

# 배치 정렬기에 기반한 ATM 스위치의 큐잉 기법 모델링 및 성능분석

박 병 섭<sup>†</sup> · 임 철 수<sup>††</sup>

## 요 약

본 논문에서는 내부 스피드업을 갖는 배치-베니안 망 유형의 스위칭 구조에 관한 분석적 모델을 제안하고 성능분석 결과를 제시한다. 또한 이 모델에서 다양한 큐잉 기법들이 다른 트래픽 분포하에서 비교된다. 높은 스위칭 스피드업을 갖는 배치 베니안 망에서, 입력 및 출력 트래픽 분포에 영향을 받는 버퍼링 기법의 성능은 전체 스위칭 망의 성능에 커다란 영향을 주게된다. 따라서 본 고에서는 제안된 모델링 기법을 사용하여 베니안 망들의 큐잉 기법을 분석하며, 큐잉 기법을 통하여 큐잉 성능과 입력 트래픽의 관계를 규명한다.

## Modeling and Performance Analysis of Queueing Mechanisms for Batcher Sorter-based ATM Switches

ByoungSeob Park<sup>†</sup> · CheolSu Lim<sup>††</sup>

## ABSTRACT

We propose an analytical modeling scheme and present a performance analysis of switching structure in the Batchersorterbased networks with internal speedup. Each queueing scheme is analyzed under different traffic distributions. In the Batchersorterbased networks with higher switching speedup, the performance of buffering techniques which is dependant upon input and output traffic distributions affects the performance of total switching system. Therefore, we investigate queueing schemes for several Batchersorterbased networks using our proposed analytical modeling techniques, and address both input traffic and associated queueing performance issues using queueing scheme.

### 1. Introduction

As ATM(Asynchronous Transfer Mode) switches have emerged as a core technique for high speed packet switching in the B-ISDN(Broadband-ISDN), numerous ATM switching architectures have been proposed [1-2]. The major goals of ATM switches are to implement cost-eff

ective switching fabric and to provide high bandwidth services for users. In realizing a B-ISDN services using ATM switches, the general physical connection between input and output ports within the switch fabric can be implemented through a time or space-division methodology. First, the use of a physical resource is multiplexed among several input-output connections, based on discrete slots. Among various spacedivision types of architectures, our particular interest is on the Batchersorterbased

† 정 회 원 : 국방정보체계연구소 선임연구원

†† 정 회 원 : 서경대학교 컴퓨터공학과 교수

논문접수: 1998년 4월 20일, 심사완료: 1998년 6월 9일

hing fabrics because of its ability of self routing and internal nonblocking capability [3]. Batcher-banyan switching scheme delivers exactly one packet per time slot to each outlet requested by one of the winner inlets.

Recently, it has been demonstrated that moderate switch speedup, in conjunction with parallel banyan and appropriate queueing scheme, provides a remarkable throughput-delay characteristic [4]. However, the acceleration of the entire switch fabric is still considered to be undesirable since switching fabric with input queueing, in spite of their simplicity, have a limited throughput by HOL(Head-Of-Line) blocking. Switches with output queueing provide optimal throughput-delay performance, but they require the speedup of switching operation in order to transmit packets arriving on all inlets into their corresponding output buffers within one time slot. For an  $N \times N$  switch, the internal fabric has to operate  $N$  times faster than the I/O trunks. However, for sufficiently large  $N$ , such a full speedup can not be achieved for ATM switches. In this regard, the switching fabrics with partial speedup are emerging [4-6]. In order to achieve partial speedup, most of this switches employ  $K$ -parallel banyan or serial connection structure of banyans. Accordingly, the output ports of these switches take a form of multiple outlets [7], and output buffering. In the switches with multiple outlets, the performance of the switching system such as throughput and latency time affects the entire system performance, and is closely coupled with the traffic distribution that is the packet arrival rate at each output link destined for a given outlet.

