

우선순위 알고리즘을 적용한 상호연결 망 구조의 ATM 스위치 설계

Design of ATM Switch-based on a Priority Control Algorithm

박병수*, 조태경**, 조동욱***

상명대학교 컴퓨터시스템공학과*, 상명대학교 정보통신공학과**,
충북과학대학 정보통신학과***

Byoung-Soo Park(bpark@smu.ac.kr)*, Tae-Kyung Cho(tkcho@smu.ac.kr)**,
Dong-Uook Cho(ducho@ctech.ac.kr)***

요약

최근 ATM 스위치를 위한 대부분의 연구는 병렬 하드웨어 자체에 규칙성과 자체 라우팅 특성을 가지는 다단계 상호 연결 네트워크에 근거하여왔다. 그러나 네트워크는 동시에 또는 병렬로 전송되지만 서로 충돌을 피할 수 없다는 측면에서는 블러킹 네트워크라고 할 수 있는데, 주로 뱀안 네트워크가 그 구조에 사용되어왔다. 뱀안 형태의 스위치에 있어서 처리율을 증가시키고, 블러킹을 제거하기 위해서 즉 내부링크의 속도를 증가시키고, 모든 스위치 노드에 버퍼를 두고, 병렬로 다중 연결링크를 두고, 그 네트워크 전에 부하를 균등하게 하는 등 여러 가지 방법들이 사용되어 왔다.

따라서 본 논문에서는 모든 블러킹이 제거되고 하드웨어 복잡도를 향상시키기 위하여 재순환 서플익스체인지 네트워크의 사용을 제안하였다. 이 구성은 하드웨어 복잡도 면에서 한층 단순하여진 구조인 재순환 서플익스체인지 네트워크와, 동일한 목적으로 전달되는 패킷들에 있어서 우선순위가 결정된 후 순위가 높은 패킷은 다음 네트워크로 보내고, 순위가 낮은 패킷들을 재순환하는 트리구조의 순위 네트워크로 구성된다. 전송된 패킷은 뱀안 네트워크에서 분할 및 합성 알고리즘을 통하여 자체 라우팅 방식으로 최종 목적지에 전송되도록 구성된다. 처리율과 대기 시간 및 버퍼 크기에 따른 패킷의 손실율은 동일한 부하에 따라 각 포트에 도달한 패킷들의 확률을 이항분포로서 적용된다. 이때, 50%의 부하 정도면 버퍼 사이즈, $B_{size} = 15$ 이상 즉, 16이면 허용 가능한 손실율을 나타낸다. 그러므로 본 논문은 하드웨어의 복잡도 측면에서 기존의 바이토닉 정렬기를 재순환 서플익스체인지 네트워크로 구성하여 단순화 시켰다.

■ 중심어 : | 서플익스체인지 | 상호연결망 | 패킷 손실율 |

Abstract

Most of the recent researches for ATM switches have been based on multistage interconnection network known as regularity and self-routing property. These networks can switch packets simultaneously and in parallel. However, they are blocking networks in the sense that packet is capable of collision with each other. Mainly Banyan network have been used for structure. There are several ways to reduce the blocking or to increase the throughput of banyan-type switches: increasing the internal link speeds, placing buffers in each switching node, using multiple path, distributing the load evenly in front of the banyan network and so on.

Therefore, this paper proposes the use of recirculating shuffle-exchange network to reduce the blocking and to improve hardware complexity. This structures are recirculating shuffle-exchange network as simplified in hardware complexity and Rank network with tree structure which send only a packet with highest priority to the next network, and recirculate the others to the previous network. after it decides priority number on the packets transferred to the same destination, The transferred packets into banyan network use the function of self routing through decomposition and composition algorithm and all they arrive at final destinations. To analyze throughput, waiting time and packet loss ratio according to the size of buffer, the probabilities are modeled by a binomial distribution of packet arrival. If it is 50 percentage of load, the size of buffer is more than 15. It means the acceptable packet loss ratio. Therefore, this paper simplify the hardware complexity as use of recirculating shuffle-exchange network instead of bitonic sorter.

■ keyword : | Shuffle-Exchange | Interconnection network | packet loss ratio |

* 본 연구는 상명대학교 교내 연구과제로 수행되었습니다.

