두개의 우선 순위를 가지는 고속 스윗칭 시스템의 설계 및 성능 분석

홍 요 훈† · 최 진 식†† · 전 문 석†††

요 약

기존 우선 순위 시스템에서는 우선 순위가 높은 패킷이 시스템에서 우선적으로 서비스를 받고 우선 순위가 낮은 패킷은 우선 순위가 높은 패킷이 없을 경우에만 서비스 받도록 되어있다. 그러나 입력 큐잉 시스템에서는 HOL(Head of Line)경쟁에 의해서 우선 순위가 높은 패킷이라도 차단 될 확률이 높다. 따라서 우선 순위가 높은 패킷이 차단됐을 경우라도 우선 순위가 낮은 패킷을 서비스 해 줌으로써 전체적으로 스윗칭 성능을 향상 시킬 수 있다. 본 논문은 고속 스위칭 시스템에서의 우선순위 기반 방식의 성능 분석을 하였다. 스윗칭 시스템 분석은 HOL(Head of line)경쟁 현상에 대한 우선순위 스케줄링이 미치는 영향을 고려 하였다. 또한 이러한 제어방식을 기반으로 시스템의 최대 처리율, 큐잉 분포 현상을 도출 하였다. 입력단 간에 서비스 의존도 때문에 스윗칭 시스템의 정확한 분석은 어려우나 상호 의존성을 갖는다는 가정과 흐름제어 규정을 두어 분석을 하였다. 각각의 입력단에서 보여주는 서비스 향상을 평가 하기위해 큐잉 시스템을 이용 하였다. 윈도우 방식을 고려하지 않고 우선순위 방식에서 정확한 결과를 구하기 위하여 Chen과Guerin[1]가 사용한 방식을 확장 하였다. 더욱이 시스템 구현과 운영 관점에서 우선순위 스윗칭 시스템에 적용하기 위하여 새로운 윈도우 제어방식을 제안한다. 그러므로 우선순위가 낮은 패킷은 지연시간과 처리율을 향상 시킬 수 있다. 성능 향상을 위해 결과치를 비교하여 등가 큐잉시스템을 사용하여 윈도우 방식을 분석 하였다.

Design and Analsis of a high speed switching system with two priority

Yo Hoon Hong† · Jin Seek Choi†† · Moon-Seog Jun†††

ABSTRACT

In the recent priority system, high-priority packet will be served first and low-priority packet will be served when there isn't any high-priority packet in the system. By the way, even high-priority packet can be blocked by HOL (Head of Line) contention in the input queueing System. Therefore, the whole switching performance can be improved by serving low-priority packet even though high-priority packet is blocked. In this paper, we study the performance of preemptive priority in an input queueing switch for high speed switch system. The analysis of this switching system is taken into account of the influence of priority scheduling and the window scheme for head-of-line contention. We derive queue length distribution, delay and maximum throughput for the switching system based on these control schemes. Because of the service dependencies between inputs, an exact analysis of this switching system is intractable. Consequently, we provide an approximate analysis based on some independence assumption and the flow conservation rule. We use an equivalent queueing system to estimate the service capability seen by each input. In case of the preemptive priority policy without considering a window scheme, we extend the approximation technique used by Chen and Guerin [1] to obtain more accurate results. Moreover, we also propose newly a window scheme that is appropriate for the preemptive priority switching system in view of implementation and operation. It can improve the total system throughput and delay performance of low priority packets. We also analyze this window scheme using an equivalent queueing system and compare the performance results with that without the window scheme. Numerical results are compared with simulations.

키워드 : QoS, 고속 스윗칭(high speed switching), HOL경쟁(HOL contention), 처리율(throughput), 윈도우(window)

1. Introduction

The advancement of high-speed semiconductors and optical technologies in the past decade has revolutionized the concept of computer communication networks. Particularly, the development of the Internet Packet Switching technologies will lead to BISDNs which offer a diversity of services for voice, data and high quality video.

To use the enormous potential of various traffics efficiently, new communications switches have to be developed to satisfy the demands of all classes of traffics. Moreover,

---

† 종신회원 : KLS연구소 소장
†† 정 회 원 : 정보통신 대학원대학교 교수
††† 종신회원 : 숭실대학교 정보과학대학 교수
논문접수 : 2001년 6월 28일, 심사완료 : 2001년 9월 10일

in order to provide services of differing requirements in a single network, the switching system should be capable of controlling the access to the network resources, according to the quality of service (QoS) demanded by the services.

One of the best methods to satisfy these requirements is the fast packet switching (FPS) technology based on priority scheduling. This can be provided with priorities in a switching system which accommodates, over a single switch fabric, real-time traffics with stringent delay requirements and data traffics which are less sensitive to delay. Such a technology is suitable for meeting the stringent performance requirements of different traffic types. Moreover, it should promise an integrated access structure with flexible allocation of bandwidth, capable of providing a broad range of future services as well as existing services.

FPS has been extensively studied in the literature. Good overviews on commonly used structures and available analysis methods can be found in [2] and [3]. Also, many papers have been devoted to the analysis of packet switches integrating a variety of traffics [4, 12]. However, most of previous analyses on FPSs use approximation techniques for switches operating with only a single priority class. In a few cases (e.g., [13-15]) where more than one priority class was introduced, the influence of the switch structure was not really taken into account. Although Chen and Guerin [1] derived the performance of a two-priority input queueing switch while taking into account of the influence of both switch structure and different classes of services, some of the results in this analysis were derived using a heuristic method.

Input Queuing scheme operates at link speed in input buffer and switching fabric. Therefore only one packet is forwarded to an output port and the others are stored in input buffer. This scheme has following benefits [16].

- Buffer memory bandwidth needs for storing packet is minimized (same as packet arrival rate of the single input link)
- Compared with its output buffering, packet loss rate of the bursty traffic is low.
- The amount of buffer required for processing multicast traffic is less than output buffering.

Howerver, it has a critical problem that the throughput will be limited to 58% due to HOL block phenomenon.

Providing QoS is another problem which has to be solved in input queueing Switch together with Improving Through put.

With switch architecture of window scheme, the packet arrived at input port is stored in arriving sequence order and packet in w(window size) is randomly transmitted at HOL location. For this, window scheme stores arriving packet orderly in the input buffer. The stored packet use input buffer structure selected and read by random order. Maximum w packet are paricipated in competetion in an input-port and only one packet can be transmitted to output port [17].

Arbitration of wN packet accomplished by sequential w degree competetions as follows. HOL packet of each input queue are participate in first competetion, and second packet of input queue unselected are participated in second competetion. Packet to be transmitted is determined by such a way without any conflicts in both input port and output port. By the way, the transmission order of packet with common destination in the same input-port is guaranted and only one packet is seleted in a input port.

