The Performance Comparison for the Contention Resolution Policies of the Input-buffered Crosspoint Packet Switch

Jung Hoon Paik and Chae Tak Lim

Abstract

In this paper, an NxN input-buffered crosspoint packet switch which selects a Head Of the Line,HOL, packet in contention randomly is analyzed with a new approach. The approach is based on both a Markov chain representation of the input buffer and the probability that a HOL packet is successfully served. The probability as a function of N is derived, and it makes it possible to express the average packet delay and the average number of packets in the buffer as a function of N. The contention resolution policy based on the occupancy of the input buffer is also presented and analyzed with this same approach and the relationship between two selection policies is analyzed in terms of the occupancy of the input buffer.

I. Introduction

Surveys of several proposed packet switching fabrics can be found in [1], [2]. Based on the placement of the buffers, these switch fabrics can be categorized into different architectures-internal buffer, input buffer, output buffer, shared buffer, or various combination of these[3]. Among them, the input-buffered architecture includes Batcher-banyan networks with ring reservation[4] or three-phase contention resolution[5], and a self-routing crossbar network with parallel, centralized contention resolution[6]. Because of head-of-line(*HOL*) blocking, its maximum throughput is only about 58 percent[5],[7], but this can be increased.

In this paper, a selection policy which selects the *HOL* packet in the input-buffered switch architecture is suggested and compared with the conventional selection policy. The conventional policy is called random selection policy which selects a *HOL* packet under the contention randomly. As conventional arbitration policy takes no account of the characteristic of the input traffic, it doesnt respond to the input traffic flexibly. Besides, the hot-spot phenomena for which the traffic is concentrated on the specific input port or output port may be popular in the broadband multimedia environments. As a result, the selection policy adapting to the fluctuation of the input traffic is desperately required to avoid the overflow in the input buffer. For this objective, a selection policy called the threshold selection policy which selects a *HOL* packet with reference to the occupancy of the input buffers is suggested. For the comparison of the performance for two selection polices, the analysis using *Markov* chain representation of the input buffer[8] is used and the probability that a *HOL* packet under contention situation is served is derived as a function of the number of input port.

Manuscript received December 20, 1996; accepted November 24, 1997.
J.H.Paik is Dept. of Radio Communication, Doowon Technical College
C.T.Lim is Dept. of Electronic Engineering, Hanyang Univ.

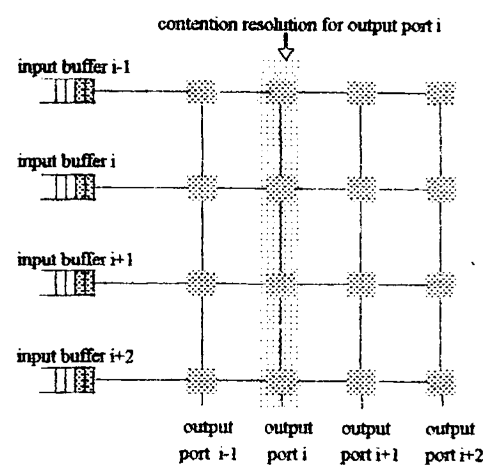


Fig. 1. Input buffered crosspoint switch.

II. Switch Architectures and Arbitration polices

1. Random selection

The input buffered crosspoint switch to be considered is shown in Fig. 1. Each arriving packet goes, at least momentarily, into a buffer on its input port.

At the beginning of every time slot, the arbitration functions check the first packet in each buffer. If every packet is addressed to different outputs, all the packets go through. If k packets are addressed to a particular output, one packet is picked to be sent; the others wait until the next time slot, when a new selection is made among the packets that are waiting. The random selection policy selects one of the k packets at random. Each packet is selected with equal probability $1/k$.

2. Threshold selection

A selection policy based on the occupancy of the buffer is presented in Fig. 2 and it is called threshold selection policy. When the threshold of the input buffer is passed, the Buffer Full, BF, signal is generated and transmitted to the cross-points on the same row. In case of contention among k input buffers, if all the k input buffers are below the threshold, one of the k packets is selected at random. If the occupancy of only one input buffer is above threshold, it wins the contention because the buffers with occupancy above the threshold get higher priority in arbitration than those below. If the occupancy of m input buffers of k input buffers are above the threshold, one is selected among the m buffers at random. This selection policy features the high adaptability to the burst traffic which concentrates on the input buffer transiently.