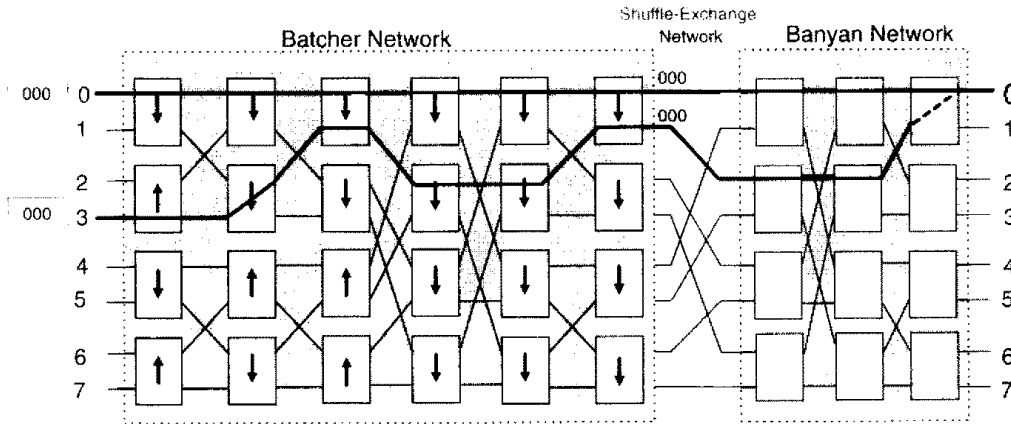
Most of current studies for the performance of Batcher-banyan networks have been mainly focused on the analysis of buffering scheme under random and uniform traffic[8-9]. However, the assumption of uncorrelated destination is

not so realistic all the time. Some works deal with the performance of the switch queueing with possibly correlated arrivals appearing in ATM services [4,8], but they analyzed the performance through a simulation method only. Also, while the general studies on switch architectures and queueing schemes have been carried out, only a few specific researches on the input and output traffic distribution that is tightly related to the performance of the queueing scheme have seldom carried out. In this paper, we propose analytical models for various buffering techniques under uniform and bursty traffic patterns. Furthermore, we attempt to capture the major properties of traffic in various queueing schemes of the Batcher-banyan networks based on our proposed analytic modeling.

The organization of the rest is as follows: Section 2 conveys the overview of the Batcher-banyan switches and queueing schemes. In Section 3, we discuss the traffic distribution. Section 4 analyzed buffering schemes for each Batcher-banyan network under uniform and bursty traffic patterns. Finally Section 5 carries some concluding remarks.

## 2. Batcher-banyan Networks

One of the drawbacks of the ordinary banyan network is that they are internally blocking in the sense that two packets destined for two different outlets may be collided at one of the intermediate switching elements. However, if packets are sorted based on destination address in advance and then routed through the banyan network, the internal blocking problem can be resolved. This is the basic idea behind the Batcher-banyan network [1]. It consists of a Batcher network(Batcher-sorter) which sorts the packets according to their destination address, followed by a shuffle exchange network and a ba



(Fig. 1) An example of output contention in the Batcher-banyan

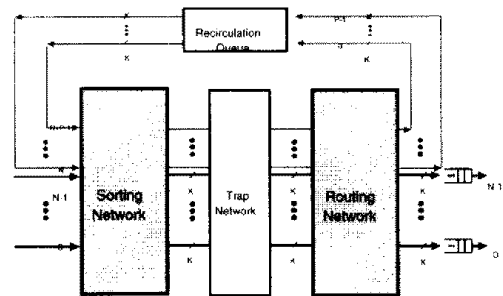
banyan network which routes the packets. However, as is shown in Figure 1, the Batcher-banyan network does not contribute to any improvement in throughput when multiple requests are destined for the same outlet.

2.1 The Starlite switch

Figure 2 illustrates the simple switch model with speedup  $K$  and shared-output queue. The Starlite switch [10] is the first case of implementation which employs the Batcher-banyan self-routing structure. It can be obtained by the general  $(N+P) \times (N+P)$  ( $P$ =number of recirculation queue input) switch structure with mixed shared-output queueing shown in Figure 2, by assuming  $K=1$ . In particular, the Starlite has both input and output queue, but we consider only shared-output queue for consistency of analysis. To overcome the output contention problem, the Starlite approach use a trap network between the sort and banyan network, which detects packet with the same destination addressed at the output of sorting network. If multiple packets are destined for the same output port, the extra packets are injected into the sorting network again on the next switch cycle through recirculation buffer size is  $P \times P$ .