The degree of freedom of packet-selection (maximum number of packet can be participated in the competetion per input-port) is w and random order transmission is possible. Hence performance is enhanced and reach at 100% in case of increasing the window size w [18].

In this paper, we propose a new enhanced model for the preemptive priority packet switch with an input queueing system to obtain a more accurate solution. Also, we propose a new window scheme appropriate for the preemptive priority packet switch with an input queueing system, which can improve the total throughput and reduce the packet delay of low-priority packets. In addition, we approximately analyze the overall performance of delay and throughput for both systems. This is done by extending some of the approximation techniques used in [1] and using an equivalent queueing model. Since the system model proposed in this paper is explicit for some mathematical manipulation without using any heuristic assumptions, the results obtained from our analysis are more accurate than that of [1].

The remainder of this paper is organized as follows. In Section 2, we describe the system structure and characterize the switch operation in terms of high- and low-traffic loads. Following this description, we propose queueing models for the preemptive priority packet switch with and without a window scheme, and briefly outline the method of analysis in Section 3. In Section 4 we present the derivation of probability of service seen by low-priority packets in the head of line (HOL) queue in detail. In Section 5, we obtain performance measures for low- and high-priority packets by using an equivalent queueing model. In this Section we describe computation of packet delay and maximum system throughput. Finally, in Section 6 we conclude and proposes possible extensions.

2. System Description

In this section, we describe the operation of an $N \times N$ preemptive priority packet switch having input queues with and without a window scheme considered. To describe the operation of the switch, we only focus on one of the input lines, which is characterized in terms of each priority class. At each time slot, the following events occur in sequence : arrival of packets at the input ports, contention among packets with the same destination when they reach at the HOL, and finally departure of the packets, which win contention. This contention is completed only when a packet successfully reaches its destination.

The arrival of each packet type is modeled by an independent Bernoulli process with λ_H and λ_L giving the probability of packet arrival for high- and low-priorities, respectively. The packets queued in input port contend for their outputs according to the preemptive priority contention procedure. For example, if high priority packets exist in the input queue, then they first contend for their outputs without consideration of low-priority packets at the same port. On the other hand, low-priority packets have to take care of the presence of high-priority packets at the same input ports as well as the their output ports. For example, if a high-priority packet exists, a low-priority packet in the same input port loses the chance to contend for its output port. Although the low-priority packet can have a chance to contend for its output, it is blocked at the output if there are high-priority packets coming from other input queues.

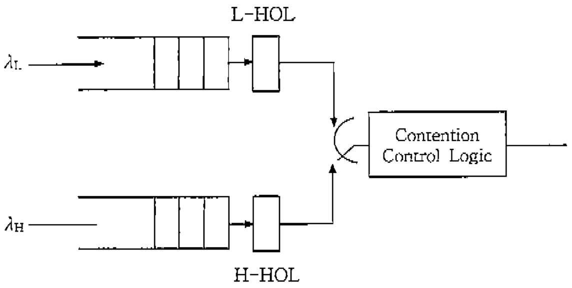
Such a preemptive priority policy can be very advantageously used for reducing the delay of high-priority packets. Nevertheless, a significant problem which must be solved for this architecture is to ensure that the switch does not suffer from the limitations of total throughput and delay of the low-priority packet. Note that, the low-priority packet is prevented from accessing an available output even when the high-priority packet is blocked due to the problem of *head of line blocking*. It limits considerably the chance to contend for low-priority packets. Therefore, for low-priority packets, the waiting time in the input queue to participate the HOL contention due to the high-priority packets is added to the contention time to access the output port. Consequently, it limits the total throughput up to 0.6, even when the uniformly distributed case is considered. In case of uniform but geometrically distributed bursty traffic, the maximum throughput can be further degraded to 0.5 [12].

In order to overcome these deficiencies, we may apply a window scheme to this preemptive priority input queueing system. This is a relaxation of the strict priority discipline of input buffers. high-priority packets at the head of the input queues contend first for access to their outputs. Low-priority packets may access their outputs even when high-priority packets exist at the same input queues, if the high-priority packets are blocked for their output contention. Consequently, the proposed window scheme can permit a request for the low-priority packets to contend for its output in the line blocked by a high-priority packet which lost contention (i.e., the high-priority packets are blocked in this time slot and outputs are not yet assigned to receive packets in this time slot.).

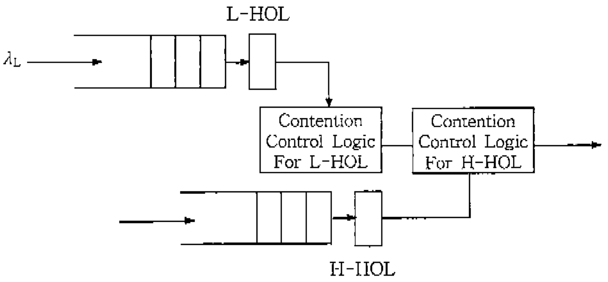
The proposed window scheme is implemented in the input queue as follows. The input queues as well as the HOL queues are divided into two independent queueing systems : one is used for low-priority packets and the other is used for high priority packets. Particularly, the HOL queue for high- and low-priority packet type is called the high-priority HOL queue (H-HOL) and the low-priority HOL queue (L-HOL), respectively. Moreover, each contention phase is also divided in two subsequent phases. First one is the contention phase only for high-priority packets. The second one is the contention phase for low-priority ones. In the first phase, only high-priority packets contend for their outputs. In the second phase, low-priority packets contend for the output only when the high-priority packet at the H-HOL is blocked in the first contention phase or the H-HOL is empty. Since high-priority packets contend for their outputs only in the first phase, they are not interfered with low-priority packets. Since the second phase is used only for low-priority packets, the low-priority packets are not interfered with the high-priority attempts which are blocked during the first phase. Moreover, they do not interfere with the high-priority requests granted during the first phase because of the difference of the priority level.

As can be seen from the above discussion, at each time slot low-priority packets can obtain more chance to contend for packet transmission comparing to that of the pure preemptive priority scheme. Although it is not truly the same as the conventional window scheme, it can also improve the maximum throughput of the switch.