접수번호 : #041108-002

접수일자 : 2004년 11월 8일

심사완료일 : 2004년 12월 3일

교신저자 : 박병수, e-mail : bpark@smu.ac.kr

I. 서론

다양한 형태의 스위치 구조설계는 지난 몇 년 동안 주로 많은 연구가 ATM 스위치를 위한 연구에 집중되어 진행되어왔다[1, 6, 8, 9]. 고도의 병렬, 분산제어 시스템에 사용되었던 다단계 상호연결 네트워크를 기본으로 한 형태로써 고성능 스위치 구조 설계를 위하여 연구되고 그 결과가 구현되어왔다[2, 6, 10]. 이러한 몇 가지 이유는 먼저, 다단계 상호연결망 구조가 ATM 스위치를 구성할 때 VLSI 칩으로 된 여러 개의 서브모듈을 이용하여 하나의 스위치로 구현됨으로서 대용량을 고속으로 처리 할 수 있다는 것이며 또한 전체 스위치가 $\log_2 N$ 스테이지로 구성되고 각각의 스테이지는 $N/2$ 의 노드로 구성되는 내부 하드웨어 구조의 규칙성 때문에 자체 라우팅 함수를 적용하여 동기 또는 비동기 방식으로 경로 설정이 가능하기 때문이다.

그러나 이러한 네트워크는 동시에 또는 병렬로 전송되지만 서로 충돌을 피할 수 없다는 측면에서는 블러킹 네트워크라고 할 수 있는데, 주로 뱅안 네트워크가 그 구조에 사용되어왔다[6]. 뱅안 형태의 스위치에 있어서 패킷의 처리율을 증가시키고, 내부 및 외부 블러킹을 제거하기 위해서는 여러 가지 방법들이 사용되어 왔다. 즉 외부 속도와 관계되는 내부링크의 속도를 증가시키고, 모든 스위치 노드에 버퍼를 두고, 차단된 패킷의 전송을 지연시키기 위하여 스테이지 사이에 핸드셰이킹 구조를 사용하고[9], 각 스위치 연결에 다중 링크를 사용하거나 병렬 다중 연결 경로로 된 다중 네트워크를 사용하고, 네트워크 앞에 부하를 균등하게 하기위하여 분산 네트워크 등을 사용한다. 뱅안의 문제점을 보완하기위한 방법 중에 하나가 소트뱅크(Sort-Banyan) 형태이다.[5-7] 이 구조는 뱅안 네트워크 앞에 Batcher의 바이트닉 정렬기를 구성하고, 패킷이 목적지 주소에 따라 정렬한 후 전송함으로써 내부 충돌 없는 네트워크를 구성한다.

본 논문에서 제안된 스위치는 정렬의 기능을 수행하기 위하여 바이트닉 정렬기 대신에 하드웨어 복잡도 면에서 아주 탁월한 단일 스테이지로 구성된 재순환 서플익스체인지 네트워크를 사용한다, 기존의 정렬기에 비

하여 경로설정 알고리즘 적용에 매우 단순하여진다. 또한 트리구조를 갖는 순위 네트워크를 구성하여 우선순위를 정한 후 중복되는 패킷이 재순환 되도록 하여, 재순환되는 패킷은 버퍼로 들어가게 되며 각 버퍼의 전단에서 충돌될 경우 유입되는 패킷의 흐름이 유지되도록 제어된다. 반면에 전송에 실패한 패킷은 다음 타임 슬롯에 반복적으로 시도되어 전송한다. 다음 단계로 뱅안 네트워크에서는 재정돈성을 적용시킨 제안된 비트 분할 및 합성 알고리즘을 적용함으로써 내부 충돌 요소가 완전히 제거된 ATM 스위치 네트워크를 구성한다.

논문의 구성은, 2절에서 스위치의 연결 네트워크를 소개하고, 3절은 연결 알고리즘을 기술하며, 4절에서는 제안된 스위치의 성능 분석을 위한 처리율 및 손실율에 따른 적절한 버퍼의 크기 등에 대하여 설명하고, 끝으로 결론을 맺는다.

II. 스위치 연결 네트워크의 내부 구조

1. 서플익스체인지 네트워크

이 네트워크는 Batcher의 정렬기를 재순환 서플익스체인지 네트워크로 구성하여 하드웨어의 복잡도를 줄이는 효과가 있다. 스위치는 두개의 입력을 비교하여 작은 수는 'L'로, 큰 수는 'H'로 연결된다. 이것은 또한 재순환하는 과정에서 출력 방향을 위하여 플래그 비트를 결정하고, 그 플래그 비트가 '0'이면 변화 없고, '1'이면 재순환 출력을 하도록 한다.

Batcher의 알고리즘을 서플익스체인지 네트워크에서 각 스테이지마다 구현하기 위해 계속적으로 증가하는 m series를 사용한다. 이는 Batcher의 정렬 네트워크의 연속적인 스테이지의 피벗 비트가 $(i_0), (i_1, i_0), (i_2, i_1, i_0), \dots, (i_{m-1}, i_{m-2}, \dots, i_1, i_0)$ 와 같이 구성되기 때문이다. 예를 들면, 그림 1의 경우, $N=8$ 이다. 따라서 이 경우의 피벗 비트는 $(i_0), (i_1, i_0), (i_2, i_1, i_0)$ 이므로 6회의 재순환 스테이지가 필요하다. 재순환 서플익스체인지 네트워크에서 재순환과 플래그 비트의 결정 알고리즘의 적용으로 출력 SO_i 는 다음의 순위 네트워크로 연결된다.

