우선순위 제어기법을 기반으로 한 재순환 Shuffle-Exchange 상호연결 ATM 스위치

박 병 수†

요 약

본 논문에서 제안된 ATM 스위치의 구성은 패킷의 내부충돌이 제거되도록 구성된 다단계 상호연결망 스위치 구조이며, 이는 하드웨어 복잡도 면에서 향상된 재순환 shuffle-exchange와, 동일한 목적지로 전달되는 패킷들은 우선순위가 정해진 후 한 개를 제외한 우선순위가 낮은 패킷들이 재순환 되는 트리구조의 순위 네트워크로 구성된다. 또한, 전달된 패킷은 벤안 네트워크에서 분할 및 합성 알고리즘으로 최종 목적지에 도달되도록 구성한다. 처리율과 대기 시간 및 버퍼 크기에 따른 패킷의 손실율은 동일한 부하에 따라 각 포트에 도달한 패킷들의 확률분포을 이항분포로서 적용한다.

Recirculating Shuffle-Exchange Interconnection ATM Switching Network Based on a Priority Control Algorithm

Byoung-Soo Park†

ABSTRACT

This paper proposes a multistage interconnection ATM switching network without internal blocking. The first is recirculating shuffle-exchange network improved on hardware complexity. The next is connected to Rank network with tree structure. In this network, after the packets transferred to the same output ports are given each priority, only a packet with highest priority is sent to the next, and the others are recirculated to the first. Rearrangeability through decomposition and composition algorithm is applied for the transferred packets in banyan network 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.

1. 서 론

지난 몇 년 동안 많은 다양한 스위칭 구조설계, 특히 ATM 스위치를 위한 연구가 활발이 진행되어왔다[1, 6, 8, 9]. 이러한 고성능 스위칭 구조 설계를 위한 현재의 접근은 주로 고도의 병렬, 분산제어 시스템에 사용되었던 다단계 상호연결 네트워크를 기본으로 한 형태로 구현되어왔다[2, 6, 10]. 이는 고속의 대용량을 요구하며, 내부 충돌로 인한 패킷 손실율을 저하시키지 않도록하기 위한 하드웨어 내부에서 수행되는 자체 라우팅 함수를 주로 이용하므로 VLSI 칩 모듈로 구현하기에 적합한 스위치 형태이다.

다단계 상호연결 네트워크중 벤안 네트워크가 주로 사용되어왔는데[6] 이것은 그 자체로는 충돌이 없는 조건을 충족시키지 못한다. 따라서 패킷들의 충돌을 감

※ 본 논문은 상명대학교 '99학년도 교내학술연구비 지원에 의하여 연구되었음
† 정 회 원 : 상명대학교 컴퓨터정보통신학부 교수
논문접수 1999년 2월 5일, 심사완료 2000년 1월 12일

소시키고, 또한 밴얀형태의 스위치 처리율을 향상시키기 위하여 외부처리 속도에 비하여 내부 연결 속도를 증가시키거나, 모든 스위칭 모드에 버퍼를 설치하고, 차단된 패킷의 전송을 지연시키기 위하여 스테이지 사이에 핸드쉐이킹 구조를 사용한다[9]. 이러한 점을 해결하기위한 형태가 sort-banyan 형태이다[5-7]. 이 형태는 밴얀 네트워크 앞에 Batcher의 bitonic sorter 를 구성하고, 패킷이 목적지 주소에 따라 정렬된 후 연결됨으로서 내부 충돌 문제를 해결하였다

본 논문에서 제안된 스위치는 sorter의 기능을 수행하기 위해 단일 스테이지로 구성된 재순환 shuffle-exchange의 네트워크을 이용하므로, 기존의 sorter에 비하여 하드웨어 복잡도 면에서 매우 단순해 진다. 또한 트리구조를 갖는 순위 네트워크를 구성하여 우선 순위를 정한 후 중복되는 패킷이 재순환 되도록 한다 이렇게 재순환되는 패킷은 버퍼로 들어가게 되며 각 버퍼의 전단에서 충돌될 경우 유입되는 패킷의 흐름이 유지되도록 제어된다. 전송에 실패한 패킷은 다음 타임슬럿에 재시도 된다. 다음 단계로 밴얀 네트워크에서 재정돈성을 적용시킨 제안된 비트 분할 및 합성 알고리즘을 이용하여 내부 충돌 요소가 제거된 스위치 네트워크을 구성한다.