(Figure 1) and (Figure 2) show the HOL queue structure

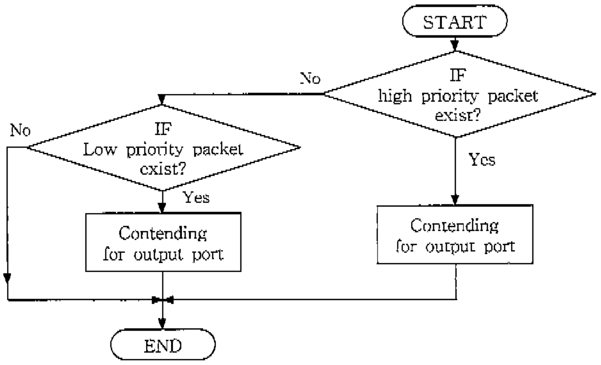


(a) HOL contention logic without a window scheme

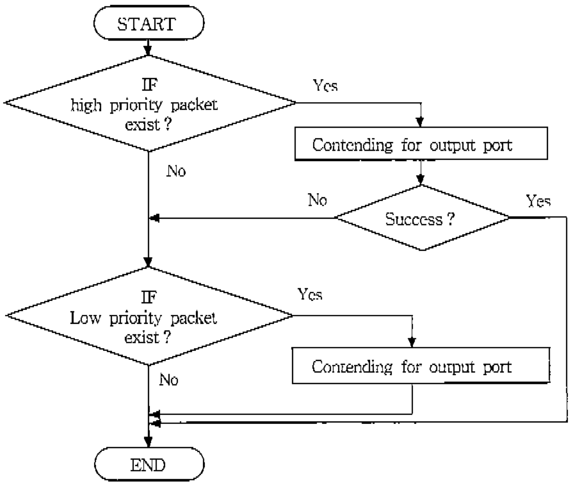


(b) HOL contention logic with a window scheme

(Figure 1) HOL queue structure



(b) HOL contention logic without a window scheme



(b) HOL contention logic with a window scheme

(Figure 2) HOL contending procedure

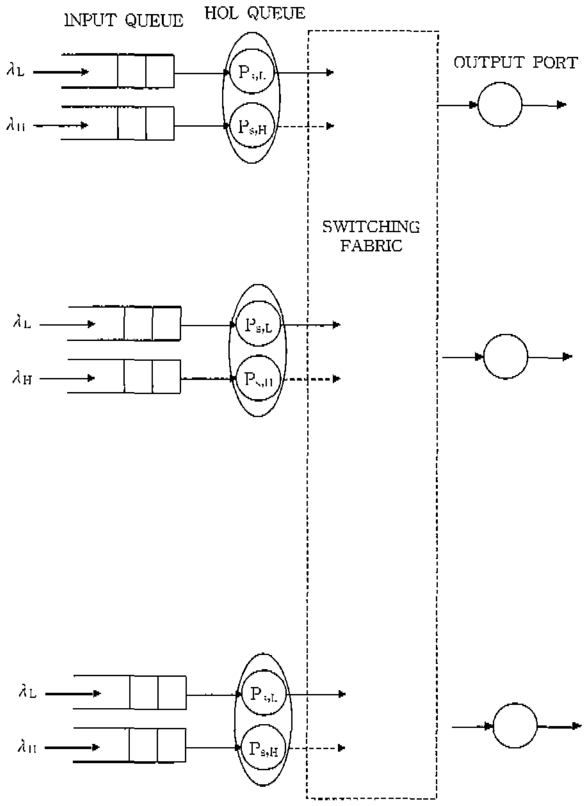
and the contending procedure with and without the window scheme. From now on, we refer to the priority handling procedure without the window scheme as the scheme A and to the procedure with the window scheme as the scheme B. In both the schemes, queued packets are served on a FCFS basis within each class.

3. SYSTEM MODEL

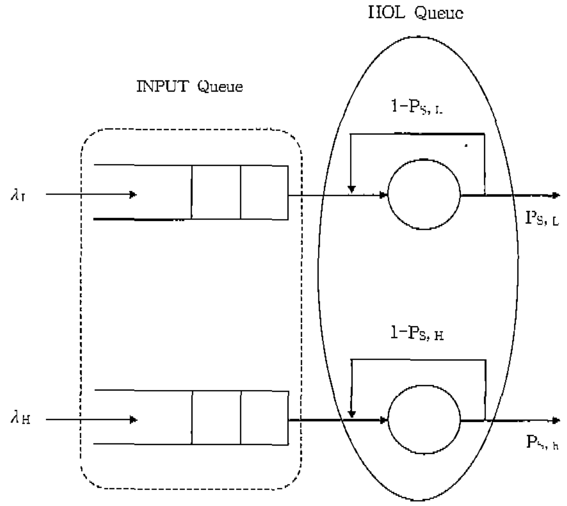
The switching system can be modeled as a system of N parallel queues (the input queues) with N feedback servers (head of line queue) contending for the outputs. Each feedback server is modeled as a geometric server with an effective service time distribution which accounts for HOL contention.

Since the state space associated with this system increases rapidly with N, an exact solution is intractable. Therefore, an approximate approach will be used, which may be summarized as follows. At first, we separate N independent queues for each input port. Moreover, the HOL queueing systems are also divided into two independent HOL systems for high- and low-priority packets, respectively. (Figure 3) and (Figure 4) show the queueing model of the overall system and individual input port. Packets of each HOL are contending for their outputs according to the conventional preemptive priority policy. In this HOL queueing model, we only use the number of packets contending for their outputs using the state probability of HOL packets at a given slot according to the contention policy with and without the window scheme.

The analysis method used in this paper is similar to the approximation method used in [14]. The analysis is done in two parts : each HOL service rate, P_s is determined as a function of λ_H and λ_L, and then the individual input queues modeled by a $Geom/Geom/1$ queueing system are analyzed. In order to obtain P_s, we first study "virtual" queues seen by the HOL packet, which are formed by HOL queues (i.e., the H-HOLs for high-priority packets and L-HOLs for low-priority packets). These virtual queues can be modeled as a *single server queue with geometric service time* [4]. The service time for this virtual queue is assumed to have a geometric distribution, which has a different value depending on the priority class and the HOL contention policy. Based on the information obtained from these virtual queues and the law of flow conservation, we derive the equivalent service probability for high- and low-priority packets under some independence and limit assumptions.



(Figure 3) Overall system model



(Figure 4) Input queue model

For the high-priority packets, the service probability at this virtual queue, $P_{s,H}$, depends only on the amount of high-priority packets contending for the same output (called output contention). In this case, the presence of low-priority packets is transparent to it. Consequently, the average service time at H-HOL is directly deduced from the analysis of the single priority packet switch [4].