2.2 순위(Rank) 네트워크

이것은 서플익스체인지로부터 연결되는 입력 RI_i 와 배안 네트워크로 연결되는 출력 RO_i 로 구성된다. 내부는 그림 2처럼 트리구조를 이용하여 입력의 순위를 각 단계(d)의 left-son을 기준으로 하고 right-son을 비교하여 기준 노드의 가중치(w)는 '0'이 되고 right-son의

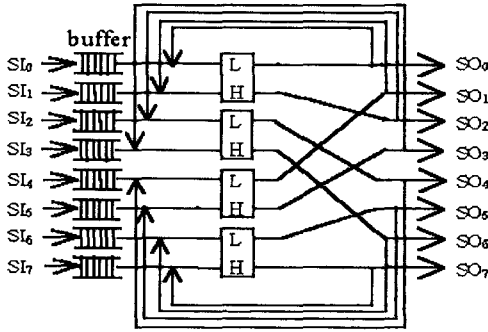


그림 1. 재순환 서플익스체인지 네트워크

가중치는 left-son과 같은 값을 나타내면 '0'이고, 크면 '1'로 설정된다. 모든 경로에 설정된 가중치에 따라 root 노드로부터 각 leaf 노드까지 최단경로의 가중치를 모두 합하면 그것이 그 입력 패킷의 순위가 된다. 먼저, 각 단계에서 기준되는 노드와 비교되는 노드의 쌍의 값은

$$(w_{m,2j}, w_{m,2j+1}) = \begin{cases} (0, 0) & \text{if } d_{m,2j} = d_{m,2j+1} \\ (0, 1) & \text{otherwise} \end{cases}$$

이며, 여기서 $1 \leq m \leq \log N$, $0 \leq j \leq 2^{m-1} - 1$ 이다.

모든 경로의 각노드에 대한 가중치의 합은

$$w_{total} = \sum_{\text{root}}^{\text{leaf node}} w_{\text{min path}}$$

로 계산된다. 계산된 순위로 중복되지 않는 패킷은 다음의 입력으로 연결되나, 중복되는 것은 이전 네트워크로 케환된 후 재전송된다.

2.3 배안-역배안 네트워크

이것은 배안-역배안의 구조로 되어있다. 순위 네트워크

크의 출력이 입력 BI_i 로 연결되고 출력 BO_i 가 패킷의 최종 목적지이다. 그 주소는 $b_{n-1} \dots b_i \dots b_0$ 로 나타내며, b_{n-1} 은 MSB고 b_0 은 LSB이다. 패킷의 분할을 위한 함수 Γ 는 다음과 같이 2개의 그룹으로 분할한다.

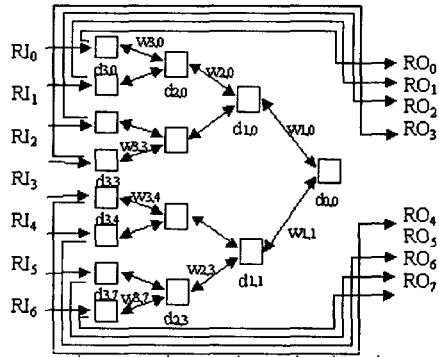


그림 2. 트리구조의 순위 네트워크

$$\Gamma(P_{k,j}) = \{P_{k+1,2j}\} \{P_{k+1,2j+1}\},$$

여기서, $0 \leq j \leq 2^k - 1$, $0 \leq k \leq n - 2$ 이다. k는 분할의 단계를 나타낸다. 그룹 $P_{k,j}$ 는 k 단계인 j번째 그룹을 의미하며, $N/2^k$ 개의 입력이 순서대로 배열된 집합으로 구성된다. k 단계의 2^k 그룹은 다음과 같다.

$$P_{k,0} = \{a_0^k, a_1^k, \dots, a_{N/2^k-1}^k\}$$

$$\vdots$$

$$P_{k,j} = \{a_{j(N/2^k)}^k, a_{j(N/2^k)+1}^k, \dots, a_{(j+1)(N/2^k)-1}^k\}$$

여기서, $(a_0^k, a_1^k, \dots, a_{N-1}^k)$ 는 k 단계에서 스위치의 입력들을 나타낸다.