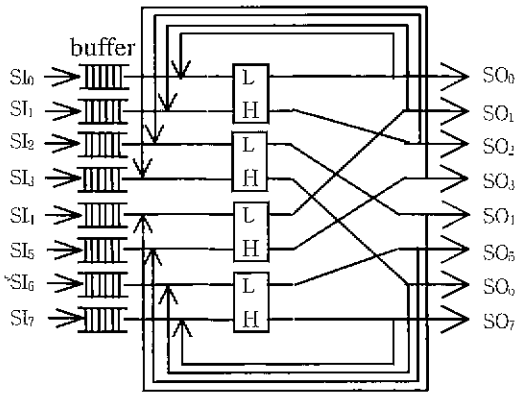
논문의 구성은, 2절에서 스위치의 연결 네트워크를 소개하고, 3절은 연결 알고리즘을 기술하며, 4절에서는 제안된 스위치의 성능 분석을 위한 처리율 및 손실율에 따른 적정한 버퍼의 크기 등에 대하여 설명하고, 끝으로 결론을 맺는다.

2. 스위치 연결 네트워크의 내부 구조

2.1 Shuffle-exchange 네트워크

이 네트워크는 Batcher의 sorter를 재순환 shuffle-exchange 네트워크로 구성하여 하드웨어의 크기를 줄이는 효과가 있다. 스위치는 두개의 입력을 비교하여 작은 수는 'L'로, 큰수는 '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 shuffle-exchange 네트워크에서 각 스테이지마다 구현하기 위해 계속적으로 증가하는 m serics를 사용한다. 이는 Batcher 의 sorting 네트워크의 연속적인 스테이지의 피봇 비트가 (i_0), (i_1, i_0), (i_2, i_1, i_0), .. , $(i_{m-1}, i_{m-2}, ..., i_1, i_0)$와 같이 구성되기 때문이다. 예를 들면, (그림 1)의 경우, $N=8$이다. 따라서 이 경우의 피봇비트는 (i_0), (i_1, i_0), (i_2, i_1, i_0)이므로 6회의 재순환 스테이지가 필요하다. 재순환 shuffle-exchange 네트워크에서 재순환과 플래그 비트의 결정 알고리즘의 적용으로 출력 SO_i는 다음의 순위 네트워크로 연결된다.



(그림 1) 재순환 shuffle-exchange 네트워크

2.2 순위(Rank) 네트워크

이것은 shuffle-exchange로부터 연결되는 입력 RI_i와 밴얀 네트워크로 연결되는 출력 RO_i로 구성된다 내부는 (그림 2)처럼 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 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

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

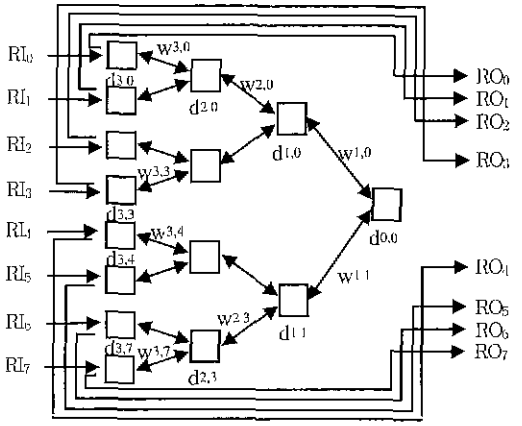
이며, 여기서 $1 \le m \le \log N,\ 0 \le j \le 2^{m-1}-1$이다

모든 경로의 각노드에 대한 가중치의 합은

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

로 계산된다. 계산된 순위로 중복되지 않는 패킷은 다음의 입력으로 연결되나, 중복되는 것은 이전 네트워

크로 궤환된 후 재전송된다.



