

A Simulation-Based Comparison of Multimedia Traffic Prioritization Schemes for High-Performance Input-Queued Packet Switches

Salman A. AlQahtani

Department of Computer Engineering, King Fahd University of Petroleum and Minerals
Dhahran, Saudi Arabia

Abstract: This study investigates the problem of enabling Quality of Service (QoS) of multimedia traffic at the input port of high-performance input-queued packet switches using a simulation-based evaluation. We focus on the possibility of assuring QoS of multimedia traffic in such switches by implementing traffic prioritization at the input port where each input-queue has been modified to provide a separate buffer for each of the service classes. The multimedia traffic can be categorized into three classes based on its real-time properties and loss tolerance and assigned a separate queue for each class. We select appropriate models for each of three types of traffic: video, voice and data. Then, we propose an efficient dynamic scheduling strategy by implementing multimedia traffic prioritization at the input port of input-queued packet switches. Simulation-based comparisons show that while the static priority scheme is beneficial for highest priority class at the expense of the others, the dynamic prioritization serves fairly well all the classes in terms of delay and loss requirements.

Key words: Traffic analysis and control, Quality of Service, Performance evaluation and Network architectures

INTRODUCTION

Advanced networking technologies have enabled the integration of multimedia transmissions, such as video, voice and data, in packet-switched networks. The multimedia traffic can be divided into two distinct types: real time, e.g., live audio and video packets and non-real time, e.g., file transfer packets. In order to meet the requirements of these types, protocol designer have to grasp the characteristics of the traffic and select the processing method suitable for its service and performance requirements. Specifically, real-time packets are loss-insensitive but delay sensitive. This means that packets of this type should be served so rapidly by the switch scheduler that they get to their destination in the shortest time possible, even if some of them are lost. On the other hand, non-real time packets are delay insensitive but loss sensitive. This means that packets of this type should be served so carefully that no packet is lost, even if the packet incurs a longer delay. The scheduler (in routers or switches), therefore, has to solve this problem by deploying efficient handling schemes that can satisfy the required QoS.

Queuing by routers and switches contributes to providing quality of service (QoS) for various network applications. A major issue in the design of packet switches is controlling the access to the switching fabric by a scheduling algorithm (SA) to avoid contention at output ports. A new packet comes in at

one of multiple input ports, is sent across switching fabric to the appropriate output port and is finally transmitted across the selected output line. But what happens when multiple newly arrived packets have to go to the same output port. This is so-called output-contention problem, which can be resolved in two ways. One is to use a very fast switching fabric^[1]. This way all packets can be transferred to the same output almost simultaneously. However, the need to queue these packets remains since all of the packets cannot be transmitted across the transmission line at the same time. This is so-called line-contention problem and corresponding generic type of packet switches are called output-queued packet switches^[2-4]. Another way is to queue newly arrived packets at the input ports and then to transfer them across switching fabric when appropriate capacities are available. The corresponding generic type of packet switches are called input-queued packet switches^[2-4] (Fig. 1).

The contribution of this study is two-fold. First, we present the appropriate traffic source model that reflects the correct behavior of each type of multimedia traffic (video, voice and data traffic). Second, we propose an efficient new scheduling strategy by implementing multimedia traffic prioritization at the input port of input- queued packet switches. We categorize the multimedia traffic into three classes in term of delay and loss requirement and assign a separate queue to each class.