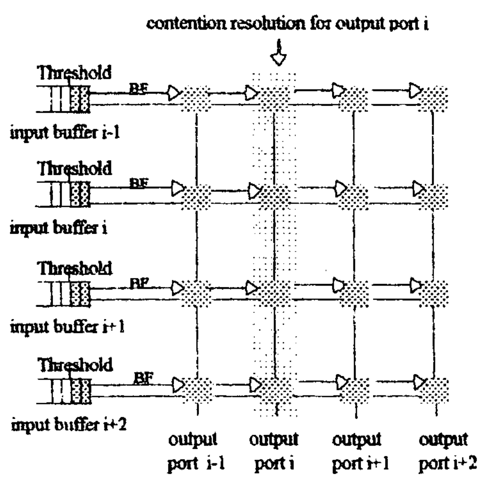
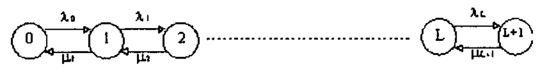


Fig. 2. Input buffered crosspoint switch with threshold.

III. Analytical Model

We assumed that packet arrivals at the N input ports are governed by independent and identical Bernoulli processes. In any given time slot, the probability that a packet arrives at a particular input is p. Each packet has equal probability $1/N$ of being addressed to any given output port. In addition, the destinations of newly arriving packets are independent of those of previous packets. This assumption simplifies the analysis without significantly affecting the results[8]. Following an approach based on [9], this buffer can be represented as a *Markov* chain with state variable k, the number of packets present at the beginning of a time slot. Fixing our attention on a *HOL* packet in a particular buffer(the tagged packet) and defining the probability that the tagged packet is served is q, the transition rate diagram of the tagged buffer is illustrated as Fig. 3 where L is the buffer length.



$$\lambda^k = \begin{cases} p & for \quad k = 0 \\ p(1-q) & for \quad 0 < k \le L \\ 0 & for \quad k > L \end{cases} \tag{1}$$

$$\mu^k = q\,(1-p) \quad for \quad k > 0$$

Fig. 3. The discrete-time Markov chain transition rate diagram for the input buffer size.

The probability pk which is the probability that there are k packets in an input buffer is given by

$$p_k = p_0 \prod_{i=0}^{k-1} \frac{\lambda_i}{\mu_{i+1}} = p_0 \frac{p}{q(1-p)} \left[\frac{p(1-q)}{q(1-p)} \right]^{k-1}$$

$$for \quad 1 \le k \le L+1 \tag{2}$$

and p_0 is obtained by means of the normalization condition

$$\sum_{k=0}^{L+1} p_k = 1 \tag{3}$$

From the definition of the p and q, the equilibrium probability pk only exists at the condition of $p<q$.
Solving (3), p_0 is obtained as a function of p, q, and L:

$$p_0 = \frac{q-p}{q - p\left[\frac{p(1-q)}{q(1-p)} \right]^{L+1}} \tag{4}$$

and by extending the sum in (3) or approaching L in (4) to infinity , the relationship among p_0 , p, and q is given by

$$P_0 = 1 - \frac{p}{q} \tag{5}$$

The average packet delay is considered as the average number of time-slots between the reception of a packet at the switch and its successful transmission. The number of packets preceding a newly arrived packet at the beginning of the first time-slot after its arrival, is given by the sum of the number of packets already present in the system, diminished by the average number of packets that leave the buffer during a slot, unless the buffer is empty.[9] By considering that each packet can be averagely transmitted in *1/q* slots and by taking the average of the position occupied in the system by a newly arrived packets, the following expression of the average packet delay, *D*, is obtained

$$D=[1+\overline{K}-q(1-p_o)]\frac{1}{q} \tag{6}$$

where $\overline{K}$ is t he average number of packets in the buffer and its expression as a function of *p* and *q* is

$$\overline{K}=\frac{p(1-p)}{q-p} \tag{7}$$

For the expression of both the average packet delay and the average number of packets in terms of *p*, the probability *q* which varies depending on the selection policy should be expressed as a function of *p*, and next chapter deals with it.

