재순환 구조를 가진 dilated 반얀 네트웍

정회원 윤 영 식*

Dilated Banyan Network with Recirculation

Young-sik Youn* *Regular Member*

요 약

반얀 (Banyan) 네트웍은 ATM 스위치의 기본 구조로 많이 사용되어 왔다. 그러나, 반얀 네트웍은 내부 블록킹이 큰 단점으로 알려져 있으며, 이 단점을 보완하기 위해 제시되고 있는 구조 중 하나가 dilated 반얀 네트웍이다. 그런데, 기존의 dilated 반얀 네트웍 스위치는 네트웍 용량의 많은 부분이 낭비되는 구조를 가지고 있다. 본 논문에서는 이렇게 낭비되는 자원을 충분히 활용하여 스위치 용량을 증가시킬 수 있는 구조를 제안한다. 제안된 스위치는 deflection 라우팅과 recirculation을 가지는 구조로, 스위치의 성능은 uniform 트래픽 가정 하에 분석되었다. 분석 및 시뮬레이션을 통한 검증 결과, 제안된 스위치의 성능이 기존의 dilated 반얀 네트웍에 비해 월등히 향상됨을 보여 준다.

ABSTRACT

Banyan network has been widely employed as a basic building block for ATM switches. But the banyan network has very low routing capacity because of the internal blocking problem. Hence, a dilated banyan network has been used as one solution that can overcome the internal blocking problem. However, tremendous network capacity is wasted in the dilated network In this paper, we propose a dilated banyan network with deflection routing and recirculation mechanism to fully utilize the wasted capacity. The performance of the proposed switch is analysed under uniform traffic assumption. Numerical and simulation results show that the proposed switch yields a significant improvement of the maximum throughput as compared that of the pure dilated banyan network.

I. Introduction

Asynchronous transfer mode (ATM) has been widely accepted as the most promising solution to implement the broadband ISDN. Recently, various switching architectures have been proposed for ATM applications[1]. Many of the proposed architectures employ the banyan network as a basic building block, because of its hardware simplicity and self-routing capability. But, one problem with the banyan network is internal blocking which occurs when two cells use common internal links, although they are not destined to the same network output.

A dilated banyan network is one of solutions [5-10] proposed to overcome blocking and to improve throughput. In the d-dilated network [11,12], input and output ports of a switching element (SE), respectively, have d parallel links

instead of a single link, so that a maximum of d cells can be forwarded between two SEs simultaneously in each time slot. Thus, the probability of blocking rapidly decreases as the dilation degree d increases. It has been shown [12] that, to achieve the loss probabilities between 10^{-6} and 10^{-9} for a 1024×1024 network, the dilation degree of 9 and 12 are required, respectively. However, it seems that such dilation degrees are too high for practical implementation. In fact, tremendous network capacity is wasted in the dilated banyan network, because input ports of SEs in the first stage of the network receive cells only on one of the d links, and the remaining d-1 links are unused.

In order to utilize the network capacity wasted in the pure dilated banyan network, the back-pressure mechanism may be introduced[13]. But, for the back-pressure mechanism, each port of an SE requires d multiple links for backward acknowledgement signals, and thus hardware complexity increases in viewpoint of chip implementation. In this paper, as another approach to utilize the wasted routing capacity, we propose a switch architecture that is a dilated banyan network using recirculation instead of the back-pressure mechanism. For the recirculation, the proposed switch employs the deflection routing [6, 9], in which the blocked packet is not dropped at the blocked SE but routed to the "wrong" output.

Thus, the blocked packet can be transferred to the switch input through the recirculation link, after arriving at the output of the dilated banyan network. That is, the switch requires no backward signal link. However, since the recirculated packets are not transmitted at the input ports at which they arrived originally, the proposed switch has the out-of-sequence problem, and thus requires a resequencer in each output queue. We use a packet resequencing method based on *delay equalization* [9], in which the resequencer delays the packet in the output queue until its total delay becomes equal to a predetermined maximum delay D_R. Here, the total delay is the number of

time slots taken for the packet to leave the resequencer from the switch input. If the packet arrives at the resequencer with delay longer than the maximum delay D_R, they are discarded. Hence, in order to minimize the packet loss, the parameter D_R has to be selected carefully.

The selection of the parameter D_R is decided based on the distribution of the delay of the packet arriving at the resequencer. But, the distribution of the delay is largely affected by the methods that recirculated packets are handled when conflicts occur in the network. Hence, we investigate two contention resolution methods, i.e., random service and priority service. In the resolution method by random service, no distinction is made between packets in the SE. On the other hand, in the resolution method by priority service, each SE will preferentially select the packets with longer delay. We also analyze the performance of the proposed switch for the two contention resolution methods, and compare their performances.

II. Switch Architecture and Operation

The proposed switch architecture is composed of input modules, a d-dilated banyan network, packet filters and output modules. The outputs of the d packet filters in the i-th output port are connected to the i-th input module for recirculating unsuccessful packets. As an example, Fig. 1 shows the switch using a 2-dilated banyan network of size 8×8. The switch operates in a time-slotted fashion with the slot size equal to the transmission time of a cell. The cell has a fixed size, and its header consists of a deflection bit, a time stamp field and an address field. The time stamp is used for resequencing the cell at the output module.