2.2 The Sunshine switch

The Sunshine [5], like the Starlite, also uses a Batcher-banyan combination to make the switch nonblocking, and a recirculation mechanism to reduce the packet loss rate. However, one distinctive feature of the Sunshine switch design is that  $K$ -multiple banyan networks are used in parallel to transmit at most  $K$  packets to the same outlet. Thus, the  $K$ -multiple banyan is employed to decrease the packet recirculation rate, and to achieve a certain degree of output buffering. In Figure 2, if the  $K$  is not equal to 1, this switch structure becomes that Sunshine.



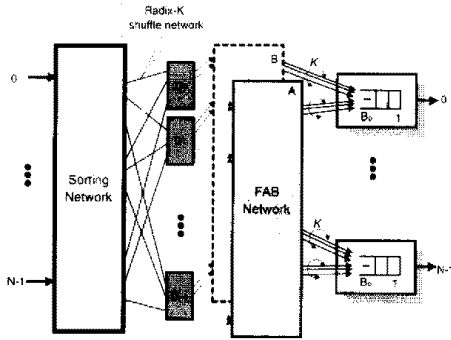
(Fig. 2) The switch architecture with speedup  $K$

2.3 The FBSF switch

The FBSF(Fat Banyan Switching Fabric) [6, 12] switch is a fast packet switching network with output buffering. Unlike the Sunshine switch with multiple outlets, where  $K$ -parallel banyan networks are used to deliver more than

one packet to each outlet, the FBSF switch uses only two FAB(Fat Banyan) networks [11] resulting in a speedup of 2K.

As is shown in Figure 3, the FBSF consists of a sorting network, packet distributors(D), two FAB networks, and output queue. The FAB network is a unified design approach based on the full utilization of switch bandwidth in a diluted banyan network. Dilution network [11] comprises the switching elements with constant link dilation and achieves enough throughput. In FAB network, each port of a switching element has multiple input and output links.



(Fig. 3) The FBSF switch architecture with speedup 2K

Moreover, the number of input and output links per port may not be equal, such an SE is called Fat SE(FSE). In FAB architecture, the number of input and output links would grow in the first few stages. However, in the remaining stages, the number of links would remain fixed or even decreased. This is based on the observation that in multibuffered-banyan networks, performance improvement has been achieved for buffer sizes up to four. In the FBSF switch, if more than 2K packets are destined for the same outlet, the packet distributor randomly selects 2K packets for delivery, and rejected packets should retry in the next time slot. The packet selection scheme is deactivates the packet at higher input port among input ports that have the same destination addresses.

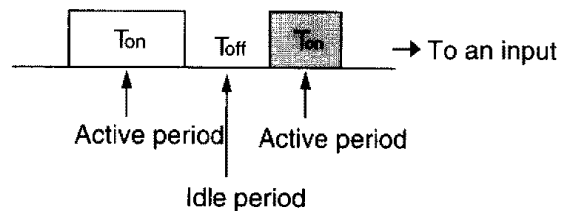
### 3. Classification of Traffic Patterns

#### 3.1 Uniform Traffic Distribution

For the uniform traffic, we model the packet arrival patterns for the N inlets based on independent and identical Bernoulli processes [3]. That is, in any given time slot, the probability that a packet would arrive on a specific input is  $p(0 \leq p \leq 1)$ . The  $p$  also corresponds to the input traffic load  $\rho$ .

#### 3.2 Bursty Traffic Distribution

In a bursty traffic type, the traffic of each input is characterized by bursty packet arrivals. Here the packet arrivals consist of bursts to different destinations. Within each burst, packets with a common destination arrive continuously in a stream. Such traffic can be modeled as a discrete-time ON-OFF Markov chain [9], as is shown in Figure 4.