On the other hand, to obtain the service probability for the low-priority packets, $P_{s,L}$, we must consider the presence of high-priority packets at the input and output ports as well as the amount of other low-priority packets contending for the same output port. According to the cell scheduling at the HOL queue, we classify the service probability into two : $P^A_{s,L}$ for scheme A and $P^B_{s,L}$ for scheme B. The difference of these two schemes is only the service probability of the low-priority packet when the high-priority packet at the same port is blocked for its output contention. This effect is reflected in the difference of $P^A_{s,L}$ and $P^B_{s,L}$.

These service probabilities mean how often, on the average, packets get served at the HOL queue. From this, we can approximate each input queue as an independent *Geom/Geom/*1 queueing system where the service time can be derived from the above virtual queue for the given packet type. The performance measures for delay and throughput are then obtained from this equivalent queueing system.

Before starting analysis, we make some assumptions about the switch model. The arrival statistic in this queueing model is assumed such that the aggregate traffics of high- and low-priority packets are mutually independent and uniformly distributed for each destination with Bernoulli rate of λ_H and λ_L, respectively. Furthermore, it is assumed that successive packet arrivals are independently destined from those of previous arrivals. As in High Speed Switching, we also assume that packets are of fixed length and the switch operates synchronously, such that, at each time slot, every input/out put can receive/transmit one packet. Therefore, packets are received at different switch input ports and transmitted to different switch output ports in the same time interval called a slot. Both high- and low-priority input queues are assumed to be infinite and to operate on the FIFO basis.

4. Equivalent Hol Service Probability

In this section, we derive the equivalent service probability at HOL, or service rate, for each priority class. This equivalent service rate is a measure of how many slots a packet takes for the first time for the HOL position until successful delivery to its destination. This time can also be viewed as the packet service time at the virtual queue modeled before. Ideally, the switch is capable of delivering one packet from each input at every time slot. However, because of the presence of high-priority packets at the input queue and the contention with other packets for the outputs, the HOL packet typically takes more than one slot to effectively go through the switch. More specifically, the successful transmission of each packet requires that the conditions listed below be met.

(1) For a high priority packet
- it occupies the H-HOL position of input queue.
- it is selected (randomly) among all high-priority packets contending for the same output.

(2) For a low-priority packet under scheme A
- it occupies the L-HOL position of input queue.
- A high-priority packet does not exist at the same input port.
- There are no high-priority packets contending for the same output port at other input queues.
- it is selected (randomly) among all low-priority packets contending for the same output.

(3) For a low priority packet under scheme B
- it occupies the L-HOL position of input queue.
- A high-priority packet does not exist at the H-HOL position, or it has to be blocked in the first contention phase.
- There are no high-priority packets contending for the same output at other H-HOL positions.
- it is selected (randomly) among all low-priority packets contending for the same output.

4.1 HOL service probability for high-priority packets

From the point of view of service conditions mentioned before, the service probability depends only on other high-priority packets destined for the same output. Consequently, the virtual queue is modeled as a HOL queueing system, which is the same as the HOL system of a conventional nonblocking packet switch. In [4], the equivalent service rate at the HOL position is given by

$$P_{s,H} = \frac{2(1-\lambda_H)}{2-\lambda_H}. \qquad (1)$$

Here, $1/P_{s,H}$ is the expected number of slots which contend for high-priority packets to go through the switch. Based on this result, the steady-state probabilities of queue size P_k is deduced [4]. It is given by

$$P-k = \begin{cases} (1-w)(1-\lambda_H), & i=0 \\ (1-w)w^{i-1}, & i>0 \end{cases} \qquad (2)$$

where $w = \lambda_H^2[2(1-\lambda H)^2]$. The complete queue length distribution of high-priority packets is used to compute the HOL service probability for low-priority packets.

4.2 HOL service probability for low-priority packets

In this subsection, the HOL service probability, $P_{s,L}$ seen by low-priority packets can be obtained from the same derivation tool used in [1], with a few modifications.

Comparing to high-priority packets, the service conditions for low-priority packets are more complicated, which are listed above. In other words, the computation of $P_{s,L}$ should take into account of the following additional properties. First, the number of packets contending for the same output in successive slots are correlated since no more than one packet can be delivered to one output in each slot. Second, at each input, high-priority packets make the switch unavailable to low-priority packets for a period of k kconsecutive slots for scheme A or 1 slot for scheme B, where k is determined by the duration of a busy period generated by the high-priority packet at H-HOL, and 1 slot for scheme B means the successful transmission time of the high-priority packet.

In the section below, the derivations of $P_{s,L}^A$ and $P_{s,L}^B$ are carried out in detail for scheme A and B, independently. Before starting the analysis, we first define random variables and notations used for our performance analysis as the following :

$H_j^{(i)}$: Number of H-HOL packets destine for output j in slot i
$Y_j^{(i)}$: Number of L-HOL packets destined for output j before contention phase in slot i
$L_j^{(i)}$: Number of L-HOL packets destined for output j that participate output contention in slot i
$Q_j^{(i)}$: Number of L-HOL packets destined for output j that are not delivered at the end of slot i.

Since the system normally operates below saturation, the average low-priority packet throughput T_L per output equals the input arrival rate λ_L and can be expressed in terms of H_j and L_j as

$$T_L = \lambda_L = \frac{1}{N}\sum_{j=1}^{N} E[\in(L_j)\overline{\in}(H_j)], \qquad (3)$$

where $\in(x)$ is the indicator function of $\{x \in \mathcal{R}, x > 0 \in\}$ and $\overline{\in}(x) = 1 - \in(x)$. Another quantity of interest is the number of L-HOL packets that are left over at the end of a contention phase. Using the fact that all outputs are identical, the expected value of Q can be related to the expected value of Y_j and the average output throughput T_L as

$$E[Q]=NE[Y_j]-NT_L \tag{4}$$

We now introduce two additional intermediate variables, M and ρ. At a random slot, M, $0\leq M\leq N$, denotes the number of L-HOL positions available for low-priority packets, and gives the probability that an input has a new low-priority packet waiting to move into an available L-HOL position. Input ports are excluded from M only if they have low-priority packets that already attempted to go through the switch. The quantity is simply the probability that an input port has a nonempty queue of new low-priority packets, given that it is among the M available ones. These two new variables can be related by the *flow conservation principle* of the L-HOL system as follows.

$$E[M]\rho=N\lambda_L \tag{5}$$

This equation states that the total arrival rate must be equal to the flow into the from non-empty L-HOL positions.

It can also be seen that $E[M]$ is the expected number of low-priority packets that fail to go through the switch in a time slot, and is equal to the difference between the total number of L-HOL and $E[Q]$. Consequently, $E[M]$ can be expressed as

$$E[M]=N-E[Q]. \tag{6}$$

By substituting (5) and (6) into (4), we obtain as follows :

$$E[Y_j]=\lambda_L+1-\frac{\lambda_L}{\rho} \tag{7}$$

Consequently, if we can obtain another expression about $E[Y_j]$ in terms of ρ and offered load, λ_N, then ρ can be expressed as the offered load. Furthermore, it can be used to determine $P_{S,L}$.