Each input module consists of three components: an input queue, a selector, and d input controllers (ICs). At the beginning of every time slot, the selector determines one of IC's in

the input module which is not occupied by recirculated cells and makes a path between the input queue and the IC. Then, a cell in the input queue is forwarded to the IC through the path. In case that all of the IC's are occupied by recirculated cells, the cell is delayed in the input queue. The cell arriving at the input queue is forwarded directly to the IC if the input queue is empty and one of IC's is available. The IC has a cell buffer in order to store a new cell from the input queue or a recirculated cell, and transmits the cell to the network. It is noted that IC begins transmission of the cell before the cell is completely received to synchronize the arrival of recirculated cells with the arrival of new cells. The IC also removes the deflection tag from the recirculated cell before its transmission. In addition, the IC updates the time stamp value for the resequencing function. That is, the IC sets the time stamp to 0 for a new cell, and increases the time stamp value by 1 for the recirculated cell. When the time stamp value of the recirculated cell is U, the cell is removed from the IC.
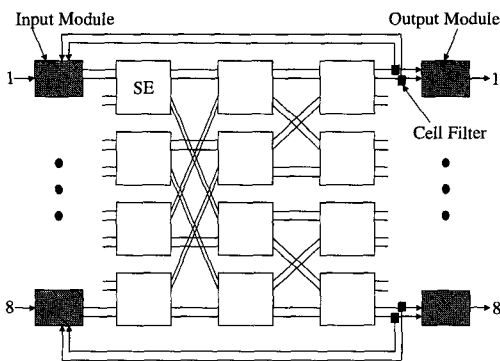

Fig. 1 Dilated banyan network with recirculation.

In order to recirculate the blocked cells due to internal conflicts, the dilated banyan network employs a deflection routing algorithm in the operation of the SE. When $i(\leq d)$ cells are destined for the same output port of an SE, all of them are routed to the i links of the output port. If $i>d$, then d cells are routed to the output

port and $i-d$ cells are deflected to the other output port with deflection tags. In the next remaining stages, the deflected cells are randomly routed so as not to affect the path of normally routed cells. At the output of the dilated banyan network, the cell filters examine the arriving cells. The normal cells are transmitted to the output modules, while the deflected cells are recirculated to the input module.

Each output module includes a resequencer and an output queue. In order to perform the resequencing function based on the delay equalization, the resequencer separates the output queue into two sub-queues, i.e., resequencing queue and transmission queue. Thus, cells arriving at the output module with the time stamp values lower than D_R are stored in the resequencing queue, but cells with the time stamp value of D_R are stored directly in the transmission queue. In case that the arriving cells have the time stamp values higher than D_R, they will be discarded. Before each time slot starts, the resequencer increases the time stamp values of the cells in the resequencing queue by 1, and then moves the cells into the transmission queue if their time stamp values become D_R. Accordingly, cells in the transmission queue can be transmitted without out-of-sequence because they have completed the resequencing process.

When there occurs internal conflicts between cells, each SE has to resolve the contention. We investigate two contention resolution methods, i.e., *random service* and *priority service*. In the resolution method by random service, no distinction is made between new and recirculated cells in the SE, so that d cells are randomly selected. On the other hand, the resolution method by priority service uses $U+1$ priority levels for distinction of the cells, so that each SE preferentially selects the cells with higher priority levels when blocking occurs. Here, the time stamp values are used for indicating the $U+1$ priority levels. Moreover, the maximum priority level U is set to the same value as D_R.

III. Contention Resolution by Random Service

3.1 Performance Analysis

Consider a d-dilated network of size $N \times N$, which is constructed with n (= $\log_2 N$) stages. We assume that the input traffic pattern is uniform, i.e., cells arriving at each input queue are equally likely to be destined for all output queues, and the cell arrivals are governed by independent and identical Bernoulli processes. Although we assume uniform traffic pattern, the traffic flow inside the network is more or less non-uniform due to recirculated cells. However, for analytical tractability, it is assumed that the traffic flow inside the network is also uniform. In addition, we assume that the maximum time stamp value U is infinite so that cells are not lost during the recirculation.

Let $q_{k,t}(l,m)$ denote the probability of finding l normal cells and m deflected cells on one output of an SE in stage k at time slot t. In particular, let $q_{0,t}(l,0)$ denote the probability of finding l normal cells in IC's of each input module at time slot t. Then, the probability that an SE in stage k receives i normal cells and j deflected cells at time slot t, $r_{k,t}(i,j)$, is given by

$$r_{k,t}(i,j) = \sum_{l=[i-d]^+}^{\min(i,d)} \sum_{m=[j-d]^+}^{\min(j,d)} q_{k-1,t}(l,m)$$
$$\times q_{k-1,t}(i-1,j-m) \quad (1)$$

where $[x]^+ \triangleq \max(0,x)$.