Ⅳ. Successful Transmission Probability

1. Random selection

In the random selection policy, *q* is given by

$$q=\sum_{i=0}^{N-1}\binom{N-1}{i}(1-P_0)^i P_0^{N-1-i}\sum_{j=0}^{i}\binom{i}{j}\left(\frac{1}{N}\right)^j\left(1-\frac{1}{N}\right)^{i-j}\frac{1}{j+1} \tag{8}$$

That is, under the condition that there are *i* non-empty buffers of *N-1* buffers and *j* *HOL* packets of the *i* non-empty buffers with the same destination as the tagged packet, the probability that the tagged packet is served is given by *1/j+1*. The probability that there are *i* non-empty buffers of *N-1* buffers is given by $(1-p_0)^i\, p_0^{N-1-i}$ and the probability that *j* *HOL* packets of the *i* non-empty buffers with the same destination as the tagged packet is given by $(1/N)^j(1-1/N)^{i-j}$. The probability, *q*, is also derived from the probabilities that the tagged packet wins the contention or experiences no contention. The probability, $p_{contention}$, that the tagged packet wins the contention is expressed as

$$p_{contention}=\sum_{i=1}^{N-1}\binom{N-1}{i}(1-P_0)^i P_0^{N-1-i}\sum_{j=1}^{i}\binom{i}{j}\left(\frac{1}{N}\right)^j\left(1-\frac{1}{N}\right)^{i-j}\frac{1}{j+1} \tag{9}$$

and the probability ,$p_{no\ contention}$, that the tagged packet experiences no contention is given by

$$P_{nocontention}=1-\sum_{i=1}^{N-1}\binom{N-1}{i}(1-P_0)^i P_0^{N-1-i}\sum_{j=1}^{i}\binom{i}{j}\left(\frac{1}{N}\right)^j\left(1-\frac{1}{N}\right)^{i-j} \tag{10}$$

So, the probability, *q* is

$$q=\sum_{i=1}^{N-1}\binom{N-1}{i}(1-P_0)^i P_0^{N-1-i}\sum_{j=1}^{i}\binom{i}{j}\left(\frac{1}{N}\right)^j\left(1-\frac{1}{N}\right)^{i-j}\frac{1}{j+1}$$

$$+1-\sum_{i=1}^{N-1}\binom{N-1}{i}(1-p_0)^i P_0^{N-1-i}\sum_{j=1}^{i}\binom{i}{j}\left(\frac{1}{N}\right)^j\left(1-\frac{1}{N}\right)^{i-j} \tag{11}$$

Both (8) and (11) are simplified to the same expression as (12).

$$q=\frac{1}{1-p_0}\left[1-\left(1-\frac{1}{N}+\frac{P_0}{N}\right)^N\right] \tag{12}$$

The equations (4) and (12) are too complex to be expressed as a function of *p*, *N*, and *L* but for the infinite buffer length, *q* can be expressed in terms of *p* and *N*

$$q=\frac{p}{N\left[1-(1-p)^{\frac{1}{N}}\right]} \tag{13}$$

As $N \to \infty$,

$$q=\frac{-p}{\ln(1-p)} \tag{14}$$

Figure 4 shows the dependency of *q* on *p* and *N*.

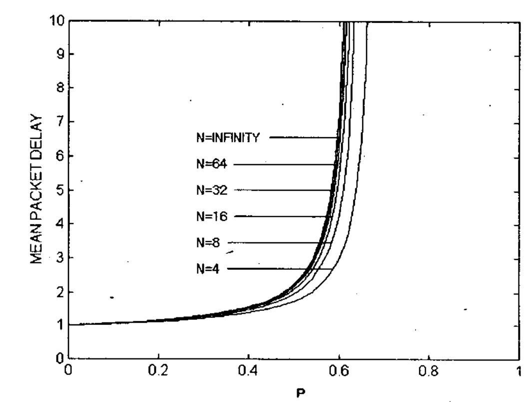


Fig. 4. The probability that a HOL packet is served.

From (5) and (13), the probability ,*p0* , is also expressed in terms of *p* and *N*

$$P_0=1-N\left[1-(1-p)^{\frac{1}{N}}\right] \tag{15}$$

As N -> ∞,

$$p_0=1+\ln(1-p) \tag{16}$$

The equilibrium probability p_k in terms of *N* and *p* is given by

$$p_k=\frac{N\left[1-(1-p)^{\frac{1}{N}}\right]\left[1-N\left[1-(1-p)^{\frac{1}{N}}\right]\right]}{1-p}\left\{\frac{n\left[1-(1-p)^{\frac{1}{N}}\right]-p}{1-p}\right\}^{k-1} \tag{17}$$

As $N \to \infty$,

$$P_k=\frac{-\ln(1-p)[1+\ln(1-p)]}{1-p}\left[\frac{-\ln(1-p)-p}{1-p}\right]^{k-1} \tag{18}$$

The probability p_k as a function of *N* for specific *k* is depicted in Figure 5.