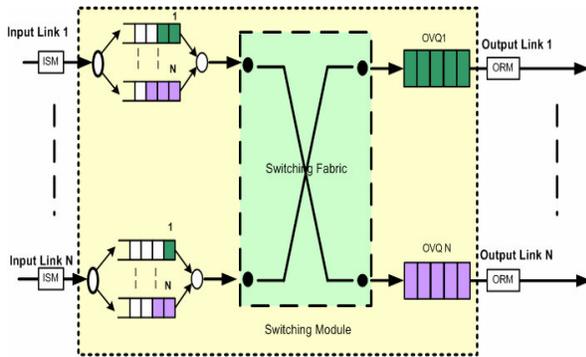


Fig. 1: Model of an Input-output Queued switch

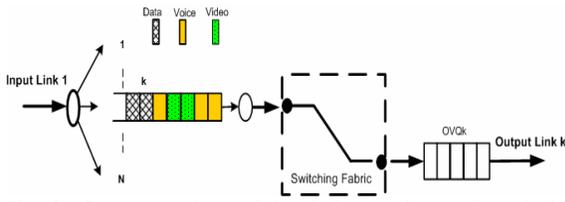


Fig. 2: Conceptual model of Input Queued switches without packet prioritization

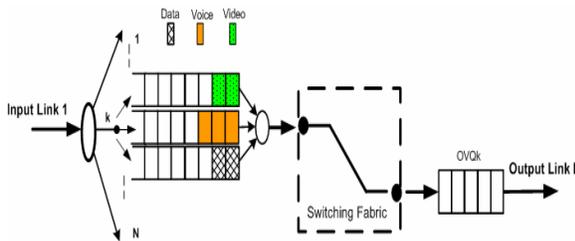


Fig. 3: Conceptual model of packet prioritization

RELATED WORKS AND PROBLEM DEFINITION

Input-Queued (IQ) switches with virtual output queuing (VOQ) based buffering schemes are often adopted as architectures of choice for high-speed switches or routers (Fig. 1). Most of the implemented high-speed IQ switches internally operate on fixed-size data units (cells): the Lucent GRF^[5], the Cisco GSR^[4], the Tiny-Tera^[1], the AN2/DEC^[6] and the MGR/BBN^[7]. IQ switches segment packets of length L into cells of size S , internally switch the cells and then reassemble the cells into packets in reassembly buffers at the output ports. “Cell train” approach to reducing packet-level delays have been studied in a number of recent research efforts^[2-4].

Input-buffering systems have a major problem: the head-of-line (HOL) blocking. In recent years, many solutions are proposed in literatures to solve the HOL blocking problem^[2-4]. Also, It is shown that by differently organizing the input buffers, we can achieve 100% throughput for input-buffered switches^[8-10]. A significant research effort has been devoted in recent

years to the design of simple and efficient scheduling policies for input queued (IQ) and combined input output queued packet switches (CIOQ)^[2]. As a result, a number of switch control algorithms have been proposed^[2-4,6,11]. As traffic loads increase, router buffers begin to fill, which adds to delay. If the buffers overflow, packets are dropped. When buffers start to fill, prioritization schemes can help by forwarding high-priority and delay-sensitive traffic before other types of traffic is helped. This requires classifying traffic and moving into queues with appropriate service levels. Therefore, appropriate traffic scheduler for packet admission and congestion are in need. In particular, in order to support multiple classes of traffic with varying delay and loss requirements, priority mechanisms must be used to guarantee the required QoS. The second important factor that can contribute in designing a scheduler in an efficient way is the appropriate traffic source model. In fact, selecting the appropriate traffic source model to reflect the behavior of the traffic is an important issue toward understanding and solving performance related problems in packet switched networks supporting multimedia application.

Previous efforts of IQ switches^[3,4,6,11] are based on the assumption that the traffic streams are of the same type (Fig. 2). Therefore, corresponding switching techniques consider only a single priority class. In a few cases^[12-14] where multiple priority classes were introduced, the influence of multimedia traffic characterization was not specifically taken into account. In^[1-14], the traffic is classified only into two classes (class 1 and class 2) and arrivals are assumed to be Bernoulli. In^[12], an optimized and prioritized iSLIP is introduced. This was extended from basic iSLIP by adding some complexity for supporting prioritized traffic. This technique faces a deep decrease in performance under non-uniform traffic. In^[14], the input queued crossbar scheduler differs from the crossbar. Each input port queue maintains a 2-level scheduler, which is responsible for scheduling and forwarding the cells in the VOQs to buffered crossbar. Therefore, the scheduling algorithm includes two parts: the TF scheduling in 2-level schedulers and the round robin scheduling in crossbar scheduler.

Presently, equal treatment of network traffic is no longer acceptable due to variation in requirements and criticality of an application. In this study, we deal with the problem of enabling QoS of multimedia traffic in input queued packet switches where each input queue has been modified to provide a separate queue for each of the service classes (Fig. 3). Our work enhances the state-of-the-art in following ways. Each input port maintains a 3-level scheduler, which is responsible for scheduling and forwarding the cells in the VOQs to buffered crossbar scheduler. The input port scheduler can be static or dynamic.

Table 1: A classification of multimedia traffic

Class	Characteristics	Multimedia Data
Class-I	Real time traffic with no packet loss and delay allowed	High quality Video
Class-II	Real time traffic with a certain degree of packet loss allowed but no delay	Voice
Class-III	Non real time traffic with no end-to-end packet loss allowed	Data traffic (ftp,web,etc.)

The input traffic is categorized into three classes (video, voice and data) and each traffic class has its own arrival distribution and its own queue. Our focus is on comparing the of static and dynamic scheduling at each input port.