Now, let us consider that l normal cells and m deflected cells are transmitted to a tagged output of an SE, when there are i normal cells and j deflected cells in the SE. In this case, assuming that b deflected cells are originally destined for the tagged output, the set of the possible b, $B_{l,m|i,j}$ is dependent on following three factors: i) the number of remaining links which are not occupied by normal cells, d-l, ii) the number of normal cells which are originally destined for the

other output port, but deflected to the tagged output, $[i\text{-}l\text{-}d]^+$, iii) the number of deflected cells which are originally destined for the other output, but routed to the tagged output, $[j\text{-}b\text{-}[d\text{-}i+l]^+]^+$. Hence, it is given by

$$B_{l,m|i,j} = \{ b \mid m = \min[d-l, b+[i-l-d]^+ + j-b-[d-i+l]^+]^+], \ 0 \le b \le j \}.$$

Consequently, the conditional probability $q_{k,t}(l,m|i,j)$ is given by

$$q_{k,t}(l,m|i,j) = \begin{cases} \sum_{a=d}^{i} \begin{bmatrix} i \\ a \end{bmatrix} 2^{-i}, \\ \qquad \text{if } l=d, \ m=0 \\ \begin{bmatrix} i \\ l \end{bmatrix} 2^{-i} \cdot \sum_{b \in B_{l,m|i,j}} \begin{bmatrix} j \\ b \end{bmatrix} 2^{-j}, \\ \qquad \text{if } l<d, \ 0 \le m \le d-l \\ 0, \quad \text{otherwise.} \end{cases} \quad (2)$$

Using (1) and (2), we have, for $1 \le k \le n$,

$$q_{k,t}(l,m) = \sum_{i=0}^{2d} \sum_{j=0}^{2d-i} q_{k,t}(l,m|i,j) \cdot r_{k,t}(i,j). \quad (3)$$

At the next time slot t+1, the load offered at IC's of each input module is given by

$$q_{0,t+1}(l,0)$$
$$= \begin{cases} (1-\lambda_{in}) \cdot R_t(l), & \text{if } l=0 \\ \lambda_{in} \cdot R_t(l-1)+(1-\lambda_{in}) \cdot R_t(l), \\ & \text{if } 0 < l < d \\ \lambda_{in} \cdot R_t(l-1)+R_t(l), & \text{if } l=d \end{cases} \quad (4)$$

where λ_{in} is the arrival rate of new cells from the input queue and $R_t(l) \equiv \sum_{x=0}^{d-l} q_{n,t}(x,l)$.

In the steady-state, the normalized throughput of the network ρ is given by

$$\rho = \sum_{l=1}^{d} \sum_{m=0}^{d-l} l \cdot q_n(l,m) \quad (5)$$

where $q_n(l,m)$ is the steady-state value of $q_{n,t}(l,m)$. Since the success probability, $Psuc$, is equal to

the ratio of the number of cells in the input module to the number of normal cells on the output port of the last stage, we have

$$P_{SUC} = \frac{\rho}{\sum_{l=1}^{d} l \cdot q_0(l,0)} . \qquad (6)$$

Let $D(t)$ denote the probability that it takes t time slots for a cell to be transmitted from the IC to the desired output port. Then, it is given by

$$D(t) = [1 - P_{SUC}]^t \cdot P_{SUC}. \qquad (7)$$

Letting K_{in} denote the buffer size of an input queue, the input queue can be modeled as a $K_{in}+1$ state Markov chain, where the state i means that the input queue has i cells. Since the input queue can transmit only a single cell, the state of the input queue can be transferred only to adjacent states or the state itself. If the input queue is in state 0 and there is no cell arrival, the input queue will remain empty. Also, although a cell arrives at the input queue, the input queue is still in state 0 if the cell is directly forwarded without being stored. Transition from state 1 to state 0 occurs when the cell in the input queue is forwarded and there is no cell arrival. Let $\pi(i)$ denote the probability that there are i cells in the input queue, and let λ denote the probability that a cell arrives at the input queue. Then, we have for state 0,

$$\pi(0) = [\overline{\lambda} + \lambda \cdot \overline{P_B}] \cdot \pi(0) + \overline{\lambda} \cdot \overline{P_B} \cdot \pi(1) \qquad (8)$$

where P_B is the probability that all of the IC's in an input module are occupied by recirculated cells, i.e., $P_B = \lim_{t\to\infty} R_t(d)$, and $\overline{\lambda} = 1 - \lambda$, $\overline{P_B} = 1 - P_B$.

