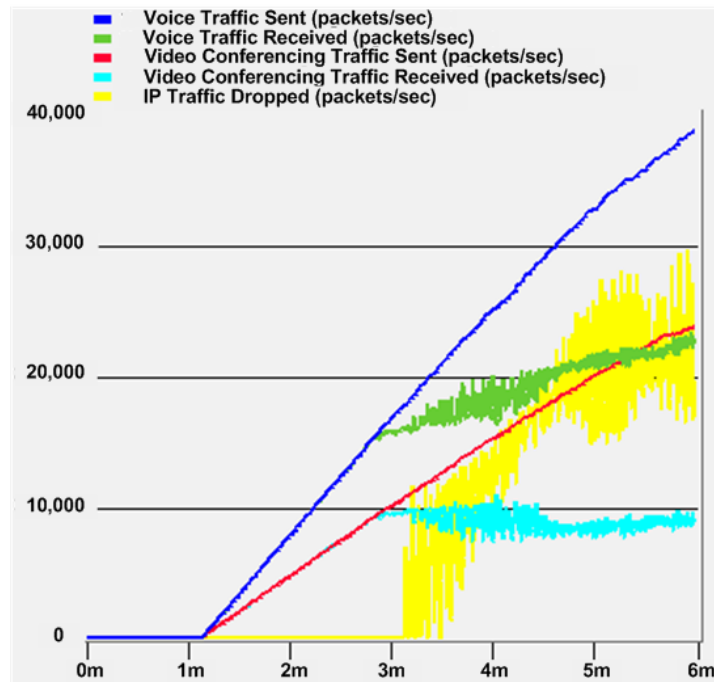

Figure 16. Global videoconferencing traffic in pps using empirical video packets
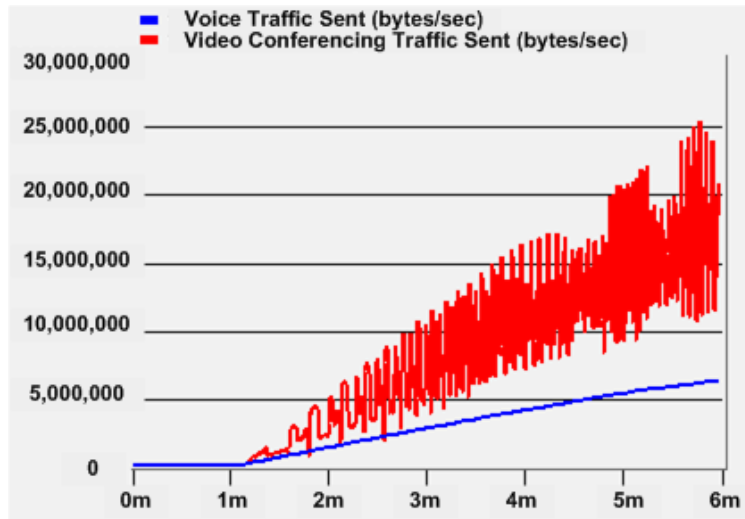
Figure 17. Generated videoconferencing traffic in bps using empirical video packets