SYSTEM MODEL

The conceptual model of a packet switch used was given in Fig. 1. It consists of (i) a set of N input ports, (ii) input buffers, (iii) a switch fabric, (iv) output buffers and a set of N output ports. A packet in the network is transmitted from its source node to destination node through one or more switches in a store-and-forward manner. A packet arrives at the incoming port and is stored in a buffer before it is forwarded to the outgoing port. The packet stays in the buffer until it is selected by the 'scheduler' according the queuing algorithm used. The switch fabric has many virtual channels that run from input ports to output ports. Finally, the packets are stored in the out buffer before dispatching to corresponding outgoing link^[2-4].

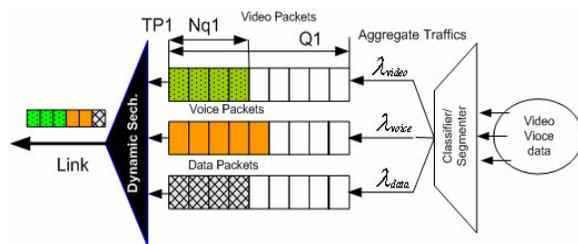


Fig. 4: Proposed Input port model

The IQ switch presented in^[2-4] does not differentiate between the traffic types. The packets that arrive to the input queue are served based on FIFO regardless of their classes. In order to maintain the QoS required by multimedia traffic, we modified the input queue at each input port of the switch so that each port provides a separate queue for each traffic type as shown in Fig. 4. Each input port has a scheduler, which is responsible for scheduling and forwarding the cells in the VOQs to buffered crossbar scheduler. The IQ switch presented in the Fig. 4 is equipped with a priority scheme that can solve this QoS problem easily. Each routers' or switches' input port maintains a set of queues, (three in our case), contain traffic of different types. An incoming packet is classified and inserted to the appropriate queue. These queues require access to a single resource, namely the output link. The crossbar

switch scheduler grants access to the link to each of the queues in an ordered fashion. So, the scheduling problem is: how to choose a queue for each packet transmission with a satisfactory performance that matches its QoS needs.

The detailed model of prioritization scheme implemented in input port is shown in Fig. 4. The proposed model consists of the following:

- * *Traffic source*: sources of different types of traffic.
- * *Classifier*: to determine the packet class (queue) and segmenter to segment the packet of length L into ceil (L/S) cells.
- * *Queues*: Three finite queues. First queue is for video, the second is for voice and the third is for data.
- * *Dynamic scheduler*: is a component in a packet switch, which is used as an effective solution for prioritizing the traffic in a network. It should have the ability of making decision effectively (dynamic decision based on Queue level or static decision).
- * *Output Link*: the outgoing transmission line used for transmitting the traffic.

PACKET PRIORITIZATION SCHEME

This section describes the prioritization scheme used, i.e., the set of rules used to decide which traffic class is granted access to the switching fabric. First, we will define the classes of multimedia traffic with their QoS requirements and then the prioritization schemes for these classes will be described.

Multimedia traffic classes: The multimedia traffic such as video, voice and data has different delay and loss requirements. According to^[15], multimedia traffic can be classified based on the delay and loss requirements into three classes shown in Table 1. We make an attempt to prioritize the traffic according to this classification.

Prioritization schemes: In order to provide each type of multimedia traffic with a consistent quality of service (QoS), an efficient packet prioritization scheme is in need. In this study, we use two types of scheduling schemes employing the priority to guarantee the consistent QoS for multimedia application over packet-switched networks. These two scheduling schemes are: (i) basic *static* priority scheme; and (ii) *dynamic* priority scheme. The former schedules packets based on traffic type and the later schedules the packets dynamically by utilizing the state of the system and the type of the traffic. These two types of schemes are described here.

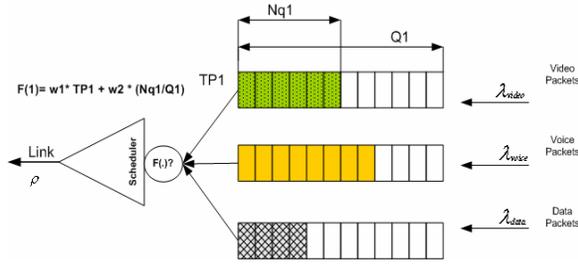


Fig. 5: Dynamic Priority Scheme model

Static priority scheme: The simplest priority scheme is a static or fixed priority scheme. In this scheme, the priority is always given to the delay and loss sensitive class (Class-I), before class-II (delay sensitive and loss insensitive) and the class-II is given the priority before class-III (loss sensitive and delay insensitive). The service policy is exhaustive, where the queue of the higher priority class will be served until it is empty, then the next priority queue is served.