정의 1. $2^n(0 \sim 2^{n-1})$ 의 정수는 고정된 비트에 따라 2^{n-1} 개의 원소들로 구성된 2개의 그룹으로 분리된다.

i번째 비트를 정하면, $(b_{n-1} b_{n-2} \dots b_{i+1} 0 b_{i-1} \dots b_0)$ 와 $(b_{n-1} b_{n-2} \dots b_{i+1} 1 b_{i-1} \dots b_0)$ 의 두 그룹으로 분리된다. 따라서, 그룹 $P_{1,0}$ 과 $P_{1,1}$ 로 나누어질 수 있으며, 반드시 2^{n-1} 개의 원소들을 포함하게 된다.

정의 2. $2^n(0 \sim 2^{n-1})$ 의 정수는 2^{n-1} 개의 그룹들로 분리될 수 있다.

정의 1을 적용하면 그룹 $P_{1,0}$ 은 2개의 그룹 $P_{2,0}$ 과 $P_{2,1}$

를, $P_{1,1}$ 은 2개의 그룹 $P_{2,0}$ 과 $P_{2,1}$ 를 각각 생성한다. 반복적으로 적용하면, 최종적으로 생성된 그룹들은 $P_{n-1,0}, P_{n-1,1}, \dots, P_{n-1,2^{n-1}-1}$ 로 총 2^{n-1} 개이다. 따라서, 정의 1, 2를 이 네트워크에 적용하기 위해, 고정 비트를 b_0 에서 b_{n-2} (역방향)로, 또는 b_{n-1} 에서 b_1 (순방향)로 순차적으로 정하면 분할된다.

특성 1. 역방향 분할과 순방향 분할은 다음 가정을 따른다.

역방향: $e_{odd}^k \wedge (2^{n-k}-1) = e_{even}^k \wedge (2^{n-k}-1)$ and

$$[e_{odd}^k \wedge (2^{n-k}-1)] \vee [e_{odd}^k \oplus 2^{n-k}] = e_{even}^k$$

순방향:

$$e_{odd}^k \wedge (2^{k+2^{n-1}-2^{k+1}}-2^k) = e_{even}^k \wedge (2^{k+2^{n-1}-2^{k+1}}-2^k),$$

and

$$[e_{odd}^k \wedge (2^{k+2^{n-1}-2^{k+1}}-2^k)] \vee [e_{odd}^k \oplus 2^k] = e_{even}^k$$

정리 1. 정의된 분할은 백안 네트워크 라우팅 알고리즘에 적용가능하다.

<증명> 백안 네트워크는 순환과정의 반복 구조를 갖고, 또한 특성 1로부터 모든 그룹들은 다른 그룹에 없는 원소들만을 포함하므로, 분할 과정은 하위 블록이 2×2 스위치의 크기와 같아질 때까지 계속되어, 고정된 비트를 제외한다면, 가장 작은 그룹들은 '1'과 '0'의 두 개의 입력을 갖게 되는 것이다. 따라서, 목적지 주소를 찾기 위해 분할의 역인 합성을 고려한다면, 정의된 분할에 따라 $0 \leq i \leq n-1$ 의 범위에서 역방향 합성은 각 스테이지에서 제어비트를 b_{n-1} 부터 b_0 로, 순차적으로 고정시키며, 순방향 합성은 b_0 부터 b_{n-1} 로 고정시키면 가능하다. 각 스위치 설정은 위에서 정의한 제어비트에 의해 정해지는데 만일 제어 비트가 '0'이면 입력은 상위출력으로, '1'이면 하위출력으로 연결되는 것이다.

위에서 제시한 역 방향 합성과 순방향 합성을 사용하여 역백안 네트워크를 위한 퍼뮤테이션은 새로운 방법을 통해 백안-역백안 다중 상호 연결 네트워크가 충돌 없이 목적지를 찾아갈 수 있도록 재배열하는 것이다.

정리 2. 제어비트가 '0'일 때는 상위출력 연결을, '1'일

때는 하위출력 연결임을 가정하면, 역 백안 네트워크의 퍼뮤테이션은 백안 네트워크의 출력이 역백안 네트워크의 입력으로 연결되었을 때 모든 스테이지에서 충돌이 발생하지 않는다.

<증명> 제어비트를 i 스테이지의 제어비트로 가정하면 $i=0$ 면 Property 1과 Lemma 1에서처럼 마지막 그룹은 고정된 비트를 제외하고 '1'과 '0'으로 구성되기 때문에 충돌 없이 교환될 수 있다. 따라서 i 단계에서 충돌이 없음을 가정하면, 다음과 같은 역 방향 합성을 얻을 수 있다.

$$[e_{odd}^k \wedge (2^{i+1}-1)] \vee [e_{odd}^k \oplus 2^{i+1}] = e_{even}^k$$

위의 식은 다음 스테이지에서는 충돌이 없음을 의미한다. 백안과 역백안 네트워크의 역상때문에, i 스테이지의 하부 블록의 다음 제어비트가 특성 1에 의해 항상 서로 보수 관계가 됨을 말한다. 결과적으로 역백안 네트워크는 위의 유도 가설에 의해 모든 스테이지에서 충돌 없이 연결됨이 증명된다.