For the state i ($1 \leq i \leq K_{in}-1$), the input queue remains in state i itself, if there is one departure and one arrival, or no departure and no arrival. Transition from $i-1$ to state i occurs when there is no departure and one arrival, and transition from state $i+1$ to state i occurs when there is one departure and no arrival. Hence, these cases are described by the following equations:

$$\pi(1) = \lambda \cdot P_B \cdot \pi(0) + [\overline{\lambda} \cdot \overline{P_B} + \lambda \cdot P_B] \cdot \pi(1)$$
$$+ \overline{\lambda} \cdot \overline{P_B} \cdot \pi(2), \quad \text{if } K_{in} \geq 2 \qquad (9)$$

$$\pi(i) = \lambda \cdot P_B \cdot \pi(i-1) + [\overline{\lambda} \cdot \overline{P_B} + \lambda \cdot P_B]$$
$$\times \pi(i) + \overline{\lambda} \cdot \overline{P_B} \cdot \pi(i+1),$$
$$\text{for } 2 \leq i \leq K_{in}-1. \qquad (10)$$

Since the probability that the input queue is in any one of the possible state K_{in} is 1, we have

$$\sum_{i=0}^{K_{in}} \pi(0) = 1. \qquad (11)$$

By solving (8)~(11), we can easily obtain the state probability $\pi(i)$. Then, the cell loss probability at the input queue, $loss_{in}$, is given by

$$loss_{in} = \pi(K_{in}) \cdot P_B. \qquad (12)$$

Also, we have

$$\lambda_{in} = \lambda \cdot \pi(0) + [1 - \pi(0)]. \qquad (13)$$

3.2 Results and Discussion

Fig. 2 shows the maximum throughputs versus the network size with the dilation degree d as a parameter for the dilated banyan networks with and without recirculation. The dilated banyan network with recirculation clearly yields better performance than the one without recirculation. Moreover, by using recirculation, the maximum throughput of about 100% is achieved with d=3. For the 1024×1024 3-dilated banyan network with recirculation, Fig. 3 shows the cell loss probability versus the input load with the input queue size K_{in} as a parameter. One can see from the figure that the loss probability is below 10^{-6} for K_{in}=1 at the input load of 0.8.

While recirculation significantly improves the performance of the dilated banyan network, the switch has to resolve the out-of-sequence in the output module. For the resequencing function, cells arriving at the output module with the network delay longer than D_R are discarded. Thus, the parameter D_R should be selected so that

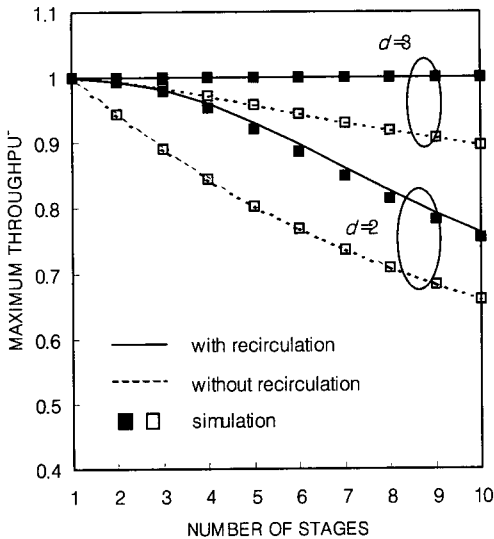the loss probability does not exceed the allowable limit.

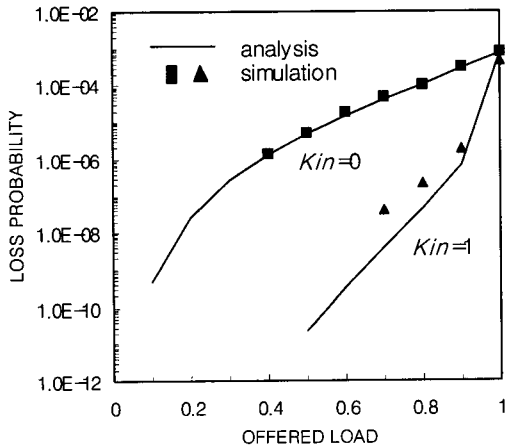Fig. 2 Maximum throughput versus the number of stages n for the banyan networks with and without recirculation.

Fig. 3 Loss probability versus the input load λ for the 1024×1024 3-dilated banyan network with recirculation

Fig. 4 shows the distribution of the network delay at full load for the dilation degrees 2 and 3 when the switch size is 1024×1024. We can see from the figure that in order to achieve the loss probability lower than 10^{-6}, D_R should be set to

the values higher than 24 and 8 for d=2 and d=3, respectively. But, considering that the resequencing delay and the required output buffer size increase as D_R increases, these values seem too large.
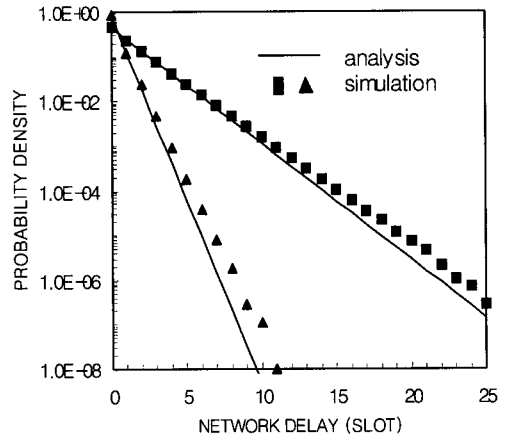

Fig. 4 Distribution of the network delay at full load for the 1024×1024 dilated banyan network with recirculation

IV. Contention Resolution by Priority Service

Priorities can be given to cells based on the number of recirculations to overcome the problem that the parameter D_R should be set to large values. In this section, we will study the performance of the dilated banyan network with recirculation when the contention resolution method by priority service is applied. As mentioned previously, the maximum priority level U is set to the same value as D_R. We first analyze the network to investigate the throughput performance. For the output queue, we will use simulations to find the performances for the delay and loss.