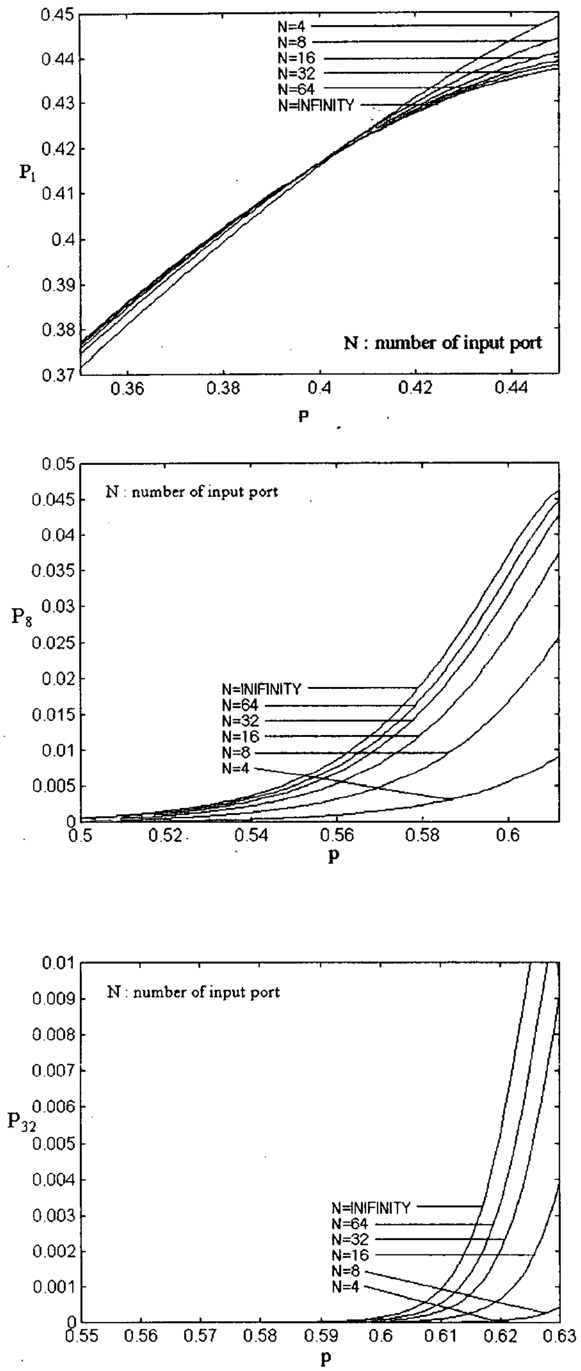


Fig. 5. The equilibrium probability p_k.

The curves of equilibrium probability ,p_k, as a function of N shows that the probability increases as N increases and it is reasonable result in that the chance that a *HOL* packet is served decreases as N increases. It follows from (6), (7) and (13) that the average packet delay, D, is given by

$$D = \frac{(1-p)N\left[1-(1-p)^{\frac{1}{N}}\right]}{\left\{1-N\left[1-(1-p)^{\frac{1}{N}}\right]\right\}p} \quad \textbf{(19)}$$

and the average number of packets in the buffer, $\overline{K}$,is expressed as

$$\overline{K} = \frac{(1-p)N\left[1-(1-p)^{\frac{1}{N}}\right]}{1-N\left[1-(1-p)^{\frac{1}{N}}\right]} \quad \textbf{(20)}$$

The relationship between equations (19) and (20) is $D=\overline{K}/p$ and this result also can be attained from Littles theorem directly. Figures 6 presents the average packet number with infinite buffer length respectively.