4.2.1 Scheme A

One may note that in order to obtain $E[Y_j]$, we must know the relationship between $E[L_j]$ (the expected number of L-HOL packets destined for output j), and (the expected number of L-HOL packets contending for output j). For scheme A, the low-priority packets at L-HOL can contend for their output only if there are no high-priority packets in the same input queue. Therefore, $E[L_j]$ can be expressed as

$$E[Lj]=E[Yj]\cdot P_0^H, \tag{8}$$

where $P_0^H=\mathrm{P_r}(K_H=0)$ means the probability that there is no high-priority packet at the input port, which is given in (2). Moreover, $E[L_j]$ was derived in [14] as

$$E[L_j]=\frac{2\lambda_L-\lambda_L^2}{2(1-\lambda_H-\lambda_L)} \tag{9}$$

Substituting (8) and (9) into (7), we obtain an expression involving and known quantities as

$$\frac{2\lambda_L-\lambda_L^2}{2(1-\lambda_H-\lambda_L)P_0^H}=\lambda_L+1-\frac{\lambda_L}{\rho} \tag{10}$$

In order to obtain the desired expression for $P_{s,L}^A$, we must use the *flow conservation principle* of the departure rate at the L-HOL queue as follows. The mean number of slots before the service of a fresh L-HOL packet is $1/P_{s,L}^A$. After a packet is served, the mean number of slots before the arrival of a fresh L-HOL packet is $(1-\rho)/\rho$. In the steady state, the sum of these two terms equals the interarrival time $1/\lambda_L$ of packets at an input port. From this relationship, we get

$$\frac{1}{\lambda_L}=\frac{1}{\rho}-1+\frac{1}{P_{s,L}^A} \tag{11}$$

Using (2), (10), and (11), we now obtain the desired expression for $P_{s,L}^A$:

$$P_{s,L}^A=\frac{(\lambda_H^2-4\lambda_H+2)(1-\lambda_H-\lambda_L)}{(2-\lambda_L)(1-\lambda_H))} \tag{12}$$

4.2.2 Scheme B

The main difference between scheme A and B is the HOL contention capability of low-priority packets. We only derive this relationship. In scheme A, a high-priority packet can disrupt the service of low-priority packets at L-HOL for several consecutive time slots because the high-priority packet not delivered at the first attempt continues to occupy the H-HOL position until it is successfully delivered. In scheme B, however, a high-priority packet disrupts the service of a low-priority packet only in the time slot of its successful delivery. Even when high-priority packets exist in the same input queue, low-priority packets can contend for their outputs if high-priority packets at the same ports are blocked during the first contention phase. Therefore, the low-priority packet at the L-HOL may have more chance to contend to go through the switch than with scheme A. Consequently, the relation between the number of low-priority packets at the L-HOL and the number of low-priority packets contending for their output becomes

$$E[L_j]=E[Y_j]\cdot P_0^H+E[Y_j](1-P_0^H)(1-P_{s,H}), \quad (13)$$

where $(1-P_0^H)(1-P_{s,H})$ means the probability that the high priority packet exists but it is blocked. By substituting (13) into (7) and manipulating in the same way as in scheme A, we have

$$\frac{2\lambda_L-\lambda_L^2}{2(1-\lambda_H-\lambda_L)(1-P_0^H)(1-P_{s,H})}=\lambda_L 1+-\frac{\lambda_L}{\rho}. \quad (14)$$

From the *flow conservation principle* of the L-HOL queue, we obtain $P_{s,L}^B$ as follows.

$$P_{s,L}^B=\frac{2(1-\lambda_H)(1-\lambda_H-\lambda_L)}{(2-\lambda_L)}. \quad (15)$$

5. ANALYSIS OF OVERALL PERFORMANCE

In this section, we investigate the overall performance of the preemptive priority packet switch with and without a window scheme. We first obtain the delay performance using a *Geom/Geom/*1 queueing model [4], and investigate the maximum throughput achievable by the switching system.

5.1 Queueing Delay

We first analyze the average queue length for high- and low-priority packets under scheme A and B. They can easily be obtained using the *Geom/Geom/*1 queueing model as follows.

For high-priority packets, since the presence of low-priority packets is transparent to the performance of high-priority packets, the queue length distribution is the same as the single-priority system. Therefore, performance measures for high-priority packets are available from the previous works on single-priority switches [9] and [4]. In this paper, we only consider the performance of low-priority packets.

Once $P_{s,L}$ is obtained, the delay performances for low-priority packets are directly deduced from the standard queueing theory. The reason is that the effect of high-priority packet preemption is considered in the value of $P_{s,L}$. Particularly for scheme A, the presence of high-priority packets affects the service time of low-priority packet at the same input port in the following two ways. At first, the high-priority packet preempts the chance of L-HOL contention, which increases the service time of a low-priority packet. Secondly, when low priority packets contend for their outputs, the high-priority packet contended from other H-HOL ports affects the acceptance of the low-priority packet. Both influences are accounted for the quantity $P_{s,L}^A$, which represents the number of time slots needed before the low-priority packet can be delivered by the switch.

For scheme B, on the other hand, the presence of a high-priority packet affects the low-priority packet only in the case that a high-priority packet wins HOL contention or the output is occupied by the high-priority packet delivered from other input. These effects are also accounted for the quantity $P_{s,L}^B$. In addition, the effect of other low-priority packets is also considered in the value of $P_{s,L}^B$.

Consequently, low and high-priority packets arrive with independent probability and at each time slot and receive service independently with a fixed probability $P_{s,L}^A$ and $P_{s,L}^B$ per time slot, for low-priority packet under scheme A and B, respectively. Therefore, the average queue length for low-priority packets can directly be obtained from the *Geom/Geom/*1 queueing model using $P_{s,L}^A$ and $P_{s,L}^B$ for scheme A and B, respectively.