3. 전송 알고리즘

본 논문에서 구성된 네트워크 상에서 패킷을 전송하기 위한 각각의 알고리즘을 다음과 같은 방법으로 연결할 수 있도록 구성한다.

```

Procedure : Rank 알고리즘
begin
  While  $0 \leq i \leq \log N - 1$  do
    // packet transmission to each leaf node //
     $d_{\log N, i} = RI_i$ 
  endwhile
  while  $\log N \geq m \geq 1$  do
    while  $0 \leq j \leq 2^{m-1} - 1$  do
      if  $(d_m, 2j = d_m, 2j+1)$ 
        // the comparison of nodes for weight //
         $w_m, 2j = 0$ 
         $w_m, 2j+1 = 0$ 
      else
         $w_m, 2j = 0$ 
         $w_m, 2j+1 = 1$ 
         $d_{m-1, j} = d_m, 2j$ 
      // transmission to upper level //
    endwhile
  endwhile
endwhile
  while  $1 \leq m \leq \log N - 1$  do
    // summation of total weight from root to leaf node//

```

```

while 0 ≤ j ≤ 2m-1-1 do
    Wm-1, 4j = Wm+1, 4j + Wm, 2j
    Wm+1, 4j+1 = Wm+1, 4j+1 + Wm, 2j
    Wm+1, 4j+2 = Wm+1, 4j+3 + Wm, 2j+1
    Wm+1, 4j+3 = Wm+1, 4j+3 + Wm, 2j+1
endwhile
endwhile
while 0 ≤ i ≤ logN - 1 do
    ROi = WlogN, i // decision of weight //
endwhile
while 0 ≤ i ≤ logN - 1 do // packet routing //
    if(ROi == ROi+1)
        SIi+1 = ROi+1
    else
        BIi+1 = ROi+1
    endwhile
endwhile
end

Procedure : decomposition 알고리즘
// Switch setting for banyan network at each stage //
begin
While n ≥ l > 1 do
    j := l mod n; // To start from b0 l
    bj := c; // don't care //
    Upper output[] := x0;
    Lower output[] := y0;
    Link_start := x0; // the initialization for link_start //
    While 1 ≤ m ≤ ⌊ N/2i+1 ⌋ do // i is stage //
        If ym-1 == xm & ym-1 ≠ Link_start
// checking the link connection //
            Upper output[] := xm;
            Lower output[] := ym;
        Else
            Upper output[] := xm;
            Lower output[] := ym;
            Link_start := xm;
        // the newly initialization for link_start //
        endwhile;
    endwhile;
end

```

```

Procedure : composition 알고리즘
// Switch setting for banyan1 network at each stage //
begin
While 0 ≤ i ≤ n - 1 do
    con_biti = b(n-i) mod n; // To start from b0 //
    If con_biti == 0
        goto upper output ;
    Else
        goto lower output ;
    endwhile;
end

```

4. 성능분석

이 스위치는 각각의 입력에 독립적인 N 개의 버퍼로 구성되며 그것은 블러킹을 고려하여 효과적인 서비스 타임 분배로 구성된다. 각 포트의 버퍼는 타임 슬롯에 따라 유입되는 패킷을 FIFO의 기법에 따라 처리된다. 이 때 유입되는 패킷은 슬롯마다 각 포트에 대한 전송

되는 트래픽, p 의 확률을 가지며 그것은 어느 한 출력 포트에 도달하는 확률은 동일한 $1/N$ 의 확률을 갖는다고 가정한다.

블러킹의 상태는 순위 네트워크의 출력의 순위에 따라 상태 함수로 나타낼 수 있다.

$$\Phi(RO_i) = \begin{cases} 1 & \text{if } RO_i \neq RO_{i+1} 0 \leq i < N-1 \\ 0 & \text{if } RO_i = RO_{i+1} 0 \leq i < N-1 \end{cases} \quad (1)$$

그러므로 배안 네트워크의 각 출력으로 도달될 수 있는 모든 패킷의 수는 다음 식으로 표현된다. (Σ 는 $\Sigma_{i=0}^{N-1}$ 을 의미함)

$$\Sigma \Phi(RO_i) \quad (2)$$

이 스위치의 구조에서 패킷의 처리율은 정상 상태에서 서 유입되는 타임 슬롯마다 확률 p 이므로 그것은 기대값으로 표현될 수 있으며, 그에 대한 식은 다음과 같다.