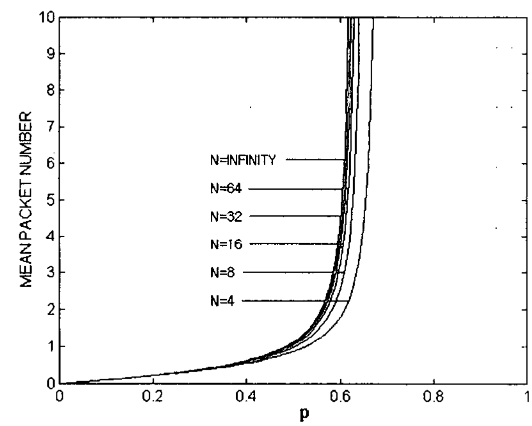


Fig. 6. The mean packet number with the random selection.

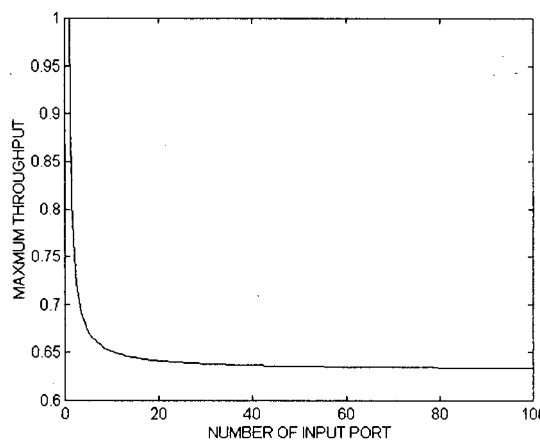


Fig. 7. The maximum throughput.

The maximum p, maximum throughput, is obtained when q approaches to p. Applying this condition to (13), the maximum throughput is gven by(21).

$$1-\left(1-\frac{1}{N}\right)^{N} \tag{21}$$

The switch throughput as a function of N is shown in Figure 7.

2. Threshold selection

In threshold selection policy, the probability that the tagged packet is served, q, is conditioned on its buffer state. So, qk is used for its dependency on the buffer state in this selection policy. In case that its buffer state is below the threshold, the tagged packet gets the opportunity to be served only when the *HOL* packets in the buffers whose occupancy is above the threshold have the different destinations and the *HOL* packets in the buffers whose occupancy is below threshold have the same destination as the tagged packet. For this case, the probability is given by (22) where $p_{0<m<TH}$ means the pk for $0<k<TH$ and TH is the threshold value in the input buffer.

$$q_r=\sum_{i=0}^{N-1}\binom{N-1}{i}p_{0<m<TH}^{i}\sum_{j=0}^{N-1-i}\binom{N-1-i}{j}p_{m\geq TH}^{j}p_0^{N-1-i-j}\left(1-\frac{1}{N}\right)^{j}\sum_{k=0}^{i}\binom{i}{k}\left(\frac{1}{N}\right)^{k}\left(1-\frac{1}{N}\right)^{i-k}\frac{1}{k+1}$$

$$\text{for} \quad r<T \tag{22}$$

On the other hand, if the tagged buffer state is above the threshold, it contends with only the buffers whose occupancy is above the threshold and whose *HOL* packets have the same destinations as the tagged packet. For this case, the probability is given by (23).

$$q_r=\sum_{i=0}^{N-1}\binom{N-1}{i}P_{m\geq TH}^{i}(1-P_{m\geq TH})^{N-1-i}\sum_{j=0}^{i}\binom{i}{j}\left(\frac{1}{N}\right)^{j}\left(1-\frac{1}{N}\right)^{i-j}\frac{1}{j+1}$$

$$\text{for} \quad r\geq TH \tag{23}$$

For the convenience of the notation, q_r for $r < TH$ and $r \geq TH$ is referred to q_a and q_b respectively.

From the (2), $p_{0<m<TH}$ and $p_{m\geq TH}$ in (22) and (23) are given by

$$p_{0<m<TH}=\sum_{i=1}^{TH-1}p_i=\frac{p_0 p}{q_a-p}\left[1-\left\{\frac{p(1-q_a)}{q_a(1-p)}\right\}^{TH-1}\right] \tag{24}$$

$$p_{m\geq TH}=\sum_{i=TH}^{\infty}p_i=\frac{p_0 p}{q_b-p}\left[\frac{p(1-q_b)}{q_b(1-p)}\right]^{TH-1} \tag{25}$$

Equation (22) and (23) are reduced to

$$q_r=\frac{1}{P_{0<m<TH}}\left[\left(1-\frac{p_{m\geq TH}}{N}\right)^{N}-\left(1-\frac{1}{N}+\frac{p_0}{N}\right)^{N}\right]$$

$$\text{for} \quad r<T \tag{26}$$

$$q_r=\frac{1}{p_{m\geq TH}}\left[1-\left(1-\frac{p_{m\geq TH}}{N}\right)^{N}\right] \quad \text{for} \quad r\geq TH \tag{27}$$