(그림 2) 트리구조의 순위 네트워크

2.3 밴얀과 역밴얀(Banyan-Banyan^{-1}) 네트워크

이것은 밴얀-역밴얀의 구조로 되어있다. 순위 네트워크의 출력이 입력 BI_i로 연결되고 출력 BO_i가 패킷의 최종 목적지이다. 그 주소는 $b_{n-1}..b_i...b_0$로 나타내며, b_{n-1}은 MSB고 b_0는 LSB이다. 패킷의 분할을 위한 함수 Γ는 다음과 같이 2개의 그룹으로 분할한다.

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

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

$$P_{k,0} = \{\alpha_0^k, \alpha_1^k, \cdots, \alpha_{N/2^k-1}^k\}$$
$$\vdots$$
$$P_{k,j} = \{\alpha_{j(N/2^k)}^k, \alpha_{j(N/2^k)+1}^k, \cdots, \alpha_{(j+1)(N/2^k)-1}^k\}$$

여기서, $(\alpha_0^k, \alpha_1^k, \cdots, \alpha_{N-1}^k)$는 k단계에서 스위치의 입력들을 나타낸다.

Definition 1 $2^n(0\sim2^n-1)$의 정수는 고정된 비트에 따라 2^{n-1}개의 원소들로 구성된 2개의 그룹으로 분리된다.

i번째 비트를 집하면, $(b_{n-1}b_{n-2}..b_{i-1}0b_{i-1}..b_0)$와 $(b_{n-1}b_{n-2}...b_{i+1}1b_{i-1}..b_0)$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.

Definition 2 $2^n(0\sim2^n-1)$의 정수는 2^{n-1}개의 그룹들로 분리될 수 있다.

Definition 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 $P_{n-1,0}$, $P_{n-1,1}$, ⋯, $P_{n-1,2^{n-1}-1}$로 총 2^{n-1}개이다. 따라서, Definition 1, 2를 이 네트워크에 적용하기 위해, 고정 비트를 b_0에서 b_{n-2}(역방향)로, 또는 b_{n-1}에서 b_1(순방향)로 순차적으로 정하면 분할된다.

Property 1 역방향 분할과 순방향 분할은 다음 가정을 따른다.

$$e_{odd}^k \in \{P_{k,j} | j\text{는 홀수}\},\ e_{even}^k \in \{P_{k,j} | j\text{는 짝수}\}.$$

만일 j가 홀수나 짝수인 그룹에서 원소가 선택되었다면, 다음 식을 만족하는 원소가 상반되는 그룹에 존재해야 한다.

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

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

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

$$\lfloor e_{odd}^k \wedge (2^{k+2^{n-1}-2^k+1}-2^r) \rfloor \vee \lfloor e_{odd}^k \oplus 2^k \rfloor = e_{even}^l$$

Lemma 1 정의된 분할은 밴얀 네트워크 라우팅알고리즘에 적용가능하다.

<증명> 밴얀 네트워크는 순환과정의 반복 구조를 갖고, 또한 Property 1로부터 모든 그룹들은 다른 그룹에 없는 원소들만을 포함하므로, 분할 과정은 하위 블럭이 2×2 스위치의 크기와 같아질 때까지 계속되어, 고정된 비트을 제외한다면, 가장 작은 그룹들은 '1'과 '0'의 두 개의 입력을 갖게되는 것이다. 따라서, 목적지 주소를 찾기 위해 분할의 역인 합성을 고려한다면, 정의된 분할에 따라 $0 \le i \le n-1$의 범위에서 역방향 합성은 각 스테이지에서 제어비트를 b_{n-1}부터 b_0로, 순차적으로 고정시키며, 순방향 합성은 b_0부터 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 역밴얀 네트워크를 위한 퍼뮤테이션은 새로운 방법을 통해 밴얀-역밴얀 다중 상호 연결 네트워크가 충돌 없이 목적지를 찾아갈 수 있도록 재배열하는 것이다.