(Fig. 4) ON-OFF model for packet arrivals

The durations of an active and idle states are both geometrically distributed with parameters  $\alpha$  and  $\beta$ , respectively. We assume that there is at least one packet in each active burst and each burst length is statistically independent. The probability that active period lasts for a duration of  $i$  time slot is then  $A(i) = \alpha(1-\alpha)^{i-1}(i \geq 1)$ . The probability that idle period lasts for  $j$  time slot is  $B(j) = \beta(1-\beta)^j(j \geq 0)$ . Unlike the duration of an active period, the duration of an idle period can be 0. Given  $\alpha$  and  $\beta$ , the mean burst length  $E\{B\}$ , the mean idle(separation) time  $E\{I\}$ , and the normalized

offered load  $\rho$  are given as

$$E[B] = \sum_{i=1}^{\infty} i \cdot A(i) = \frac{1}{\alpha} \cdot E[I] = \sum_{j=0}^{\infty} j \cdot B(j) = \frac{1 - \beta}{\beta} \cdot \rho$$

$$\rho = \frac{E[B]}{E[I] + E[B]} = \frac{\beta}{\beta + \alpha - \alpha \beta} \quad (1)$$

#### 4. Performance Analysis

##### 4.1 Performance under Uniform Traffic

###### (1) Analysis of Output Queueing

Under the independent and uniform input traffic, every outlet should have the same distribution for queue length. To analyze the performance of output queueing, we adopt the FBSF switch model with speedup  $2K$ . The basic assumptions and definitions associated with the FBSF switch are in [6,12]. For simplicity, we choose the unit of time to be the length of a time slot, and assume infinite output buffer space. Consider a tagged output port  $i$ . Defining the random variable  $A$  as the number of packet arrivals destined for the tagged outlet in a given time slot. We have

$$a_k \equiv \Pr[A=k] = \binom{N}{k} \left(\frac{\rho}{N}\right)^k \left(1 - \frac{\rho}{N}\right)^{N-k}, \quad k=0, \dots, N \quad (2)$$

as  $N \rightarrow \infty$ , it become

$$a_k \equiv \Pr[A=k] = \frac{\rho^k e^{-\rho}}{k!}, \quad k=0, 1, \dots, \infty \quad (3)$$

where  $\rho$  is the offered input load. Let  $M_t^i$  denote the total number of packets in the tagged queue at the end of  $t$ -th time slot and  $\psi_{2K}(M_t^i)$  denote the value of  $M$  during the  $t$ -th time slot. Also, let  $\mu(M)$  be the number of packets served in the given time slot. Since the maximum number of packet served in the given time slot are one, we have  $\mu(M) = \min(1, M)$ . Hence, the two random variables,

$M_t^i$  and  $M_{t+1}^i$  are related each other by the equation

$$M_{t+1}^i = M_t^i - 1 + \mu(M) + \psi_{2K}(M_t^i) \quad (4)$$

At first, to evaluate the throughput and waiting time, we obtain the moment generating function  $g_{M_t^i}(z)$  of  $M_t^i$ . It is easily verified that there is no correlation between  $M_t^i$  and  $M_{t+1}^i$  as  $N \rightarrow \infty$ . Consequently, the moment generation function is given by

$$g_M(z) \equiv E[z^M] = E[z^{M-\mu(M)}] \cdot E[z^{\psi_{2K}(M)}] \quad (5)$$

Define  $\omega_m \equiv \Pr[\psi_{2K}(M) = m]$  for  $m = 0, 1, \dots, 2K$ , since the number of packets arriving at tagged output can not exceed  $2K$ . It is self-evident that,  $\omega_m = a_m$ , for  $m = 0, 1, \dots, 2K-1$

and  $\omega_{2K} = 1 - \sum_{m=0}^{2K-1} a_m$  for  $m = 2K$ . As a consequence,

$$E[z^{\psi_{2K}(M)}] = z^{2K} - \sum_{m=0}^{2K-1} a_m (z^{2K} - z^m).$$

In equilibrium, the output trunk has the same idle probability, i.e.,  $\Pr[M = 0] = 1 - \rho$ . Thus, we obtain

$$g_M(z) = \frac{(1-\rho)(z-1)(z^{2K} - \sum_{m=0}^{2K-1} a_m (z^{2K} - z^m))}{z - z^{2K} + \sum_{m=1}^{2K-1} a_m (z^{2K} - z^m)} \quad (6)$$