4.1 Performance Analysis

Similarly to the previous section, we assume that the input traffic pattern and the traffic flow inside the network are uniform. But the maximum

priority level U is set to a finite value, and thus the cell recirculated with the priority level U is discarded by the IC.

To represent the number of cells on x parallel links, we use the vector notation as $\vec{a} \equiv (a_0, a_1, \cdots, a_{2U+1})$ where a_i and a_{i+U+1}, $0 \le i \le U$, are the numbers of normal and deflected cells with priority level i, respectively. For x parallel links, the vector $\vec{a}$ must satisfy the following two conditions: i) $0 \le a_i \le x$ for $0 \le i \le 2U+1$, and ii) $0 \le \sum_{i=0}^{2U+1} \vec{a} \le x$. Let us define the set Λ_x as:

$$\Lambda_x \equiv \{ \vec{a} \mid 0 \le \vec{a} \le x \text{ for } 0 \le i \le 2U+1,$$

$$0 \le \sum_{i=0}^{2U+1} \vec{a} \le x \}. \qquad (14)$$

Then, the set Λ_d and Λ_{2d} will be the state spaces for one output port and two input ports of an SE, respectively.

Let $q_{k,t}(\vec{a})$ denote the probability that there are $\vec{a}$ cells on one output port of an SE in stage k ($1 \le k \le n$) at time slot t. In particular, let $q_{0,t}(\vec{a})$ denote the probability that there are $\vec{a}$ cells in an input module at time slot t. Then, the probability that two input ports of an SE in stage k receives $\vec{c} (\in \Lambda_{2d})$ cells at time slot t, $r_{k,t}(\vec{c})$, is given by

$$r_{k,t}(\vec{c}) = \sum_{\vec{a}, \vec{c}-\vec{a} \in \Lambda_d} q_{k-1,t}(\vec{a}) \cdot q_{k-1,t}(\vec{c}-\vec{a}) \qquad (15)$$

where $q_{k-1,t}(\vec{a}) \cdot q_{k-1,t}(\vec{c}-\vec{a})$ is the probability that there are $\vec{a}$ and $\vec{c}-\vec{a}$ cells from two SE's in stage k-1, respectively.

Now, to derive the probability $q_{k,t}(\vec{a})$, we introduce the following vector notations:

$$\vec{a}^n \equiv (a_0^n, a_1^n, \cdots, a_U^n)$$

$$= (a_0, a_1, \cdots, a_U)$$

$$\vec{a}^d \equiv (a_0^d, a_1^d, \cdots, a_U^d)$$

$$= (a_{U+1}, a_{U+2}, \cdots, a_{2U+1})$$

$$\vec{a}_i^n \equiv (a_i^n, a_{i+1}^n, \cdots, a_U^n)$$

$$\vec{a}_i^d \equiv (a_i^d, a_{i+1}^d, \cdots, a_U^d)$$

Then, probability $q_{k,t}(\vec{a})$ can be expressed as

$$q_{k,t}(\vec{a}) = \sum_{c \in \Lambda_{2d}} \Big\{ \prod_{i=0}^{U} q_{k,t}(a_i^n \mid \vec{a}_{i+1}^n, \vec{c})$$

$$\times \prod_{i=0}^{U} q_{k,t}(a_i^d \mid \vec{a}_{i+1}^d, \vec{a}^n, \vec{c}) \Big\} \cdot r_{k,t}(\vec{c}) \qquad (16)$$

where $q_{k,t}(a_i^n \mid \vec{a}_{i+1}^n, \vec{c})$ is the probability that a_i^n out of c_i normal cells with priority level i are routed to the output port, when there are $\vec{c}$ cells at two input ports and $\vec{a}_{i+1}^n$ normal cells are routed to one output port. Since the routing of the priority i cell is not affected by the cells with priority level lower than i, we have

$$q_{k,t}(a_i^n \mid \vec{a}_{i+1}^n, \vec{c})$$

$$= \begin{cases} \begin{bmatrix} c_i \\ a_i^n \end{bmatrix} 2^{-c_i}, & \text{if } a_i^n < d - m(\vec{a}_{i+1}^n) \\ \sum_{j=a_i^n}^{c_i} \begin{bmatrix} c_i \\ j \end{bmatrix} 2^{-c_i}, & \text{if } a_i^n = d - m(\vec{a}_{i+1}^n) \end{cases}$$

$$(17)$$

where $m(\vec{a}_i^n) \equiv \sum_{j=i}^{U} a_j^n$.