Lemma 2 : 제어비트가 '0'일 때는 상위출력 연결을, '1'일 때는 하위출력 연결임을 가정하면, 역 밴얀 네트워크의 퍼뮤테이션은 밴얀 네트워크의 출력이 역밴얀 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.

$$\lfloor e^k_{odd} \wedge (2^{i+1}-1) \rfloor \vee \lfloor e^k_{odd} \oplus 2^{i+1} \rfloor = e^k_{even}$$

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

3. 라우팅 알고리즘

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

Procedure : 순위 알고리즘
begin
 While $0 \le i \le \log N - 1$ do
// packet transmission to each leaf node //
  $d_{\log N, i} = RI_i$
 endwhile
 while $\log N \ge m \ge 1$ do
  while $0 \le j \le 2^{m-1}-1$ do
   if ($d_{m,2j} == d_{m,2j+1}$)
// the comparison of nodes for weight //
    $w_{m,2j} = 0$
    $w_{m,2j+1} = 1$
   else
    $w_{m,2j} = 0$
    $w_{m,2j+1} = 1$
    $d_{m-1,j} = d_{m,2j}$
 // transmission to upper level //
  endwhile
 endwhile
 while $1 \le m \le \log N - 1$ do
 // summation of total weight from root to leaf node//
  while $0 \le j \le 2^{m-1}-1$ do
   $w_{m-1,4j} = w_{m+1,4j} + w_{m,2j}$
   $w_{m+1,4j+1} = w_{m+1,4j+1} + w_{m,2j}$
   $w_{m+1,4j+2} = w_{m+1,4j+2} + w_{m,2j+1}$
   $w_{m+1,4j+3} = w_{m+1,4j+3} + w_{m,2j+1}$
  endwhile
 endwhile
 while $0 \le i \le \log N - 1$ do
  $RO_i = W_{\log N, i}$ // decision of weight //
 endwhile
 while $0 \le i \le \log N - 1$ do // packet routing //
  if($RO_i == RO_{i-1}$)
   $SI_i = RO_i$
  else
   $BI_{i-1} = RO_{i-1}$
 endwhile
end

Procedure : 분할 알고리즘
// Switch setting for banyan network at each stage //
begin
 While $n \ge I > 1$ do
  $I = I \bmod n$, // To start from b_0 //
  $b_I = c$, // *don't care* //
  Upper output[] := x_0,
  Lower output[] := y_0,
  Link_start := x_0, // the initialization for link_start //
  While $1 \le m \le \lfloor N/2^{i+1} \rfloor$ do // i is stage //
   If $y_{m-1} == x_m$ & $y_{m-1} \ne$ *Link_start*
// checking the link connection //
    Upper output[] := x_m;
    Lower output[] := y_m;
   Else
    Upper output[] := x_m;
    Lower output[] := y_m;
    Link_start := x_m,
// the newly initialization for link_start //
  endwhile,
 endwhile,
end

Procedure : 합성 알고리즘
// Switch setting for banyan^{-1} network at each stage //
begin
 While $0 \le i \le n - 1$ do
  con_bit$_i$ = $b_{(i+1)} \bmod n$, // To start from b_0 //
  If *con_bit*$_i$ == 0
   goto upper output ;
  Else
   goto lower output ;
 endwhile,
end

4. 성능분석

이 스위치는 각각의 입력에 독립적인 N개의 버퍼로 구성되며 그것은 블러킹을 고려하여 효과적인 서비스 타임 분배로 구성된다 각 포트의 버퍼는 타임 슬럿에 따라 유입되는 패킷을 FIFO의 기법에 따라 처리된다. 이 때 유입되는 패킷은 슬럿마다 각 포트에 대한 전송되는 트래픽, p의 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.

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

그러므로 랜얀 네트워크의 각 출력으로 도달될 수 있는 모든 패킷의 수는 다음 식으로 표현된다 ($\sum$는 $\sum_{i=0}^{N-1}$를 의미함)

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

이 스위치의 구조에서 패킷의 처리율은 정상 상태에서 유입되는 타임 슬럿마다 확률 p이므로 그것은 기대값으로 표현될 수 있으며, 그에 대한 식은 다음과 같다

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

또한, 한 타임 슬럿 동안 블러킹된 패킷의 수를 $RO_{blocked}$라 하고, 이전 타임 슬럿에 블러킹 되지않은 패킷의 수를 RO_i^{pre}라고 하면

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

을 얻는다. 위의 (4)식을 스위치 처리율의 관계식으로 유도하기 위하여, 양변에서 기대값을 구하고, 양변을 N으로 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 입출력 패킷의 보존 관계식을 이용한다. 우선, 이전에 타임 슬럿에 블러킹 되지않은 패킷의 확률을 정상 상태에서 α라고 가정한다.