The end-to-end delay of both voice and video is shown in Figure 18. With similar analysis as that of Figure 11 ((a) and (b)) the total videoconferencing sessions to be supported is 162. The delay curves resemble those of Figure 11. Also, a successful simulation run (with no packet loss and a delay of 22.5 ms) of 144 videoconferencing sessions were obtained when using empirical video packet sizes.
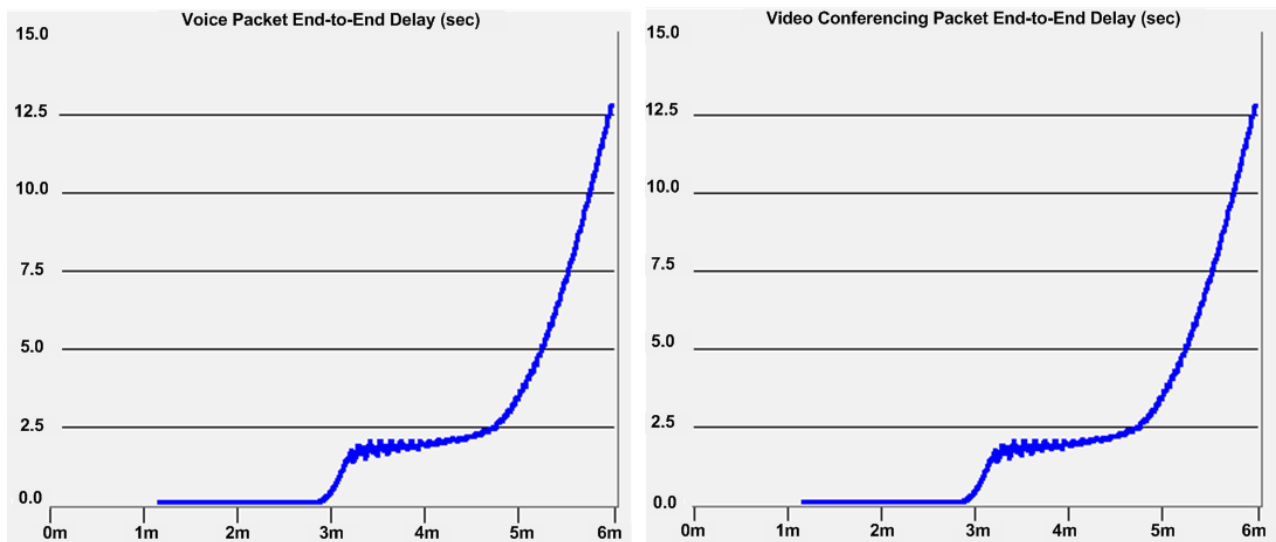

Figure 18. Global videoconferencing end-to-end delay using empirical video packets

4   Conclusion

The paper described in great detail how OPNET can be utilized to assess the readiness of existing IP networks to support desktop videoconferencing. The paper offered extensive interpretations and analysis of simulation results and showed how to draw proper conclusions. Two scenarios of traffic were used for OPNET simulation (viz. fixed and empirical video packet sizes). Videoconferencing traffic of fixed size of 1344 bytes for video packets gave similar outcome (in terms of number of videoconferencing sessions to support) to that of video packets of empirical size distribution.   However, there was a difference in resulted simulation graphs between the two scenarios.   For verification purposes, we compare OPNET simulation results for fixed-size packets reported in this paper with the results of an analytic approach described in [11]. We observe that the difference in number of videoconference sessions to be supported is only 9, and thereby reflecting the correctness of our simulation work.   It is to be noted that the analytic approach described in [11] is an approximation and our simulation work presented in this paper is more accurate as OPNET offers more detailed and realistic simulation of a real-life network environment.

In this paper, one network topology was considered and only peer-to-peer videoconferencing calls were considered. However the simulation techniques and details offered in this paper can be extended and applied on any generic network. Simulation work presented in this paper can also be extended to other popular and important types of multimedia and videoconferencing (such as real-time web conferencing and broadcast/multicast videoconferencing as well as streaming stored audio and video).

References

[1] Recommendation H.323, "Packet-based Multimedia Communication Systems," ITU, January 1998.

[2] Eurescom H.323 Studies http://www.eurescom.de/~public-webspace/P800-series/P807/start/printerversion.asp

[3] NetIQ, Vivinet Assessor, http://www.netiq.com

[4] BMCsoftware, "PATROL DashBoard," http://www.bmc.com

[5] Spirent Communications Resources, http://www.spirentcom.com

[6] Radvision H.323 Protocol Toolkit, http://www.radvision.com

[7] Omegon, "White Paper: NetAlly," http://www.omegon.com/

[8] Lucent Technologies, "VitalSuite," http://www.qip.lucent.com

[9] VideNet Scout Resources, http://scout.video.unc.edu

[10] P. Calyam, W. Mandrawa, M. Sridharan, A. Khan, P. Schopis, "H.323 Beacon: An H.323 application related end-to-end performance troubleshooting tool", ACM SIGCOMM 2004 Workshop on Network Troubleshooting (NetTs), 2004.