If the threshold of one or infinity is applied to (26) and (27), $p_{0<m<TH}$ and $p_{m\geq TH}$ become 1 or 0 respectively, and q for both cases tends to (12) which is the q for the random selection policy. Therefore, the performance of the threshold policy with threshold of one or infinity will be same as that of the random policy. This result is anticipated in that the threshold selection policy with threshold of one or infinity shows identical operations with the random selection policy. From the definition of q, the maximum throughput over the entire N input ports, S , is considered as the q when p_0 is 0.

$$S=\left(q_a\big|_{p_0=0}P_{0<m<TH}+q_b\big|_{p_0=0}P_{m\geq TH}\right)N \tag{28}$$

Applying (24) and (25) to (28), S is given by

$$\left(1-\left(1-\frac{1}{N}\right)^{N}\right)N \tag{29}$$

For the random selection policy, the maximum throughput over the entire N input ports is attained by multiplying (21) by N. As the multiplying result is equal to (29), it is concluded that the maximum throughput over the N input ports for the threshold selection policy is same as for the threshold selection policy. Although the overall maximum throughput is same for the two policies, the throughput for the specific input port is different according to the condition on the occupancy of the input buffers. The curves q_a, q_b *and* q as a function of $p_{m\geq TH}$ are presented in figure 8. First, as N increases, all the curves decreases. This is resulted from more contentions caused by the increase of N. The figure also shows that q_a and q_b decrease as $p_{m\geq TH}$ increases. As in (22) and (23), $p_{m\geq TH}$ eans that the probability that the occupancy of the other buffers except the tagged buffer is above the threshold. So, the increase of $p_{m\geq TH}$ results in that of contention between the tagged buffer and others. The more contention happens, the more q decreases. When $p_{m\geq TH}$ is 1, the figure shows that qb is identical to q. As the value 1 of $p_{m\geq TH}$ means that the occupancy of all the input buffers is above the threshold, there is no difference between two selection policies under this conditions. So, q_b should be same as q. The figure also shows that q_a and q have the same values when $p_{m\geq TH}$ is 0. As value 0 of $pmTH$ means 1 of $p_{0<m<TH}$, this case is that all the input buffers are in the state of $p_{0<m<TH}$. In this case, there is no difference between two policies and it is proved in the figure. The simulation for the correctness of the analysis is performed and presented in the figure. The simulation is only performed for the case that $p_{m\geq TH}$ is 0, 0.25, 0.50, 0.75, and 1. It is learned from the figure that the analysis and the simulations match.

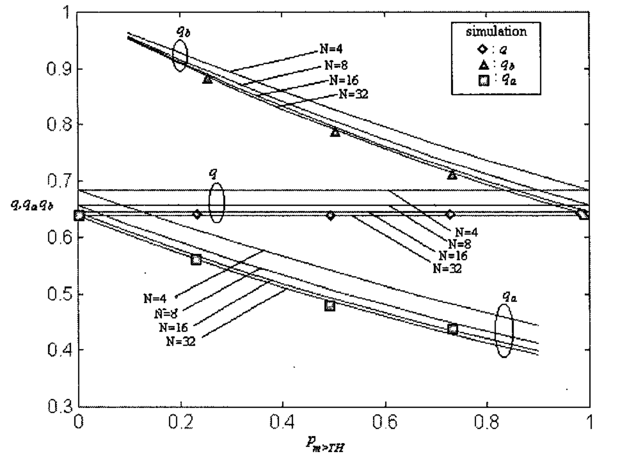


Fig. 8. Comparison of q for threshold policy and random policy.

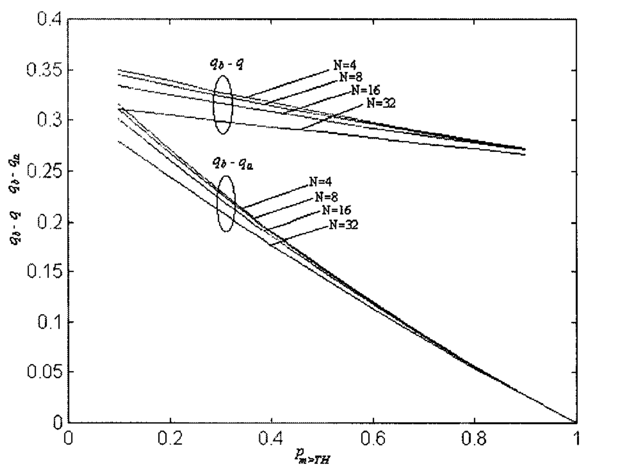


Fig. 9. Difference between q, qa and qb.