Based on the values of $P_{s,L}^A$ and $P_{s,L}^B$ derived above, the *Geom/Geom/*1 queueing model can be solved using a standard queueing method (e.g., [4]). Therefore, we find that the generating function $G(z)$ of the steady-state queue length K is

$$G(z)=\frac{(P_{s,L}-\lambda)(1-\lambda+\lambda_z)}{P_{s,L}(1-\lambda)-\lambda(1-P_{s,L})z} \quad (16)$$

From the above equation, we can obtain several quantities of interest. The first one is the average queue length $E[K]$ of low-priority packets,

$$E[K]=G'(z)|_{z=1}=\begin{cases}\dfrac{\lambda_L(1-\lambda_L)}{P_{s,L}^A}, & \text{for scheme A}\\ \dfrac{\lambda_L(1-\lambda_L)}{P_{s,L}^B}, & \text{for scheme B}\end{cases} \quad (17)$$

The expected delay $E[D]$ can also be directly deduced from (17) using Little's formula as follows.

$$E[D]=\frac{E[K]}{\lambda_L}=\begin{cases}\dfrac{\lambda_L(1-\lambda_L)}{P_{s,L}^A}, & \text{for scheme A}\\ \dfrac{\lambda_L(1-\lambda_L)}{P_{s,L}^B}, & \text{for scheme B}\end{cases} \quad (18)$$

The queue length distribution is also obtained by the inversion of the generating function in (16) as

$$P_r(K=i)=\begin{cases}(1-a)(1-\lambda), & i=0\\ (1-a)(1-\lambda)a^i+\lambda(1-a)a^{i-1}, & i>0\end{cases} \quad (19)$$

where a is given by

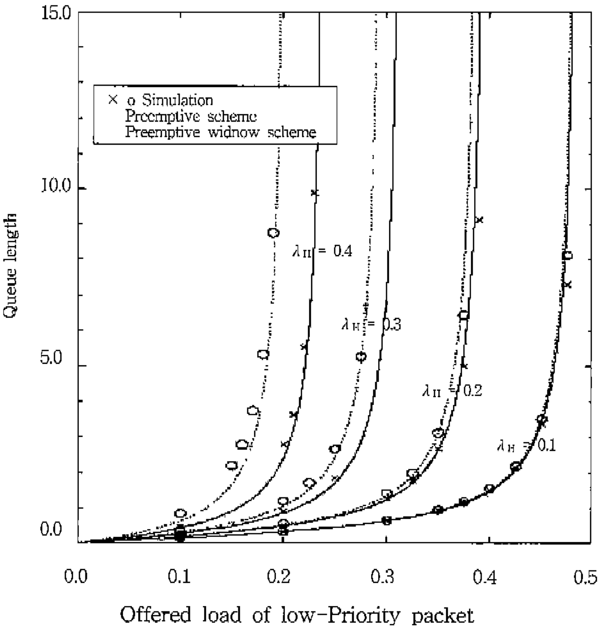
$$a = \begin{cases} \dfrac{\lambda_L(1-P^A_{s,L})}{[P^A_{s,L}(1-\lambda_L)]}, & \text{for scheme A} \\ \dfrac{\lambda_L(1-P^B_{s,L})}{[P^B_{s,L}(1-\lambda_L)]}, & \text{for scheme B} \end{cases} \tag{20}$$

The performance results provided by this analysis are presented in (Figure 5)~(Figure 7) Particularly, (Figure 5) shows a comparison of the mean queue lengths and (Figure 6) shows the average delay time under scheme A and B. (Figure 7) shows the queue length distribution of low-priority packets in a given traffic condition of λ_H 0.2 and λ_L 0.3 which shows the same trend as (Figure 6). The accuracy of the analytical results is also shown by providing the results obtained through computer simulation of several system configurations. For simplicity, we assumed the switch size of 64×64 in our simulation.

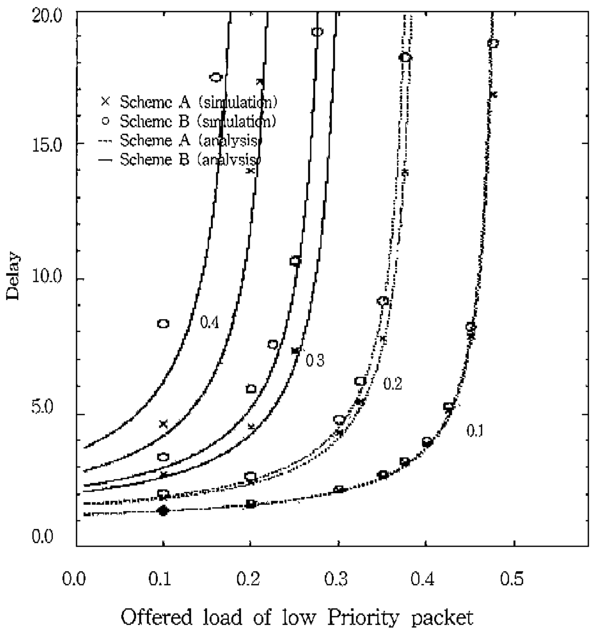
As shown in these figures, we get good agreements between analysis and simulation. It indicates that the analysis techniques used in this paper are valid for the given system.

5.2 System throughput

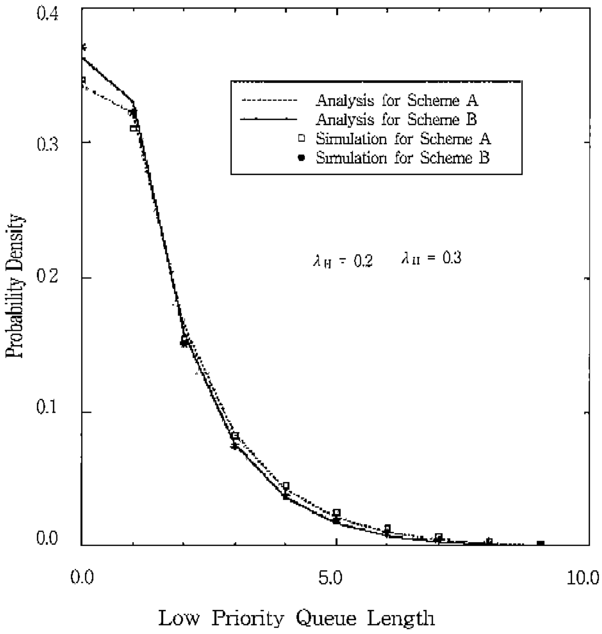
We now investigate the maximum throughput of the system under scheme A and B. We first obtain a stability criterion which gives the maximum high- and low-priority loads that can be simultaneously supported by the system. The maximum throughput corresponds to the maximum admissible input load under which the system remains stable. For high-priority packets, the maximum throughput has been limited up to 0.586, the same value as for the single-priority switch of Hui and Arthurs [4]. However, the maximum throughput for low-priority packets can be found as an expression that a system remains stable under the given arrival rate of high-priority packets. This relation can be obtained from the parameter ρ, which was defined as the probability that there is a new lowpriority packet waiting to move into an available HOL position. The value of ρ can never exceed unity. Therefore, setting $\rho=1$ in (10) gives the maximum throughput of the low-priority packets, $\lambda_{\max,L}(\lambda_H)$