$$p = (1/N) \Sigma E\{\Phi(RO_i)\} = E\{\Phi(RO_i)\} \quad (3)$$

또한, 한 타임 슬롯 동안 블러킹된 패킷의 수를 $RO_{blocked}$ 라 하고, 이전 타임 슬롯에 블러킹되지 않은 패킷의 수를 RO_i^{pre} 라고 하면

$$RO_{blocked} = \Sigma RO_i - \Sigma \Phi(RO_i) + \Sigma RO_i^{pre} \quad (4)$$

을 얻는다. 위의 (4)식을 스위치 처리율의 관계식으로 유도하기 위하여, 양변에서 기대값을 구하고, 양변을 N 으로 동일하게 나누고 유도하면,

$$E\{\Phi(RO_i)\} = E\{RO_i\} + E\{RO_i^{pre}\} - E\{RO_{blocked}\} / N \quad (5)$$

을 얻을 수 있다. 이와 같이 처리율을 얻고, 모든 패킷의 블러킹된 패킷의 기대값을 구하기 위하여 임출력 패킷의 보존 관계식을 이용한다. 우선, 이전에 타임 슬롯에 블러킹되지 않은 패킷의 확률을 정상 상태에서 α 라고 가정한다.

$$E\{\Sigma \Phi(RO_i)\} \alpha = p \Sigma RO_i \quad (6)$$

그러면 위와 같은 식을 얻는다. 식 (5)를 다시 식 (6)에 대입하면 다음과 같은 식을 구할 수 있다.

$$E(RO_{blocked})/N = 1 + E(RO_i^{pre})/N - p/\alpha \quad (7)$$

식 (7)은 패킷이 성공적으로 목적지에 도달한다는 조건에 따라 도착율을 측정하는 트래픽 비율 p 와 포화 정도 α 를 나타내는 식이다. 따라서 블러킹되지 않은 패킷의 수를 최대로 하기 위하여 식 (6)로부터 $\sum RO_i = N$ 은 상수이므로 α 의 값을 정하면 된다. 스위치 처리율 p 는 α 가 "1"의 값을 가질 조건에서 최대의 처리율을 나타내고, 또한 정상 상태에서 입력 포트의 수는 최대 N 이상을 가질 수 없으므로 $E(RO_i^{pre}) = p$ 의 조건을 만족하고, $N \gg 1$ 의 조건에 따라서 식 (7)을 다음과 같은 관계식으로 나타내진다.

$$E(RO_{blocked})/N = 1 - p \quad (8)$$

이러한 과정으로써, 한 타임 슬롯 동안 i 번째 패킷이 임의의 버퍼에 도착할 확률을 χ_i 라고 하면 다음과 같은 식으로 표현 가능하다.

$$\chi_i = \binom{N}{i} \left(\frac{p}{N}\right)^i \left(1 - \frac{p}{N}\right)^{N-i} \text{ for } i = 0, 1, \dots, N \quad (9)$$

또한, Random Variable, χ_i 의 확률 생성 함수, $X(z)$ 로 큐, $Q(z) = (1-p)(1-z)/(X(z)-z)$ 에서 패킷의 수에 관한 확률 생성 함수를 얻기 위하여 사용될 수 있다. 이 함수로부터 평균 큐의 크기를 계산하기 위하여 dQ/dz 를 구한 후 르피탈 정리를 이용하여 극한 값을 구하면 다음과 같은 결과를 얻는다.

$$\lim_{z \rightarrow 1} dQ/dz = p^2 N / 2(1-p)(N-1) \quad (10)$$

이상에서 본 바와 같이, 만일 $N \rightarrow \infty$ 로 접근한다면, 이 스위치 네트워크에 있어서 패킷이 목적지 RO_i 로 향하여 이동할 때 그것은 확률 p 에 따라 포아송 분포를 이룬다. 이것은 곧, 정상 상태에서 식 (10)에서 $N \rightarrow \infty$ 일 때, $E(RO_i)$ 의 값이 출력 큐잉에서 설명되는 M/D/1 시스템을 적용시키면 다음과 같은 식을 얻을 수 있다.

$$E(RO_i) = p^2 / 2(1-p) \quad (11)$$

위에서 계산결과로 각 포트 마다 최대 스위치 처리율을 계산하기 위하여 식 (3)과 식 (8) 그리고 식 (11)를

(5)에 대입하면 처리율 p 만의 방정식으로 나타내고, $N \rightarrow \infty$ 일 때, 제안된 스위치의 처리율은 최대 $p_{max} = 0.586$ 임을 알 수 있다.

따라서, 입력큐를 사용하면 최대로 얻어질 수 있는 부하율이 58.6%이므로 일반적으로 출력 큐를 사용하였을 때 나타나는 부하율 약 80%보다는 성능면에서 많이 떨어지는 단점이 있다. 이 결과로부터 입력 큐의 패킷의 수를 Random Variable K 라고 가정하고, 그 큐의 평균 길이를 구하기 위하여 다음과 같이 적용한다.