Differentiating (6) with respect to  $z$ , and setting  $z = 1$ , we get  $g'_M(1)$ . Now, we can calculate the average queue length  $E[L]$ . Since the mean number of packets in services is  $\mu = \rho h$  (where,  $h$  is the service time), the average queue length is given by

$$E[L] = g'_M(1) - \rho \quad (7)$$

By Little's formula in (3), we obtain the average waiting time as

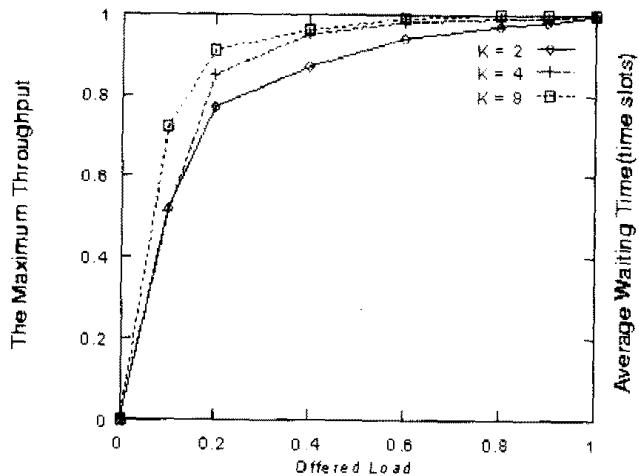
$$E[W] = E[L]/\rho = g'_M(1)/\rho - 1 \quad (8)$$

Since the service rate of output queue is  $\mu \leq 1$ , the maximum throughput can be determined by fixing  $\rho = \mu$ .

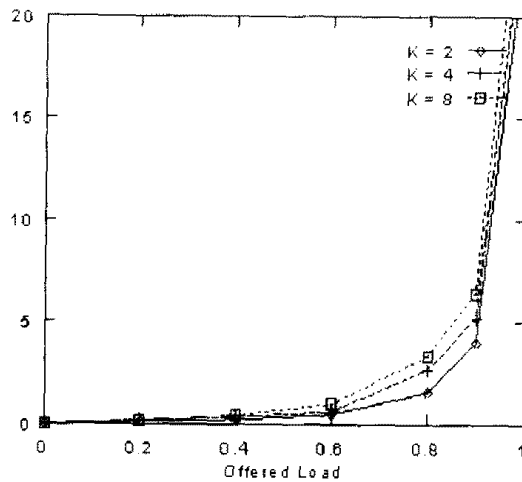
###### (2) Analysis of Shared-output Queueing

For the case of  $K = 1$ , the analysis for

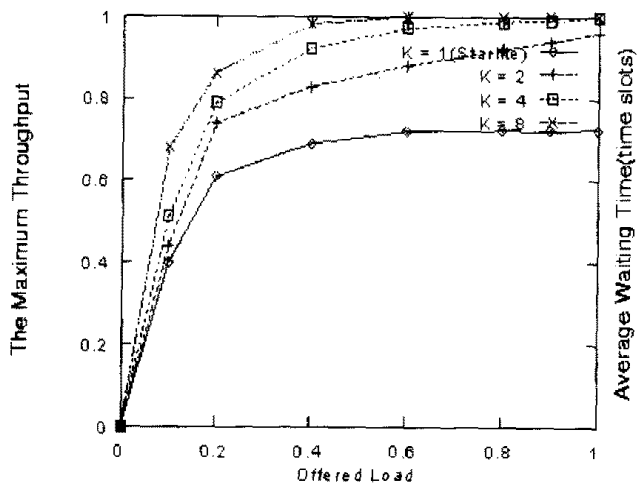




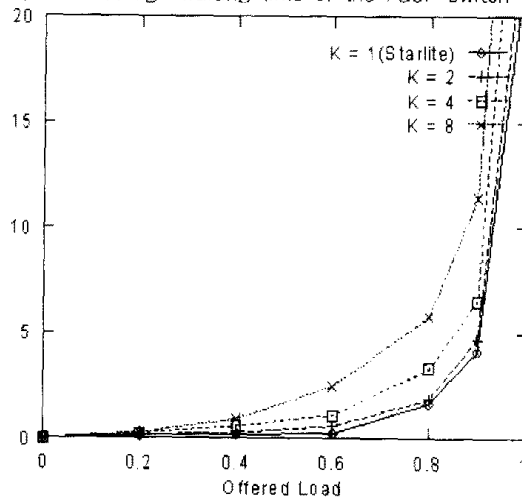
(Fig. 5) The maximum throughput of the FBSF switch