(Figure 5) Mean Queue length Low-priority packet



(Figure 6) Average waiting time for Low-priority packets



(Figure 7) A Queue length distribution of Low-priority

For scheme A, the resultant value of the maximum throughput is given by

$$\lambda_{\max,L}(\lambda_H)=\frac{\lambda_H^2-6\lambda_H+4-\sqrt{-3\lambda_H^4+12\lambda_H^3-16\lambda_H+8}}{2[1-\lambda_H]} \qquad (21)$$

where $0 \le \lambda_H \le 2-\sqrt{2} = 0.586$. This equation gives $\lambda^A_{\max,L}(0)=0.586$ and $\lambda^A_{\max,L}(0.586) \fallingdotseq 0$ which is in good agreement with the result for single-priority switches. Moreover, this maximum throughput for low-priority packets can also be obtained from the stability condition of the equivalent queueing system, $P_{s,A}=\lambda^A_{\max,L}$ These two results are identical. We shows $\lambda^A_{\max,L}$ as a function of λ_H in (Figure 8).

Consequently, the total system throughput $T^A_{\max}(\lambda_H)$ as a function of the offered loads can also be obtained from (21) as follows.

$$T^A_{\max}(\lambda_H) \le \lambda^A_{\max}(\lambda_H)+\lambda_H, \quad 0 \le \lambda_{H \le 0.586} \qquad (22)$$

This total system throughput under scheme A is illustrated in (Figure 9). It can be seen that $T^A_{\max}$ starts at 0.586 for $\lambda_H=0$ and increases to a maximum of about 0.607 for $\lambda_H \cong 0.45$, after which it decreases to reach again 0.586 for $\lambda_H=0.586$.

Under scheme B, the maximum achievable throughput of low-priority packets, $T^B_{\max}$, can be obtained in the same way as for scheme A. By setting to one, we can get.

$$T^A_{\max}(\lambda_H)=2-\lambda_H-\sqrt{2-\lambda_H^2}. \qquad (23)$$

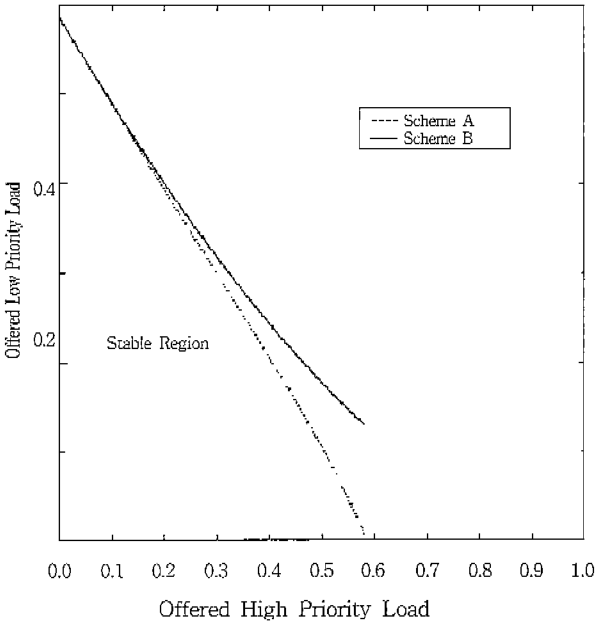
This maximum throughput for low-priority packets can also be obtained from the stability condition of the equivalent queueing system, $P_{s,B}=\lambda^B_{\max,L}$. The resultant value for $\lambda^B_{\max,L}$ is plotted, together with $\lambda^A_{\max,L}$, in (Figure 8) as function of λ_H.

The total system throughput $T^B_{\max}$ under scheme B is given by

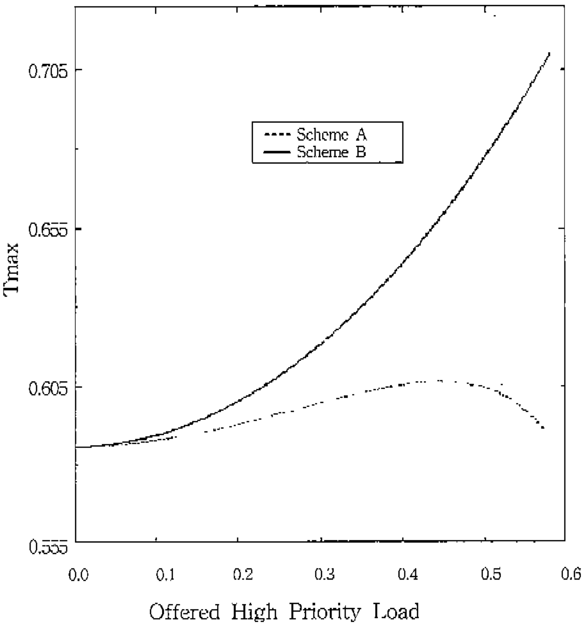
$$T^B_{\max}(\lambda_H) \le \lambda^B_{\max}(\lambda_H)+\lambda_H, \quad 0 \le \lambda_H \le 0.586 \qquad (24)$$

(Figure 9) also shows the total system throughput under scheme B. As shown in this figure, the maximum system throughput of scheme B is much higher than that of scheme A, because the throughput of the low-priority packets does not degrade with the traffic load of the high priority packets.

In the recent priority system, High priority packet will be served first and low priority packet will be served when there isn't any higher priority packet(in the system)



(Figure 8) Offered High Priority Load vs Offered Low Priority Load



(Figure 9) Maximum of Low-priority packet as a function of high priority packet load

By the way, even high priority packet can be blocked by HOL(Head of Line) contention in the Input queueing System.

Therefer the whole switching performance can be improved by serving low priority packet even though high priority packet is blocked.

This indicates the importance of the service policy not only in selecting between competing inputs as studied in [1], but also in arbitrating between packets from the same input queue. Such a priority window scheme is further extended to have the maximum throughput up to 1 with ω contention phases which corresponds to the "windowing system" for ω w priority classes. Packets with ω priority classes are allowed to contend for the switch sequentially until the first packet is successfully delivered.

6. Conclusion

In this paper we presented performance results for a preemptive priority packet switch with and without a window scheme. The main contribution of the paper is the derivation of performance measures for low-priority packets that take into account of the influence of priority scheduling and window policy for HOL contention.