Dynamic priority scheme: The above scheme is based on static conditions, where the priority is not affected by the state of the system or the load applied. The problem with this scheme is that as video and voice traffic load increases, the data traffic backlog will continue to grow. Therefore, a dynamic priority assignment is needed. Dynamic priority considers the current system state and traffic characteristics. In this scheme the priority level changes dynamically over time. In order to model this scheme, the packet priority is calculated based on a dynamic linear cost function $F(i)$. The input parameters of this function are the number of packets in the queue of each class and the traffic priority. This cost function is given as:

$$F(i) = w(i) * TP(i) + w(i) * TL(i); 0 < F(i) \leq 1 \quad (1)$$

Where $TP(i)$: Traffic priority of packets of type i and its value is in the range: $[0, 1]$; $TL(i)$: Traffic load of packet of type i and calculated as (Number of queued packet of class i / its queue size); and $w(i)$: a weight factor. The rationale behind this cost function is that a packet has the highest priority as it has higher traffic load and requests the service with high TP value. The video traffic will have the highest TP value and the data traffic will have the lowest TP value. The TL value will depend on the current state of the queue and available queue size. Each traffic class can have different queue size. The packets are scheduled optimally using the result of the cost function as shown in Fig. 5. The class with highest value $F(i)$ is served first. In particular when a queue is in high state (large number of packets are queued), the value of TL approaches 1, a packet in the queue is to be served prior to any other packets because it may be discarded sooner or later. However, if more than one queue is simultaneously in high state or low state, then the TP will be the dominant factor in the cost

function where the traffic with higher priority will be served first.

MODELING OF TRAFFIC SOURCES

Traffic source modeling and characterization constitute important steps towards understanding and solving performance-related problems in current and future packet switches networks. Therefore, In order to perform a successful design of new networks, it is an important issue to select the appropriate traffic source model to reflect the behavior of the traffic in the system. The traffic source models include three types of traffic generation: video, voice and data traffic. The purpose of this section is to examine the most appropriate models proposed for data, video and voice sources.

Data Traffic Model: In the last few years, an exponential growth of the Internet is observed. Most of the traffic volume consists of WWW related data transfers. Therefore, www-traffic is considered to be an important traffic source for future high speed packet-switched networks^[16,17]. Until mid 90s, it was assumed that the packet length of data traffic follows exponential distribution. A number of studies based on traffic measurements showed that the pattern of many types of data traffic did not follow exponential distribution. And in many cases, data traffic length followed heavy tailed distribution^[10,16,18]. In this study, we model the data traffic packet sizes using Pareto distribution with exponential inter-arrivals times as in^[16,18,19]. The well-known Pareto-distribution exhibits heavy-tailed behavior and has no maximum value and infinite variance. In contrast, packet size measurements from real data networks (wireless data networks) are bounded by finite minimum and maximum values and exhibit a high but also finite variance. Thus, a modified Pareto-distribution matching these properties is introduced in many research efforts. In detail the Pareto-distribution is normalized to cover values from a minimum k to a maximum m . The gradient of the distribution is given by a parameter α . Packet size is defined as: $PacketSize = \min(P, m)$, where P is a normal Pareto distributed random variable ($\alpha = 1.1$, $k = 81.5$ bytes) and m is maximum allowed packet size. Based on many measurements^[17,20], we observe that the packet of the sizes 40 bytes, 576 bytes and 1500 bytes dominate the traffic streams. Therefore, we will choose m (maximum allowed packet size) to be 1500 bytes. The normal Pareto distribution without cut-off is defined by:

$$f(x) = \frac{\alpha k^{-\alpha}}{x^{\alpha+1}}, x \geq k; \quad (2)$$

Table 2: Traffic model parameters

Characteristics	Traffic patterns
Video Traffic (2- state MMPP)	
Frame generation in state 1	25 VF/s
State 1 duration	Exponential distribution with average 25 ms
Frame generation in state 2	30 VF/s
State 2 duration	Exponential distribution with average 30 ms
Video packets size distribution	Gamma (104.8 cell, 29.7)
Voice Traffic (2- state MMPP)	
Bit rate in state 1	128 Kbps
State 1 duration	Exponential distribution with average 300 ms
Bit rate in state 2	64 Kbps
State 2 duration	Exponential distribution with average 700 ms
Data Traffic	
Average bit rated	Poisson process with rate 1Mbps
Packet size distribution	Pareto ($\alpha = 1.1$, $k=81.5$ bytes) and maximum size of 1500 bytes