[11] K. Salah, "An Analytical Approach for Deploying Desktop Videoconferencing," IEE Proceedings Communications, vol. 153, no. 3, June 2006, pp. 434-444.

[12] R. Abler and G. Wells, "Supporting H.323 Video and Voce in an Enterprise Network," Proceedings of the 1st USENIX Conference on Network Administration, Santa Clara, Clifornia, April 7-10, 1999.

[13] Cisco Systems, "Implementing QoS Solutions for H.323 Video Conferencing over IP," Technical Report, Document ID: 2155, available at: http://www.cisco.com/warp/public/105/video-qos.html

[14] B. Goode, "Voice over Internet Protocol (VoIP)," Proceedings of IEEE, vol. 90, no. 9, Sept. 2002, pp. 1495-1517.

[15] B. Liao, "IP Videoconferencing Embracing the Next Era of Visual Collaborations," Proceedings of the 5th IEEE Asia-Pacific Conference on Optoelectronics and Communications Conference, Beijing, October 1999, pp. 853-863.

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

[17] P. Calyam and C. G. Lee, "Characterizing voice and video traffic behavior on the Internet", Proceedings of the International Symposium on Computer and Information Sciences (ISCIS), October 2005.

[18] S. Kouremenos, S. Domoxoudis, V. Loumos, A. Drigas, "Modeling of Unconstrained and Constrained H.26x Traffic over IP Networks", Elsevier Computer Communications Journal, 2006.

[19] S. Domoxoudis, S. Kouremenos, A. Drigas, V. Loumos, "Frame-based Modeling of H.264 Constrained Videoconference Traffic over an IP Commercial Platform", IEEE TridentCom, 2006.

[20] A. K. Alasmari, "A New Video Compression Algorithm for Differential Videoconferencing standards," International Journal of Network Management, vol 13, no. 1, November 2003, pp. 3-10.

[21] J. Sprey, "Videoconferencing as a Communication Tool," IEEE Transactions on Professional Communications, vol 40, no. 1, March 1997, pp. 41-47

[22] D. Tynan, "RTP Payload Format for BT.656 Video Encoding," RFC 2431, October 1998.

[23] Recommendation G.114, "One-Way Transmission Time," ITU, 1996.

[24] W. Jiang, K. Koguchi, and H. Schulzrinne, "QoS Evaluation of VoIP End-Points," Proceedings of IEEE International Conference on Communications, ICC'03, Anchorage, May 2003, pp. 1917-1921

[25] M. Baldi and Y. Ofek, "End-to-End Delay Analysis of Videoconferencing over Packet-Switched Networks," IEEE/ACM Transactions on Networking,, vol. 8, no. 4, August 2000, pp. 479-492.

[26] R. Steinmetz, "Human Perception of Jitter and Media Synchronization," IEEE Journal on Selected Areas in Communications, vol. 14, no. January 1996, pp. 61-72.

[27] OPNET Technologies, http://www.mil3.com

[28] K. Salah and A. Alkhoraidly, "An Opnet-based Simulation Approach for Deploying VoIP," International Journal of Network Management, vol. 16, no. 3, 2006, pp. 149-183.

[29] K. Pawlikowski, H. Jeong and J. Lee, "On Credibility of Simulation Studies of Telecommunication Networks", IEEE Communications Magazine, vol. 40, no. 1, January 2002, pp. 132-139.

[30] A. Law and W. Kelton, Simulation Modeling and Analysis, McGraw-Hill, 2nd Edition, 1991.