Moreover, we analyzed this scheme with an equivalent queueing system for each packet class using an approximation technique. With the service discipline rule for the HOL queue, we proposed two methods based on the service capability of low-priority packets. The first method is the preemptive priority policy without considering a window scheme, which is the same as the conventional preemptive priority policy. The second method is the preemptive priority policy with a window scheme. This window scheme can improve the total system throughput and delay performance of low-priority packets.

In the analysis, we used some new procedures as follows. At first, we separated the input queue for each priority class. Second, the HOL queueing systems were also divided into two independent HOL systems for high- and low-priority packets, respectively. In this HOL queueing model, we only use the number of packets contending for their outputs using the state probability of HOL packet at a given slot according to the contention policy with and without the window scheme.

Although this model characterizes accurately the average behavior of the HOL queueing system, it neglects the correlation effect of consecutive time slots. For example, when the L-HOL contention does not exist at a given time slot because of nonempty high-priority packets, it is more likely to be non-existent in the next time slot, resulting in more "bursty" behavior than reflected in the "random server" model used here. Consequently, a new preemptive priority policy based on the window scheme would give higher performance than the results presented here. Therefore, further works are needed to consider these effects of correlated and unbalanced input traffics and to derive packet delay performance for these traffics.

References

[1] J. S.-C. Chen and R. Guerin, "Performance study of an input queueing packet switch with two priority classes," *IEEE Trans. Commun.*, Vol.39, pp.117-126, Jan. 1991.

[2] H. Ahmadi and W. E. Denzel, "A survey of modern high-performance switching techniques," *IEEE J. on Selected Areas Commun.*, Vol.SAC-7, No.7, Sept. 1989.

[3] Y. S. Hui, switching and Traffic Theorg for Intefrated BroadBand Network, Kluwer Academic publisher, Boston, 1990.

[4] J. Y. Hui and E. Arthurs, "A broadband packet switch for integrated transport," *IEEE J. Select. Areas Commun.*, Vol.SAC-5, pp.1264-1273, Oct. 1987.

[5] I. Iliadis, "Performance of a packet switch with input and output queueing under unbalanced traffic," *in Proc. INFOCOM '92*, pp.743-752. June, 1992.

[6] J. N. Giacopelli,W. D. Sincoskie andW. Littlewood, "Sunshine : a high performance self routing broadband packet switch architecture," *in Proc. XIII International Switcing Symposium, '90*, Stockolm Sweden, May, 1990.

[7] H. Suzuki et al., "Output-buffer switch architecture for asynchronous transfer mode," *in Proc. ICC '89*, pp.99-103 June, 1989.

[8] M. Yamamoto, H. Tode, H. Okada and Y. Tezuka, "A control-ahead ATM switch architecture and its performance," *IEEE J. on Selected Areas Commun.*, Vol.SAC-9, No.9, Dec. 1991.

[9] M. G. Hluchyj and M. J. Karol, "Queueing in high-performance packet switching," *IEEE J. Select. Areas Commun.*, Vol.SAC-6, pp.1587-1597, Dec. 1988.

[10] M. Karol and M. G. Hluchyj, "Using a packet switch for circuit-switched traffic : A queueing system with periodic input traffic," *IEEE Trans. Commun.*, Vol.37, pp.623-625, June, 1989.

[11] M. Karol, M. G. Hluchyj, and S. P. Morgan, "Input versus output queueing on a space-division packet switch," *IEEE Trans. Commun.*, Vol.35, pp.1347-1356, Dec. 1987.

[12] S. Li and M. J. Lee, "A study of traffic imbalances in a fast packet switch," *in Proc. INFOCOM '89*, 1989, pp.538-547.

[13] Y. Lim and J. Kobza, "Analaysis of a delay-dependent priority discipline in a multi-class traffic packet switching node," *IEEE Trans. Commun.*, Vol.38, pp.659-665. May, 1990.

[14] A. Iyengar and M. E. Zarki, "Switching prioritized packets," *in Proc. GLOBECOM '89*, pp.1181-1186, 1989.

[15] A. Y.-M. Lim and J. A. Silverster, "Priority queuing strategies and buffer allocation protocols for traffic control at an ATM integrated broadband switching system," *IEEE J. on Selected Areas Commun.*, Vol.SAC-9, pp.1524-1536. Dec. 1991.

[16] G. Kesides and N. Mckeown, "output buffered ATM packet switching for Integrated services communication Networks," IEEE ICC '97, 1997.

[17] Y. K. Park and G. Lee, : NN based ATM Cell Scheduling

with Queue length_Based priority schene. IEEE JSAC. Vol.15, No.2, pp.261-270, Feb. 1997.

[18] Adisak Mekkittikul, and Nick McKeown, "A practical scheduling Algorithm to Achieve 100% Through put in input-Queued switches," IEEE infocom '98, 1998.

[19] A. K Gupta and N. D Georgans, "Priority performance of ATM packet switches," in Proc. IINFOCOM '92, pp.772-733, June, 1992.

[20] K. W. Sarkies, "The bypass queue in fast packet switching," IEEE Trans. Commun, Vol.39, pp.766-774, May, 1991.

[21] A. Y. M. Lim and J. A. Silverster, "Priority queucing strategies and buffer allocation protocols for traffic control at an ATM intefrated broadband switching system," IEE J. on Selected Areas Commum, Vol.SAC-9, pp.1524-1536, Dec. 1991.

홍요훈

e-mail : yhhong@klsc.co.kr
1981년 인하대학교 전자공학과 학사
1981년~1985년 LG 전자연구원
1985년~1994년 ETRI 연구원
1994년 서강대학교 대학원
1994년~1998년 한국통신 연구개발원
2001년 숭실대학교 컴퓨터 학과 박사과정
1998년~현재 KLS연구소 소장
관심분야 : 인터넷 QoS, 고속 라우터 기술

최진식

e-mail : jin@ns.icu.ac.kr
1985년 서강대학교 학사
1987년 한국과학기술원 석사
1987년~1991년 LG 정보통신연구원
1995년 한국과학기술원 박사
1995년~2001년 공주대 교수
1998년~2000년 NIST 방문연구원
2001년~현재 한국정보통신대학원대학교 교수
관심분야 : 스위칭 시스템,인터넷 QoS, 고속 라우터

전문석

e-mail : mjun@computing.ssu.ac.kr
1980년 숭실대학교 전자계산학과 졸업(학사)
1986년 University of Maryland 전산과 졸업(석사)
1989년 University of Maryland 전산과 졸업(박사)
1989년 Morgan State University 전산수학과 조교수
1989년~1991년 New Mexico State University 부설 Physical Science Lab. 책임연구원
1991년~현재 숭실대학교 정보과학대학 부교수
관심분야 : 컴퓨터 알고리즘, 병렬처리, VLSI 설계, 암호학