In (16), $q_{k,t}(a_i^d \mid \vec{a}_{i+1}^d, \vec{a}^n, \vec{c})$ is the probability that a_i^d deflected cells with priority level i are routed to the output port after $\vec{a}^n$ normal cells and a_{i+1}^d deflected cells are routed to the port. Under the conditions, let X and Y denote the numbers of empty links on the tagged output port and the other output port, respectively. Then, we have

$$X = d - m(\vec{a}^n) - m(\vec{a}_{i+1}^d) \qquad (18)$$

$$Y = [d - \{ m(\vec{c}^n) - m(\vec{a}^n) \}$$

$$- \{ m(\vec{c}_{i+1}^d) - m(\vec{a}_{i+1}^d) \}]. \qquad (19)$$

If $X=0$, there is no space for deflected cells with priority level i on the tagged output port. It means that a_i^d should be equal to 0. If $X>0$ and $Y=0$, then all of the deflected cells with priority

level i will be routed to the tagged output port. Thus a_i^q is equal to $c_{i+U+1}+D$ where D is the number of priority i cells deflected in this SE, and it is given by

$$D = \left[m(\vec{c}_i^n) - m(\vec{a}_i^n) - d \right]^+$$
$$- \left[m(\vec{c}_{i+1}^n) - m(\vec{a}_{i+1}^n) - d \right]^+. \qquad (20)$$

Finally, in case that $X>0$ and $Y>0$, assuming that b deflected cells with priority level i are originally destined for the tagged port, the set of the possible b, B, is given by

$$B = \{ b \mid a_i^d = \min[X, \ b+[c_{i+U+1}-b-Y]^+],$$
$$0 \le b \le c_{i+U+1} \}. \qquad (21)$$

Hence, the probability $q_{k,t}(a_i^d \mid \vec{a}_{i+1}^d, \vec{a}^n, \vec{c})$ is obtained as

$$q_{k,t}(a_i^d \mid \vec{a}_{i+1}^d, \vec{a}^n, \vec{c})$$

$$= \begin{cases} I(a_i^d), & X=0 \\ I(a_i^d - c_{i+U+1}-D), & X>0, \ Y=0 \\ \displaystyle\sum_{b \in B} \left[\begin{array}{c} c_{i+U+1} \\ b \end{array} \right] 2^{-c_{i+U+1}}, & X>0, Y>0 \end{cases}$$

$$\qquad (21)$$

where $I(x) = \begin{cases} 1, & \text{if } x=0 \\ 0, & \text{otherwise.} \end{cases}$

At the next time slot $t+1$, each IC removes the deflection tag from the recirculated cell, and increases its priority level by 1. Hence the probability $q_{0,t+1}(\vec{a})$ is given by

$$q_{0,t+1}(\vec{a})$$

$$= \begin{cases} \lambda_{in} \cdot \displaystyle\sum_{b \in \Lambda(\vec{a})} q_{n,t}(\vec{b}), \\ \qquad \text{if } a_0^n=1, \ m(\vec{a}^d)=0 \\ (1-\lambda_{in}) \cdot \displaystyle\sum_{b \in \Lambda(\vec{a})} q_{n,t}(\vec{b}), \\ \qquad \text{if } a_0^n=0, \ m(\vec{a}^n)<d, \ m(\vec{a}^d)=0 \\ \displaystyle\sum_{b \in \Lambda(\vec{a})} q_{n,t}(\vec{b}), \ \text{if } a_0^n=0, \ m(\vec{a}^n)=d \\ 0, \quad \text{otherwise.} \end{cases}$$

$$\qquad (23)$$

where

$$\Lambda(\vec{a}) \equiv \{ \vec{b} \mid \vec{b} \in \Lambda_d, \ b_i^d = a_{i+1}^n$$
$$\text{for } 0 \le i \le U-1 \}$$

In the steady state, the normalized throughput of the network ρ is given by

$$\rho = \sum_{a \in \Lambda_d} m(\vec{a}^n) \cdot q_n(\vec{a}) \qquad (24)$$

where $q_n(\vec{a})$ is the steady state value of $q_{n,t}(\vec{a})$. Since cells recirculated with priority level U are discarded by IC's, the probability that a cell is lost inside the network, $loss_{net}$, is given by

$$loss_{net} = \frac{\displaystyle\sum_{a \in \Lambda_d} a_U^d \cdot q_n(\vec{a})}{\rho + \displaystyle\sum_{a \in \Lambda_d} a_U^d \cdot q_n(\vec{a})}. \qquad (25)$$

Also, λ_{in} is obtained from (13) with

$$P_B = \sum_{\substack{a \in \Lambda_d, \\ m(\vec{a}^d)-a_U^d = d}} q_n(\vec{a}).$$

4.2 Results and Discussion

We first investigate the throughput and lost cell performances of the network. In Fig. 5, the maximum throughput of the 2-dilated network is shown as a function of U for $n=6$ and $n=10$, respectively. The maximum throughput increases as U increases, and then it is almost saturated at $U=2$. This phenomenon is very different from that of the network with back-pressure mechanism [13] in which the maximum throughput slowly decreases as U increases. It is because recirculated cells are not transmitted at the input ports at which they were transmitted in the previous time slot, and thus the traffic flow inside the network is properly distributed.