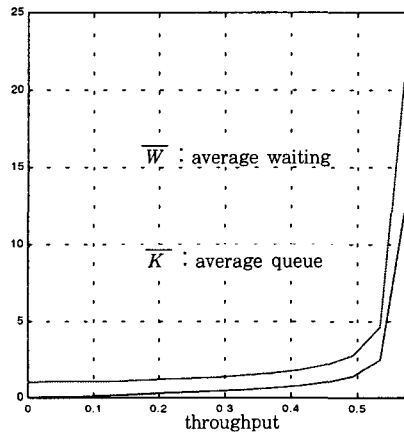


그림 3. 처리율의 평균 큐의 길이 및 대기시간

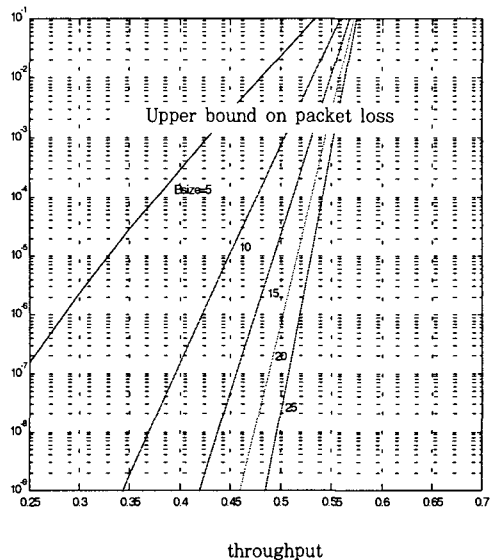


그림 4. 패킷 손실율과 버퍼의 크기

결과적으로, K 의 평균 기대값이 다음과 같이 주어지고

$$\bar{K} = \frac{p(2-p)(1-p)}{(2-\sqrt{2}-p)(2+\sqrt{2}-p)} \quad (12)$$

위의 식으로부터 다음과 같은 결과를 얻는다.

$$\begin{aligned} p_0 &= (1-\bar{\omega}_p)(1-p) \\ p_k &= (1-\bar{\omega}_p)(1-p)\bar{\omega}_p + p)\omega_p^{k-1} \\ &= \frac{p(2-p)}{2(1-p)}(1-\bar{\omega}_p)\bar{\omega}_p^{k-1}; k > 0 \end{aligned}$$

큐의 평균 길이 \bar{K} 와 평균 대기 시간 \bar{W} 을 구하기 위하여 주어진 p 에 대하여 Little's 이론($\bar{W} = \bar{K}/p$)을 식 (12)과 같이 적용하여 그림 3에 \bar{K} 와 \bar{W} 를 도시한다. 그림 3에서와 같이 처리율은 58%에서 최대를 나타내고 있다.

또한, 유한한 버퍼 크기(B_{size})에 대한 버퍼 초과 확률은 $\bar{K} > B_{size}$ 의 확률에 따라 패킷 손실율의 상위 한계점이 된다. 따라서,

$$P(\text{loss}) < P(\bar{K} > B_{size}) = \frac{p(2-p)}{2(1-p)} \omega_p^{B_{size}} \quad (13)$$

그림 4는 식 (13)에 따라 패킷 손실율에 대한 상위 한계점을 나타낸다. 이것으로 스위치의 최대 부하율인 58.6%가 부하되면, 무한 버퍼 사이즈임을 알 수 있다. 그러나 50%의 부하 정도면 버퍼 사이즈, $B_{size} = 15$ 이상 즉, 16이면 허용 가능한 손실율을 나타내며, $B_{size} = 20$, 또는 25 이상에서는 패킷의 손실율이 극히 낮은 것으로 나타난다. 또한 45% 정도면 버퍼 사이즈, $B_{size} = 10$ 이라도 무방함을 알 수 있다.

5. 결론

본 논문에서는 정렬기로서 서플익스체인지 네트워크, 순위 네트워크 그리고 뱅안 네트워크로 구성된 ATM 스위치를 제안하였다. 우선, 하나의 스테이지로 구성되어서 하드웨어의 복잡도 면에서는 탁월하게 단순화된 서플익스체인지 네트워크를 구성하고, 이를 통하여 목