The differences among q, q_a and q_b are presented in figure 9. It is learned from the figure that the difference between q_b and q is about 0.3 and the difference decreases as $p_{m \geq TH}$ increases. The increase of $p_{m \geq TH}$ causes the decrease of the qb but the q is constant, so the results is obtained. It is observed from the analysis and simulations that the service rate of the threshold selection policy on the *HOL* packet in the buffer whose occupancy is above the threshold is about 30 % higher than that of the random selection policies. Therefore, the threshold selection policies are suitable to the environments in which the input traffics are so bursty and concentrated to some input ports, so the probability of overflow for the concentrated buffers is very high. As the threshold selection policies service the *HOL* packets in the concentrated buffers preferentially, the overflow caused by the transient bursty traffics is more reduced than the random selection policies.

V. Simulation and Discussion

The simulation is performed for the dependency of the threshold selection policy on the threshold and the comparison for two policies. The simulation is performed under the condition that the number of input and output ports is 8 and the size of input buffer is 32 and the iteration numbers are 200,000. In the case that the input traffics and those destinations are evenly distributed, the blocking probability for the two selection policies is simulated in figure 10.

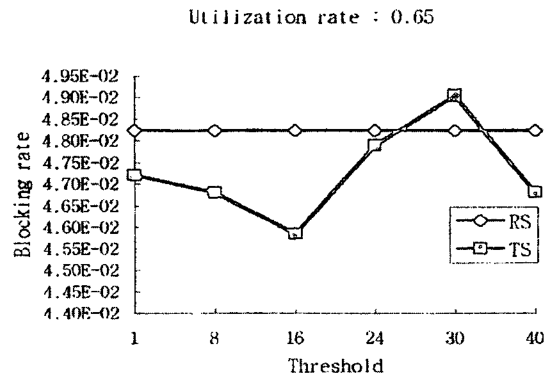


Fig. 10. Blocking rate for evenly distributed traffics.

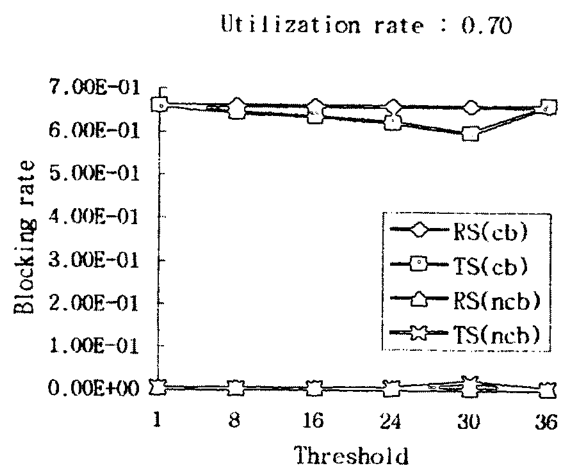


Fig. 11. Blocking rate for concentrated traffics.

It is observed from the figure that two polices behave identically for this kind of traffics. Under the above traffic patterns, all the input buffers have the same occupancy. So, two selection policies operate identically. On the other hand, when the input traffics are concentrated to a specific input port and the traffics of input traffics are evenly distributed to output ports, the dependency of the blocking probability on the threshold in threshold selection policy is given in figure 11. As the analysis is indicated, the figure also shows that two selection policies have the same blocking probability when the threshold is 1 or above the buffer length. It is also learned from the figure that the blocking probability of the concentrated input buffer and non-concentrated input buffer decreases and increases respectively as the threshold increases. When the input rates to the concentrated buffer are high, so the mean buffer lengths of the input buffer approache to the buffer size, the increase of the threshold causes that of $p_{m\geq TH}$ of the concentrated input buffer. In turn, this results in more service of the *HOL* packet in the concentrated buffer. As the destinations of the input traffics are evenly distributed, the increase of the service probability for the concentrated buffer means the decrease of the non-concentrated input buffers which have the *HOL* packets whose destinations are same as those of *HOL* packts in the concentrated buffers.