Voice traffic model: An arrival process of packets from a voice source is fairly complex due to strong correlation among arrivals. It is widely accepted that the arrival of a new voice call can be characterized by a Poisson process and its duration can be represented by an exponential distribution. A single voice source may be modeled by the well-known ON-OFF process. When N independent voice sources are multiplexed, aggregated packet arrivals are governed by the number of voice sources in the ON state. The aggregate Poisson packet arrival process from superposition of many voice sources may be represented by a doubly stochastic Poisson process, which is modulated in a Markovian manner^[9,16]. In^[9] and^[16], the aggregate packet arrival process is approximated by a 2-state Markov Modulated Poisson Process (MMPP) (Fig. 6). The approximating MMPP model is chosen in such a way that its statistical characteristics match those of the aggregate traffic from the voice sources. In^[16], Heffes and Lucantoni used an MMPP to successfully model average delay of voice packets through an infinite buffer multiplexer. It has been shown that the approaches based on MMPPs are many orders of magnitude better than modeling the superimposed traffic sources simply as a Poisson process.

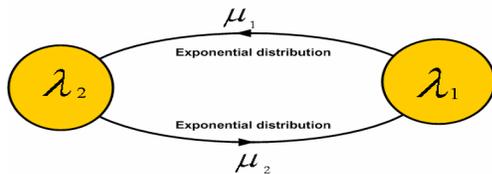


Fig. 6: 2-State MMPP used for Voice Source Generation

There are four parameters for the chosen two states MMPP, namely, the mean sojourn times in states 1 and 2 and the Poisson arrival rates in states 1 and 2 as follows: The arrival process in arrival state 1 and 2 are Poisson process with rate λ_1 and λ_2 and the inter arrival times between packets within state 1 and 2 are λ_1^{-1} and λ_2^{-1} , respectively. The sojourn time in each

state is exponentially distributed with the mean sojourn time in arrival state1 and 2 being μ_1^{-1} and μ_2^{-1} respectively.

Video traffic model: Video traffic typically requires large bandwidth. Therefore, outputs of video sources are usually compressed by using an inter-frame variable rate coding scheme which encodes only the significant differences between successive frames. This introduces a strong correlation among cell arrivals from successive frames^[16]. Like a voice sources, a video source generates correlated cell arrivals; however, its statistical nature is quite different from a voice source. Video sources with uniform activity level may be represented by the model proposed by^[17]. In this model, a video source is represented by a continuous-time, discrete state Markov chain. The bit-rate from a source is quantized into M discrete levels of step-size γ . The model switches between various levels spending exponentially distributed time in each level. As noted in^[16] the continuous-time, discrete state Markov chain may be constructed from the superposition of M mini-sources, where each mini-source is in one of two states: ON or OFF. When ON, it generates packets at a constant rate and when OFF, it does not generate any packets. Thus an ON-OFF characterization is given to the video traffic as well and following the same approach as in the case of voice, the superposition of video sources may also be approximated by a two state MMPP as proposed in^[15].

The studies of actual video conferencing traffic report that video frames (VFs) were found to be generated periodically and contained a very large number of cells in each frame^[15]. The distribution of the number of cells per VF was found to be described by gamma (or Pareto since it is self-similar traffic behavior) distribution. New VFs are assumed to arrive every 40 ms in state 1 (25 VFs per second) and 33.3 ms in state 2 (30 VFs per second). In^[17], the results of the analysis are that the digitized video transmission

exhibits a self similar character and that the frame length conforms to Pareto distribution. The variable length of the frames is due to the nature of the compression/encoding algorithm.