(Fig. 6) The average waiting time of the FBSF switch



(Fig. 7) The maximum throughput of the Sunshine and Starlite



(Fig. 8) The average waiting time of the Sunshine and Starlite

average waiting times  $E(W)$  are below the 2 time slots when  $0 \leq \rho \leq 0.65$ .

In Figure 7, the maximum throughput of shared-queuing schemes varies according to

offered load. Next, Figure 8 shows that the number of outlets  $K$  affects the performance of output queue.

4.2 Performance under Bursty Traffic

In the case of bursty traffic, we analyze the switch architecture with speedup  $K$  for the  $i$ th outlet. For infinite queue, let the total number of packet arrivals destined for output  $i$  in  $t$ th time slot be  $B_i^t$ . Then, we obtain

$$B_i^t = B_i^t(1) + B_i^t(2) + \dots + B_i^t(N) \quad (11)$$

where  $B_i^t(m)$  is the number of arriving packets from input  $m$  destined for that particular outlet  $i$  at the beginning of time slot  $t$ , and  $N$  is network size. Therefore, the random variable  $B_i^t(m) (1 \leq m \leq N)$  is assumed to have independent and identical distribution. For analysis, we will use the following definitions:

- $s(k) = \text{Pr}\{\text{burst-size} = k\}$
- $\lambda_B = \text{Pr}\{\text{there is a bulk arrival}\}$
- $E\{S\} = \text{the average of } s(k)$
- $\sigma^2 = \text{Var}(s)$  is variance of burst size
- $\rho = \lambda_B E\{S\}$  is traffic load density

With the assumption of the burst arrivals each input in a time slot, we have

$$b_k = \text{Pr}\{B_i^t(m) = k\} = \begin{cases} (1 - \lambda_B) + \frac{N-1}{N} \lambda_B, & k=0, \\ \frac{\lambda_B}{N} s(k), & k \neq 0. \end{cases} \quad (12)$$

In (12), for the case of  $k=0$ , the first equation

denotes the probability of no burst arrival in input  $m$ ,  $1 - \lambda_B$  plus the probability that there is a burst arrival, but the burst is not destined for the given outlet  $i$ . For the  $k \neq 0$ ,  $b_k$  is simply the probability that there is a burst arrival at input  $m$  destined for the outlet  $i$ . Thus, the moment generating function for  $B_i^t(m)$  is as follows:

$$B_i^t(m)(z) = 1 - \frac{\lambda_B}{N} + \frac{\lambda_B}{N} S(z) \quad (13)$$

Since  $B_i$  is the sum of i.i.d.  $B_i^t(m)$ , we have

$$B_i^t(z) = [1 - \frac{\lambda_B}{N} + \frac{\lambda_B}{N} S(z)]^N \quad (14)$$

At first, by evaluating the average queue length and waiting time under bursty traffic, we can obtain  $g_{B_i}$  as follows:

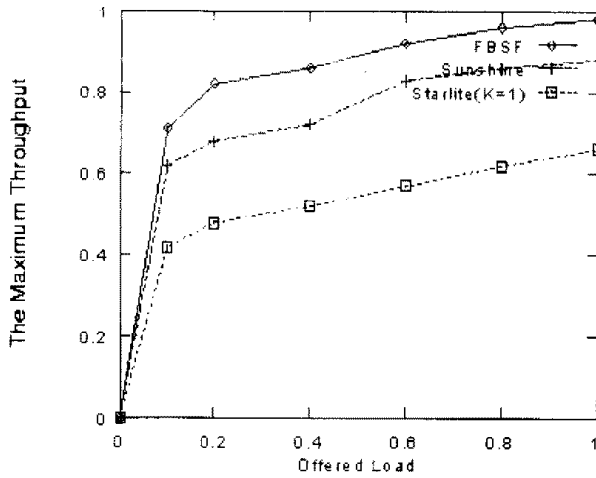
$$g_{B_i}(z) \equiv E[z^{B_i}] = E[z^{B_i - \mu(B_i)}] \cdot E[z^{\delta_B(B_i)}] \quad (15)$$