The increase of the maximum priority level U also improves the lost cell performance of the network. Fig. 6 shows the loss probability as a function of U for $d=2$ and $d=3$ with $n=10$ and $\lambda =1$. We can see in the figure that the loss probability rapidly decreases as U increases. Thus, the loss probability below 10^{-6} can be achieved with $U=6$ and $U=3$ for the networks of $d=2$ and $d=3$, respectively. These are very small compared

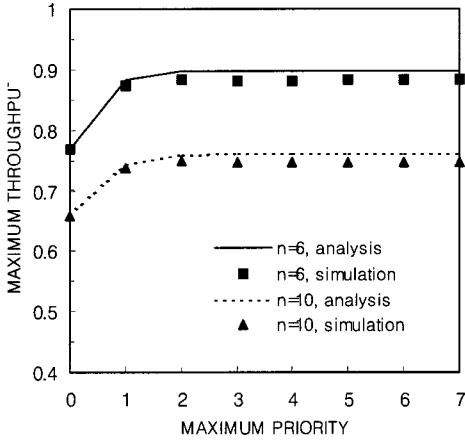to the values 24 and 8 required in the networks using the contention resolution method by random service.

Fig. 5 Maximum throughput of the 2-dilated banyan network versus the maximum priority level U for n=6 and n=10, respectively.
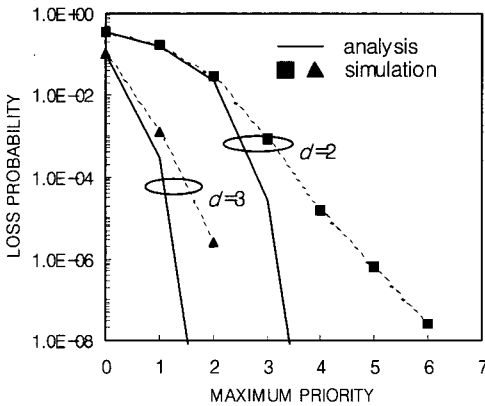

Fig. 6 Loss probability versus the maximum priority level U for d=2 and d=3 with n=10 and λ =1.

The figure also shows that the deviation between analytic and simulation results rapidly increases as U increases. It is caused by the assumption that the traffic flow inside the network is uniform. In fact, since the paths of recirculated cells are concentrated on some network outputs as the recirculation is repeated, the possibility of blocking is much higher than that under the uniform assumption. Accordingly, as U increases, the analytic results become more optimistic than simulation.

Fig. 7 plots the maximum throughput versus the number of stages for d=2 and d=3, respectively. The maximum priority level U is set to 2, because the maximum throughput of the network is saturated at that point. Comparing Figs. 2 and 7, we can see that the throughput performances for two contention resolution methods are almost the same. Therefore, by employing the priority service, we can effectively lower the maximum time stamp value D_R without losing the throughput performance.
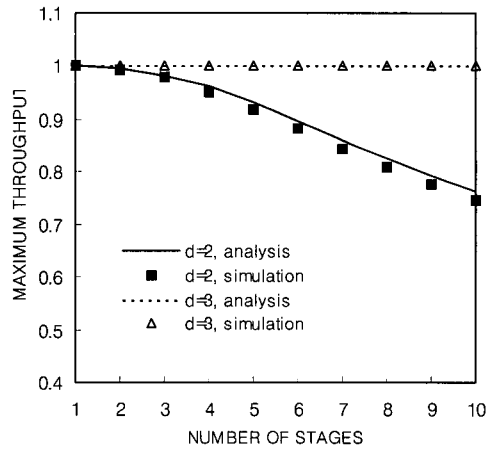

Fig. 7 Maximum throughput versus the number of stages n for d=2 and d=3, respectively

Now, we investigate the delay and lost cell performances in the output queue by means of simulation. The simulation results are obtained from a 1024×1024 3-dilated banyan network with K_{in}=1. Fig. 8 shows the loss probability versus the number of output buffers for different input loads, when the parameter U is set to 4. The loss probability decreases as the number of output buffers increases and as the input load decreases. Fig. 9 presents the loss probability in the output queue versus the number of output buffers K_o for U=2, 3 and 4, respectively, when λ =0.8. As expected, it is shown that more output buffers are

required to achieve a given loss probability when U is set to a higher value. That is, for the loss probability of 10^{-6}, the figure shows that output buffers of K_o=30, 32 and 33 are required in case of U=2, 3 and 4, respectively. Therefore, we can conclude that increasing the parameter U gives negative effects to the loss performance of the output queue, while it gives positive effects to the throughput and loss performances of the network.
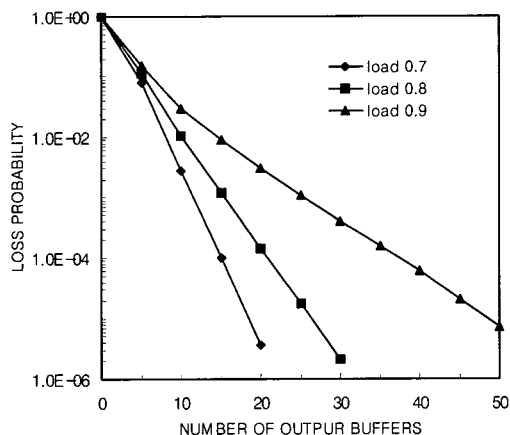