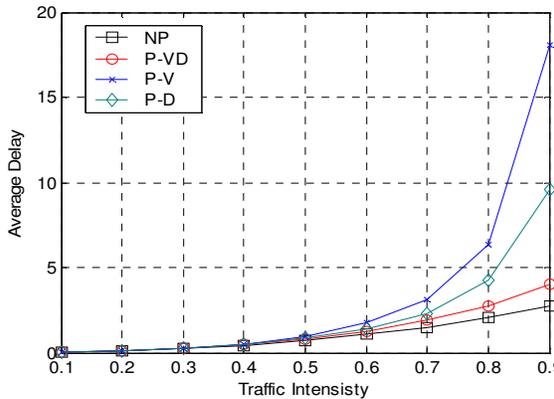


Fig. 7: Average Delay of all class with and without prioritization

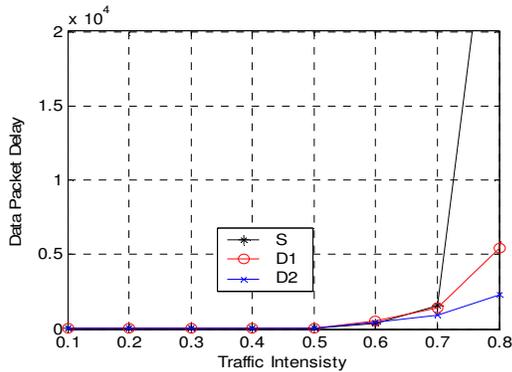


Fig. 8: Data average normalization delay using static (S) and dynamic (D1, D2) prioritization scheme

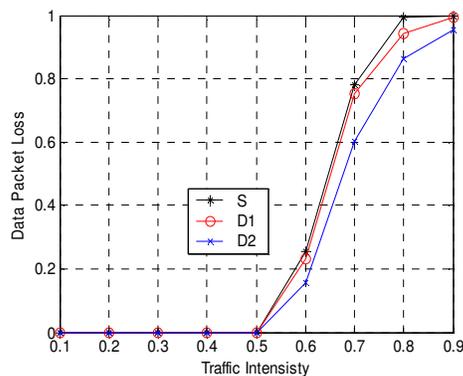


Fig. 9: Data loss probability delay using static (S) and dynamic (D1, D2) prioritization scheme

PERFORMANCE EVALUATION

In this section, we compare the performance of static priority scheme and the proposed dynamic priority scheme. First, we will present the simulation environments assumptions. Then, we will state the

performance measures used to evaluate each scheme. Finally, the simulation results will be presented and compared.

Simulation environment and parameters: The simulation program is written using visual basic language. All the simulation results reported in this project are based on the following assumption that the transmission rate of the output link is 1Gbps and the cell size is 64 byte. In the simulation, we also assume that the queue capacity of each traffic class is 100 cells. Other assumptions are shown in Table 2.

Performance measurements: The performance metrics estimated by this simulation are: (i) the probability of packets loss; and (ii) the normalized average packet delay. These parameters will be used in the presentation of numerical results and are defined as follows:

* *Packet loss probability (PL)*: this probability is the fraction of packets lost within the model:

$$PL = \frac{\sum Pkt\ Lost}{\sum Pkt\ arrived} \quad (3)$$

* *Average delay (Del)* : average packet delay is the time a packet spends from the time it arrives at the queue until the time it departs from it:

$$Del = \frac{\sum Pkt\ Delay}{(\sum Pkt\ arrived)} * slot\ time \quad (4)$$

Analysis of simulation results: In this section, we provide numerical results using the developed simulation to evaluate the system performance. The static and dynamic priority schemes will be compared in terms of packet loss rate and average delay for varying traffic intensity.

First we run the simulation without using the priority where all traffic has the same traffic class (as in Fig. 2) and then we run the simulation using the static prioritization scheme shown in Fig. 3. Model parameters used are shown above in Table 2. As shown in Fig. 5, using static priority scheme, the average delay for high priority video packets and second priority voice packets has decreased as compared to no-priority scheme. However, the low priority data packets experience higher average delay. So, the average delay of video and voice traffic decreases at the expense of small increase in the data traffic average delay.

As the above figures show, we can say that, static priority scheme is good at only one specific class at the expense of the others. The problem with the static priority is that as the higher traffic priority loads increases, the lower priority traffic would become increasingly congested and therefore dynamic priority scheme is needed. On the static priority considered above, the priority is not affected by the state of the system or the load applied. Dynamic priority considerations are concerned with priority schemes that are related to the system load and traffic characterizations. In the following dynamic scheme is

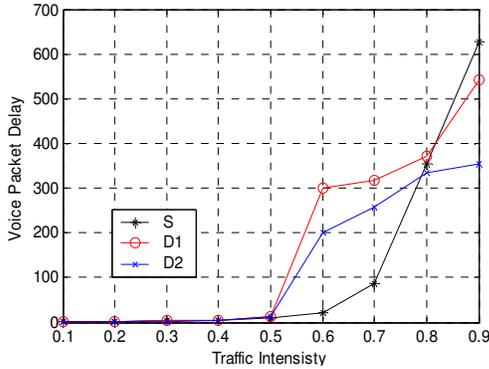


Fig. 10: Voice average normalization delay using static (S) and dynamic (D1, D2) prioritization scheme.