While the FBSF switch can transmit  $2K$  packets to the particular outlet and the Sunshine can deliver up to  $K$  packets to the given outlet, the Starlite switch can transmit only one packet to the given outlet. Thus, let  $\omega_j \equiv \text{Pr}\{\delta_B(B) = j\}$  for  $j = 0, 1, \dots, s (s = 1, K, 2K)$ .  $\omega_j = b_j$  for

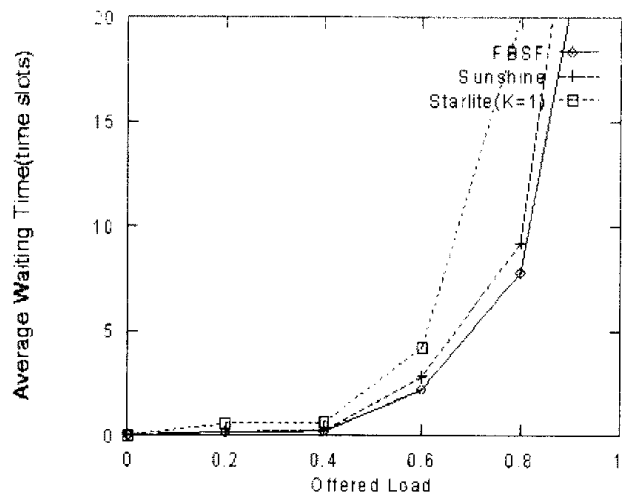
<Table 2> Summary of the numerical results for each switch(bursty traffic)

	Load	FBSF	Sunshir	Starlite
Grouping Size(radix-4)		K=4	K=4	K=1
Max. Throughput	0.2	0.82	0.68	0.48
	0.4	0.86	0.72	0.52
	0.6	0.92	0.83	0.57
	0.8	0.96	0.862	0.59
	1.0	0.98	0.88	0.64
Avg. Waiting Time	0.2	0.18	0.19	0.58
	0.4	0.21	0.24	0.64
	0.6	2.2	2.81	4.2
	0.8	7.8	9.2	20.1
	0.9	17.4	18.1	≈ inf.
1.0	≈ inf.	≈ inf.	≈ inf.	





(Fig. 9) Throughput vs. Offered load (under bursty traffic)



(Fig. 10) Waiting time vs. Offered load (under bursty traffic)

$j = 0, 1, \dots, s-1$ , and  $\omega_s = 1 - \sum_{j=0}^{s-1} b_j$ , for  $j = s$ . Consequently, we can obtain average queue length,  $E[z^{\delta_H(B)}] = z^s - \sum_{j=0}^{s-1} b_j (z^s - z^j)$  and  $g'_H(1)$  using  $b_j$  instead of  $a_m$  in equation (6).

(1) Numerical results under bursty traffic

The following graphs and Table 2 illustrate the numerical results of performance analysis for the Batchier-banyan network with multiple outlets under bursty traffic. For  $N = \infty$ , average burst length  $E(B)=10$ , Separation(idle)  $E(I)=20$ , and  $K=4(K=1$ , in Starlite), the throughput and average waiting time are shown in Figure 9 and Figure 10. Figure 9 plots the relationship between throughput and the offered load when  $K=4$ . Figure 10 shows that the FBSF switch is better, and that of the Sunshine employing the shared-output queueing follows. Thus, our analytical results indicate that the throughput under bursty traffic is lower than those under uniform traffic as is depicted by Figure 5 and Figure 7. This is mainly due to the fact that bursty traffic condition makes it likely that the next packet has the same output

address. However, we can perceive that the performance of the FBSF switch is hardly affected the traffic patterns. Figure 10 shows the packet latency and all the values are within 2 units of time slot when  $\rho \leq 0.45$ .