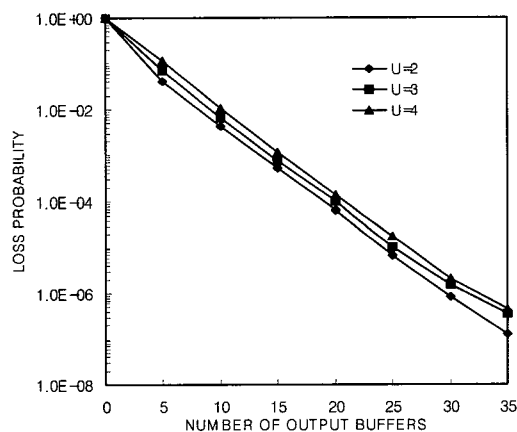Fig. 8 Loss probability versus the number of output buffers for different input loads when U=4

Fig. 9 Loss probability in the output queue versus the number of output buffers for U=2,3 and 4, respectively, when λ =0.8

V. Conclusion

In this paper, we have proposed a dilated banyan network with recirculation. The proposed switch utilizes the wasted routing capacity of the pure dilated banyan network by employing the deflection routing algorithm and recirculation links. In the output queue, a resequencing function is implemented to handle the out-of-sequence problem occurring in the network. Due to the resequencing, the performance of the switch is largely dependent on the methods that recirculated packets are handled when conflicts occur in the network. We first considered a contention resolution method by random service for processing of the recirculated packets, and analyzed the performance. For analysis, we assumed that the input traffic pattern and the traffic flow inside the network are uniform. From numerical and simulation results, we can see that the maximum throughput of the switch is greatly improved as compared to that of the pure dilated banyan network. However, to perform the resequencing function with the allowable loss probability, the switch requires a large number of output buffers and long resequencing delay. Hence, in order to reduce the problems, we considered a contention resolution method by priority service, and analyzed the throughput performance of the network under the uniform assumption. Numerical and simulation results show that the throughput performances of two contention resolution methods is almost the same. But, by employing the priority service, we can see that a smaller value of maximum time stamp value D_R is possible, and thus the required output buffers and resequencing delay decrease significantly.

Reference

[1] Ahmadi and W.E. Denzel, "A survey of modern high-performance switching techniques," IEEE J. Selected Areas Comm., 7, pp. 1091-1103, September 1989.

[2] A. Tobagi, "Fast packet switch archtectures for broadband integrated services digital net-

works," *Proc. IEEE*, 78, pp. 133-167, January 1990.

[3] Huang and S. Knauer, "Starlite: A wideband digital switch," *Proc. Globecom*, pp. 121-125, December 1984.

[4] Y.-S. Yeh, M. G. Hluchyj and A. S. Acampola, "The Knockout switch: A simple, modula architecture for high-performance packet switching," *IEEE J. Select Areas Comm.*, 5, pp. 1274-1283, October 1987.

[5] S. Turner, "Design of a broadcast packet switching network," *IEEE Trans. Comm.*, 36, pp. 734-743, June 1988.

[6] A. Tobagi and T. C. Kwok, "The tandem banyan switching fabric: A simple high-performance fast packet switch," *Proc. Infocom*, pp. 1245-1253, 1991.

[7] X. Brown and K.-H. Liu, "Neural network design of a banyan network controller," *IEEE, J. Select. Areas Comm.*, 9, pp. 1428-1438, October 1990.

[8] T. Lee, "Design and analysis of a new self-routing network," *IEEE Trans. Comm.,* 40, pp. 171-177, January 1992.

[9] Decina, P. Giacomazzi and A. Pattavina, "Shuffle interconnection networks with deflection routing for ATM switching: the closed loop shiffleout," *Proc. Infocom*, pp. 1254-1263, April 1991.

[10] Y. S. Youn and C. K. Un, "Performance analysis of cut-through buffered banyan networks with finite buffer size," *Performance Evaluation*, 25, pp. 293-311, 1995.

[11] A. Kumar and J. R. Jump, "Performance of unbuffered shuffle exchange networks," *IEEE Trans. Comput.*, 35, pp. 573-577, June 1986.

[12] T. Lee and S. C. Liew, "Broadband packet switchs based on dilated interconnection networks," *Proc. ICC*, pp. 255-261, 1992.

[13] Y. S. Youn and C. K. Un, "Performance of dilated banyan network with back-pressure mechanism," *Computer Comm.*, 19, pp. 972-981, October 1996.

윤 영 식(Young-sik Youn)   정회원
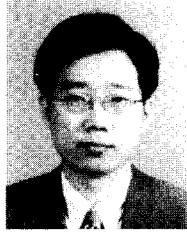
1986년 2월 : 서울대학교 전자공학과 졸업(학사)

1988년 2월 : 한국과학기술원 전기 및 전자공학과 (석사)

1993년 8월 : 한국과학기술원 전기 및 전자공학과 (박사)

1988년 3월~1995년 8월 :(주)디지콤 선임연구원
1995년 10월~현재 : (주)LG정보통신 책임연구원
<주관심 분야> ATM, 이동통신시스템