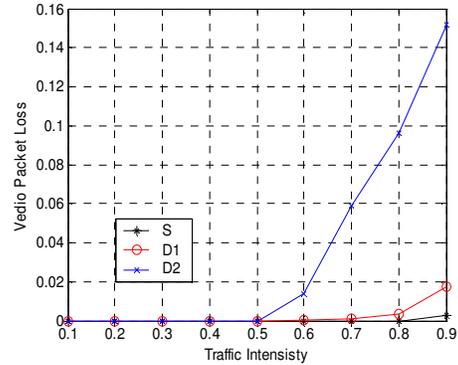


Fig. 13: Video loss probability using static (S) and dynamic (D1, D2) prioritization scheme

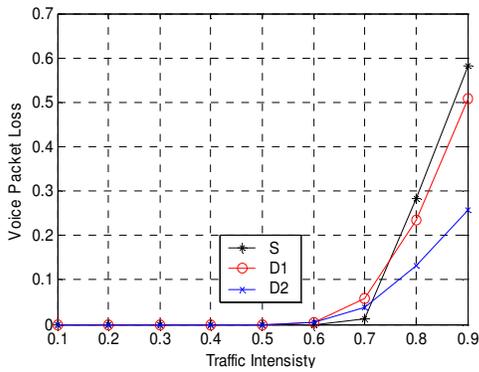


Fig. 11: Voice loss probability using static (S) and dynamic (D1, D2) prioritization scheme

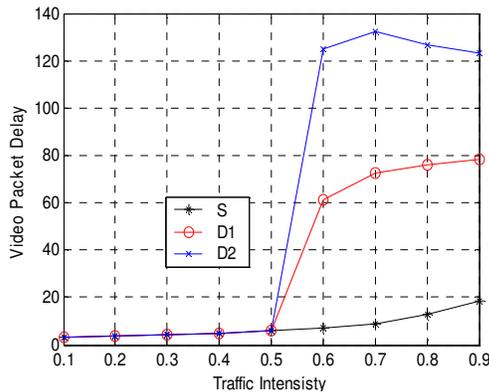


Fig. 12: Video average normalization delay using static (S) and dynamic (D1, D2) prioritization scheme

investigated and compared with the static priority scheme. As mentioned earlier, the dynamic prioritization decision is based on the value of the calculated function. The function has two main parts. The first part (TP(i)) is to maintain the traffic priority (TP(3) < TP(2) < TP(1)) factor. By this we make sure that the video traffic still has the highest priority then voice and then data. The second factor, TL(i), reflects the traffic loads of each class. This part will increase the value of the calculated function for specific class as its traffic loads increase. This part is used to avoid the

congestion issue. To compare the system performance under these types of schedulers, we take following steps. First we increase the loads of video traffic by 50% under static prioritization then we collect the performance measurements of each class. After that, under same condition we use the dynamic scheduling with two different configurations referred to as D1 and D2. In the First configuration (D1), we assume that the weighed factor (w) of each class traffic load (w(i)TL(i)) is unity and (TP(1)=0.7 , TP(2)=0.4 , TP(3)=0.2). In the second configuration (D2), we assume that the weighted factor of each traffic load w(i)TL(i) class is as follows (w(1)<w(2)<w(3) : 0.6:0.7:0.9). As shown in Fig. 8 and 9, the data traffic performance will improve under dynamic scheduling. This performance will increase as weighted factor of its traffic loads become larger than the higher priority class.

The same explanation applies to voice as shown in Fig. 10 and 11. The voice traffic will suffer from data traffic only at medium traffic load situation (0.6-0.7).

This improvement in performance for voice and data will at small expense of decreasing the performance of video traffic at high traffic conditions as shown in Fig. 10 and 11.

CONCLUSION

In this study, we developed appropriate traffic source model for each of three types of traffic. The appropriate traffic source model that reflects the correct behavior of each type of multimedia traffic was employed for subsequent simulation-based evaluation. In addition, we proposed a dynamic priority packet scheduling mechanism to efficiently serve the QoS requirements of multimedia traffic in packet switched networks. In packet switched networks, due to the effects of high speed channels and the efficient bandwidth allocation by statistical multiplexing, a large number of packets are allowed to enter the network. This may cause severe network congestion. Therefore, appropriate control of packet admission and congestion are in need. In particular, in order to support multiple

classes of traffic with varying delay and loss requirements, the priority mechanisms must be used to guarantee the required QoS.