5. Conclusions

We have so far analyzed the performance of the Batchier-banyan networks with multiple outlets for the various input traffic patterns. For the analysis and the evaluation of ATM switch architectures, we have built analytic model for the switch performance under various queueing schemes. Moreover, we have examined the traffic distribution according to the number of outlets  $K$ . The switching networks with internal speedup used for analysis are the FBSF, the Sunshine, and the Starlite switch. The numerical analyses have shown that the FBSF switch has the best throughput in a high degree of non-uniform traffic distribution. Since our proposed FBSF switch architecture with internal speedup can support the various input traffic patterns, it is appropriate for the future B-ISDN applications that require high bandwidths.

References

[1] F.A. Tobagi, "Fast Packet Switch Architecture for Broadband Integrated Services Digital network". *Proceedings of the IEEE*, vol.38, pp.133-167, Jan. 1990.

[2] C. Dhas, V.K. Konagi, and M. Sreetharan, "Broadband Switching : Architectures, Protocols, Design, and Analysis." *IEEE Computer Society press*, 1991.

[3] M.J. Karol, M.G. Hluchyj, "Input Versus Output Queueing on a Space-Division Packet Switch." *IEEE Trans., on Commun.*, vol. COM-35, No.12, pp.1347-1356, Dec., 1987.

[4] D.X. Chen and J.W. Mark, "SCOQ : A Fast Packet Switch with Shared Concentration and Output Queueing." *IEEE/ACM Trans., on Networking*, vol.1, No.1, pp.142-151, Feb. 1993.

[5] J.N. Giacopeli, J.J. Hickey, W.S. Marcus, W.D. Sincoskie, and M. Littlewood, "Sunshine : A high performance self-routing broadband packet switch architecture." *IEEE Journal on selected Areas in Communications*, vol.9, No.8, pp.1289-1298, Oct., 1991.

[6] ByoungSeob Park and SungChun Kim, "FB SF : A New Fast Packet Switching Fabric-based-on Multistage Interconnection Network with Multiple Outlets." *IEICE Trans. on Information and Systems*, vol. E80-D, No. 9, pp.847-853, Sep. 1997.

[7] T. Hanawa, H. Amano, and Y. Fujikawa, "Multistage Interconnection Networks with Multiple Outlets." *ICPP*, pp. 1-1-1-8, 1994.

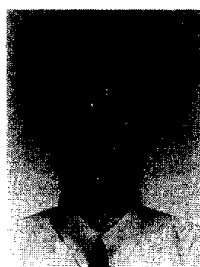
[8] M. Hamdi and J. Muppala, "Performance Evaluation of ATM Switches Under Various Traffic and Buffering Schemes." *IEEE GLOBECOM'95*, pp.828-832, 1995.

[9] S.C. Liew, "Performance of Various Input-Buffered and Output Buffered ATM Switch Design Principles under Bursty Traffic : Simulation Study." *IEEE Trans., on Com-*

*mun.*, vol.42, No.2/3/4, pp.1371-1379, Feb/Mar/Apr. 1994.

[10] A. Huang and S. Knauer, "Starlite : A Wideband Digital Switch." *IEEE GLOBECOM'84*, pp.121-125, 1984.

[11] M. Alimuddin, H.M. Alnuweiri, and R.W. Donaldson, "The Fat Banyan ATM switch." *IEEE INFOCOM'95*, pp.659-666, 1995.



박 병 섭

1989년 충북대학교 공과대학 컴퓨터공학과(학사)  
 1991년 서강대학교 대학원 전자계산학과(석사)  
 1991년 9월~1997년 서강대학교 대학원 전자계산학과(박사)

1997년~현재 국방정보체계연구소 선임연구원  
 관심분야 : ATM망 성능평가, ATM 스위치 설계, 무선 ATM



임 철 수

1985년 서울대학교 계산통계학과(학사)  
 1988년 Indiana University(석사)  
 1995년 서강대학교 대학원 전자계산학과(박사)

1985년~1986년 (주)DACOM 연구원  
 1989년~1993년 (주)아시아나항공 선임연구원  
 1994년~1996년 (주)신세기통신 선임과장  
 1997년~현재 서경대학교 공과대학 컴퓨터공학과 교수  
 관심분야 : 멀티미디어 시스템, 고속 통신망(ATM), PCS망, 무선망 설계