적지 정보로 정렬된 패킷의 결과에 따라 트리 구조인 순위 네트워크에서 각각 동일한 목적지 주소를 갖는 패킷들에 대하여는 우선순위를 부여하기 위한 알고리즘 적용하였다. 우선순위가 결정되면, 그 순위가 가장 높은 패킷을 제외한 나머지 모든 패킷은 재전송되기 위하여 서플익스체인지 네트워크로 재입력된다. 우선순위가 가장 높은 패킷은 다음 단계인 뱅안 네트워크로 전송되어 서로 구별되는 목적지 주소를 갖게 된다. 이 때 역뱅크 자체 라우팅 네트워크가 되도록 뱅안 네트워크에서 스테이지마다 목적지 주소의 기준 비트를 정하여 두개의 데이터 집합 형태로 반복 분리함으로써 다음 스테이지의 정보 없이도 스위치의 연결이 가능하도록 알고리즘을 설계하였다. 또한 역뱅크 네트워크에서는 앞의 기준 비트의 역순으로 스테이지마다 스위치 연결 설정을 위한 알고리즘을 제안하였다. 이렇게 구성된 스위치의 처리율과 대기 시간 및 버퍼 사이즈에 따른 패킷의 손실율을 분석한 결과 스위치의 부하율은 최대 58.6%까지 가능하나, 이 때의 버퍼의 크기는 현실성이 배제되므로, 실용 가능한 버퍼의 크기로 구현하기 위하여 부하율을 다음과 같이 적용하면, 55%인 경우 대략 40이며 45%의 경우는 버퍼 사이즈가 10이면 수용할 만한 패킷의 손실율을 갖는다. 그러므로 본 논문은 하드웨어의 복잡도 측면에서 기존의 바이토닉 정렬기를 재순환 서플익스체인지 네트워크로 구성하여 단순화 시켰다.

참고 문헌

- [1] H. Ahmadi, and W. E. Denzel, "A survey of Modern High-performance Switch techniques," IEEE Journal on Selected Areas in Communication, Vol.7, No.7, Sept. 1989.
- [2] J. W. Byun, and T. T. Lee, "The design and analysis of an ATM multicast switch adaptive traffic controller," IEEE/ACM Trans. Networking, Vol.2, June 1994. pp. 288-298.
- [3] W. T. Chen, and Y. W. Deng, "PPCN : A

high-performance copy network for large scale ATM switching systems," IEICE Trans. Commun., Vol. E82-B, No.1 pp. 1-13, Jan. 1999.

- [4] K. Y. Eng., M. G. Hluchyj, and Y. S. Yeh, "A growable packet switch architecture : Design principles and applications" IEEE Trans. Commun., Vol.40, pp. 423-430, Feb 1992.
- [5] J. N. Giacomelli, W. D. Sinkoskie, and M. Littlewood, "Sunshine : A high performance self-routing broadband packet switch architecture," in Proc. ISS'90, pp. 123-129, 1990.
- [6] A. Huang and S. Knauer, "Starlite : A Wideband Digital Switch," Proceeding of Globecom '84, pp. 121-125
- [7] Y. N. J. Hui and E. Arthurs, "A Broadband Packet Switch for Integrated Transport," IEEE Journal on Selected Areas in Communication, Vol.5, No.8, pp. 1264-1273, Oct. 1987.
- [8] R. Kannan, R. Bartos, K. Lee, and H. Jordan, "SXmin : A self-routing high-performance ATM packet switch based on group-knockout principles," IEEE Trans. Commun., Vol.45, No.6, pp. 710-722, June 1997.
- [9] J. S. Turner, "Design of a Broadcast packet switching network," in Proc. INFOCOM'86, pp. 667-675. Also, Dep. Comput. Sci., Wastomgton Univ., St. Louis, MO, Tech. Rep. Wucs-84-4, Mar. 1985.
- [10] J. S. Turner and L. F. Wyatt, "A packet network architecture for integrated services," in Proc. GLOBECOM'83. San Diego. CA. pp. 211-216, Nov. 1986.

저자 소개

박 병 수(Byoung-Soo Park)

정회원



- 1986년 2월 : 한양대학교 전자공학(공학사)
- 1989년 8월 : 한양대학교 전자공학(공학석사)
- 1994년 5월 : 텍사스 A&M(공학박사)

• 1995년 3월~현재 : 상명대학교 컴퓨터시스템공학과 교수

<관심분야> : 임베디드 시스템, 병렬 알고리즘

조 태 경(Tea-Kyung Cho)

정회원



- 1984년 2월 : 한양대학교 전자통신공학과(공학사)
- 1986년 2월 : 한양대학교 전자통신공학과(공학석사)
- 2001년 9월 : 한양대학교 전자통신공학과(공학박사)

• 2003년 9월~현재 : 상명대학교 정보통신공학과 교수

<관심분야> : 초고속정보통신망, e-Learning

조 동 옥(Dong-Uk Cho)

정회원



- 1983년 2월 : 한양대학교 전자공학(공학사)
- 1985년 2월 : 한양대학교 전자공학(공학석사)
- 1989년 2월 : 한양대학교 전자통신공학과(공학박사)

• 2000년 3월~현재 : 충북과학대학 정보통신학과 교수

<관심분야> : 멀티미디어 시스템, 영상처리