For three classes of multimedia traffic, the proposed dynamic priority scheme schedules the most urgent class first by the priority of each class and the state of each queue determines by the number of packets in the queue and the available of queue size. In order to evaluate the performance of the dynamic priority scheme, its packet loss rates as well as the average delay are compared with the conventional scheduling schemes such as static priority scheme. According to our simulation-based evaluation, both the packet loss rates and average delays of dynamic priority scheme measured for each class are lower in most cases than those of the static scheme. The result also shows that while the static priority scheme is beneficial for only one specific class at the expense of the others, the dynamic serves fairly well all the classes in terms of delay and loss requirements.

ACKNOWLEDGMENT

The author acknowledges King Fahd University of Petroleum and Minerals, Dhahran, Saudi Arabia for support.

REFERENCES

1. Krishna, P., N.S. Patel, A. Charny, R.J. Simcoe, 1999. On the speedup required for work-conserving crossbar switches. *IEEE J. Select. Areas Commun.*, 17: 1057-1065.
2. Ajmone, M.M., P. Giaccone, E. Leonardi and F. Neri, 2003. Local scheduling policies in networks of packet switches with input queues. *IEEE INFOCOM*.
3. Ajmone, M.M., P. Giaccone, E. Leonardi and F. Neri, 2001. Packet scheduling in input-queued cell-based switches. *IEEE INFOCOM*.
4. Ajmone, M.M., P. Giaccone, E. Leonardi and F. Neri, 2002. Packet-mode scheduling in input-queued cell-based switches. *IEEE/ACM Tran. in Networking*, 10: 5.
5. GRF-multiGigabit routers. Product Overview, April 2000. Available : <http://www.lucent.com>.
6. Hung, G.K. and N. McKeown, 1998. ATM input-buffered switches with guaranteed-rate property. *Proc. 3rd IEEE Symp. Computers and Communications*, pp: 331-335.
7. Partridge *et al.*, 1998. A 50-Gb/s IP router. *IEEE/ACM Trans. Net.*, 6: 237-248.
8. Jeong T, J. Jeong and S.J. Kim, 1997. An efficient scheduling mechanism using multiple thresholds for multimedia traffic in ATM switching node. *IEEE Local Computer Networks*, pp: 263-268.
9. Heffes, H. and D. Lucantoni, 1986. A Markov modulated characterization of packetized voice and data traffic and related statistical multiplexer performance. *IEEE J. Selected Area in Comm.*, 4: 856-867.
10. Paxson, V. and S. Floyd, 1995. Wide-area traffic: The failure of poisson modeling. *IEEE/ACM Trans. Networking*, 3: 226-244.
11. McKeown, N., V. Anantharam and J. Walrand, 1996. Achieving 100% throughput in an input-queued switch. *Proc. IEEE INFOCOMM*.
12. Minagar, A.R. and S.M. Safayi, 2003. The optimized prioritized iSLIP scheduling algorithm for input-queued switches with ability to support multiple priority levels. *Proc. 10th IEEE Intl. Conf. on Telecommunications*, Feb. 23-Mar. 1, 2: 1680-1685.
13. Schonen, R., G. Post and G. Sander, 1999. Prioritized arbitration for input-queued switches with 100% throughput. *Proc. ATM Workshop*, May 24-27, pp: 253-258.
14. Yiping, G. and B. Liu, 2002. Performance evaluation of a parallel-poll virtual output queued switch with two priority levels. *IEEE*, pp: 669-674.
15. Heyman, T.V. Lakshman, A. Tabatabai and H. Heeke, 1994. Modeling teleconference traffic from VBR video coders. *Intl. Conf. Communications'94*, 3: 1744-1748.
16. Selvakumaran, N. and Tho Le-Ngoc, 1995. Traffic modeling in a multimedia environment. *IEEE CCECE/CCGEI'95*, pp: 838-841.
17. Garrett, M. and W. Willinger, 1994. Analysis, modeling and generation of self-similar VBR video traffic. *Proc., SIGCOMM 94*.
18. Zukerman, M., T. Neame and R. Addie, 2003. Internet traffic modeling and future technology implication. *IEEE INFOCOM*.
19. Choi, H. and J. Limb, 1999. A behavioral model of web traffic. *Intl. Conf. Networking Protocol 99, (ICNP 99)*.
20. Trajkovic, L. and S.J. Golestani, 1992. Congestion control for multimedia services. *IEEE Network Mag.*, pp: 20-26, Sep.