So, the service probability of the non- concentrated buffer decreases. It is interesting that the blocking rates in the threshold selection policy are dependent on the threshold as figure 11. The blocking rate of the concentrated buffer in the threshold selection policy is minimum when the threshold is 30. But, at this threshold, the blocking rate of the non-concentrated buffers is maximum. This situation is caused by the policy of the threshold selection which gives higher service priority to the concentrated buffers than the non-concentrated buffers. So, it is concluded that the threshold makes an effect on the performance of the blocking rate.

VI. Conclusion

In this paper, two arbitration polices of an *N*x*N* input-buffered crosspoint packet switch are compared with a new analysis. The analysis shows that the maximum throughput over the *N* input ports is same between two selection policies. But, the throughput for a specific input port is different between two policies. The analysis also shows that the threshold selection policy serves the *HOL* packets in the input buffer whose occupancy is above the threshold with higher probability than the conventional selection policy and this situation of higher probability is evident as the input traffic increases. So, it is concluded that the threshold selection policy is suitable to the broadband service environments in which input traffics to the switch are so fluctuated that an input buffer is overloaded transiently because it shows better performance for the blocking rate than the random selection policy.

References

[1] H. Ahmadi and W. E. Denzel, A survey of modern high-performance switching techniques, IEEE J. Select. Areas Commun., Vol.7, No. 7, pp.1091-1103, Sept. 1989.

[2] F. A. Tobagi, Fast packet switch architectures for broadband integrated services digital network, Proc. IEEE, Vol. 78, No. 1, pp.133-167, Jan. 1990.

[3] M. Hluchyj and M. Karol, Queueing in high-performance packet switching, IEEE J. Select. Areas Commun., Vol. 6, No. 9, pp. 1587-1597, Dec. 1988.

[4] B. Bingham and H. Bussey, Reservationbased contention resolution mechanism for Batch-banyan packet switches, Electron. Lett., Vol. 24, No. 13, pp.772-773, June, 1988.

[5] Joseph Y. Hui and Edward Arthurs, "A Broadband Packet Switch for Integrated Transport", IEEE JSAC. , Vol. SAC-5, no. 8 pp.1264-1273, Oct. 1987.

[6] A. Cisneros, Large packet switch and contention resolution device, Proc. ISS90, Stockholm, Sweden, May 1990, pp.77-83.

[7] M. J. Karol, M.G. Hluchyj and S. P. Morgan, Input versus output queueing on a space-division packet switch, IEEE Trans Commun., Vol. 35, No. 12, pp. 1347-1356, Dec. 1987.

[8] J.H. Patel, Performance of processor- memory inter-connections for multiprocessors, IEEE Trans. Comput., Vol., C-30, Oct. 1981.

[9] G. Corazza and C. Raffaelli, Performance Evaluation of Input-Buffered Replicated Banyan Networks, IEEE Trans. Commu., Vol., 41, No. 6, June 1993.

[10] Joseph Y. Hui and Edward Arthurs, "A Broadband Packet Switch for Integrated Transport", IEEE JSAC., Vol. SAC-5, no. 8 pp.1264-1273, Oct. 1987.

Jung Hoon Paik received the B.S. and the M.S. degrees in Electronic Eng. from Hanyang University, Korea, in 1986 and 1988 respectively. From 1988 to 1996, he was a senior engineer for ETRI. Since 1996, he has been at Doowon technical college. He is currently working towards the Ph.D. degree in electronic engineering from Hanyang University. His research interests are ATM switching systems and transmission systems. He is a member of KITE.

Chae Tak Lim received the B.S. in Electrical Eng. from Junbuk National University, the M.S. degree from University of Hawaii, and his Ph.D. degree in Electronic Eng. from Hanyang University, Korea, in 1957, 1969 and 1974, respectively. From 1962 to 1976, he was a associate professor in the department of Electrical Engineering at Junbuk National University. Since 1967, he has been at Hanyang University. He was a visiting professor at Technology University of Denmark from 1975 to 1976. He was the president of Korea institute of Telematics and Electronics in 1991. Currently he is a professor at the department of Electronic Engineering, Hanyang University. His research interests include switching theory, circuit & systems, image processing and parallel processing. He is a member of the KITE, KISS, and KICS.