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

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 p와 포화정도 α를 나타내는 식이다. 따라서 블러킹되지 않은 패킷의 수를 최대로 하기 위하여 식 (6)로부터 $\sum RO_i = N$은 상수이므로 α의 값을 정하면 된다. 스위치 처리율 p는 α가 "1"의 값을 가질 조건에서 최대의 처리율을 나타내고, 또한 정상 상태에서 입력 포트의 수는 최대 N 이상을 갖을 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 슬럿 동안 i번째 패킷이 임의의 버퍼에 도착할 확률을 χ_i라고 하면 다음과 같은 식으로 표현 가능하다

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

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

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

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

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

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, $N \to \infty$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 K라고 가정하고, 그 큐의 평균 길이를 구하기 위하여 다음과 같이 적용한다. 즉,

$$\sum_{k=0}^{\infty} p_k S^k = (1-\tilde{\omega}_p)(1-p+pS)/(1-S\tilde{\omega}_p)$$

$$\text{for } \tilde{\omega}_p = (p^2/2(1-p)^2$$

$$= (1-\tilde{\omega}_p)(1-p+pS)\sum_{k=0}^{\infty}\tilde{\omega}_p^k S^k$$

결과적으로, K의 평균 기대값이 다음과 같이 주어지고

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

위의 식으로부터 다음과 같은 결과를 얻는다.

$$p_0 = (1-\tilde{\omega}_p)(1-p)$$

$$p_k = (1-\tilde{\omega}_p)((1-p)\tilde{\omega}_p + p)\omega_p^{k-1}$$

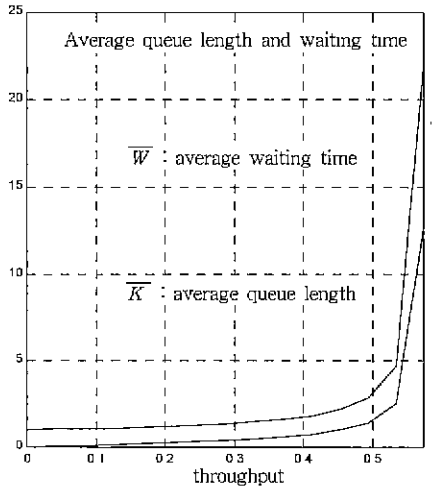
$$= \frac{p(2-p)}{2(1-p)}(1-\tilde{\omega}_p)\tilde{\omega}_p^{k-1}\ ;\ k>0$$

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

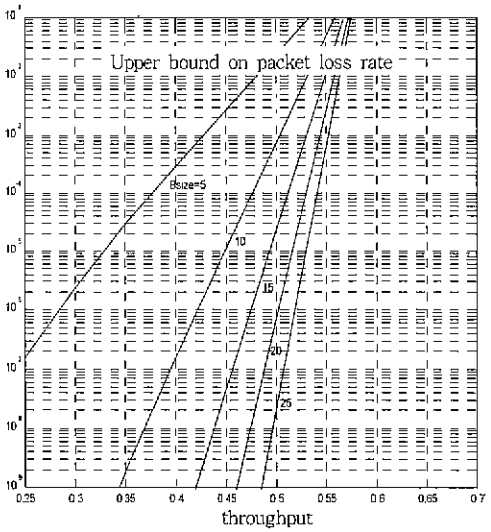
또한, 유한한 버퍼 크기(B_{size})에 대한 버퍼 초과 확률은 $\overline{K} > B_{size}$의 확률에 따라 패킷 손실율의 상위 한계점이 된다. 따라서,

$$\mathrm{P(loss)} < \mathrm{P}(\overline{K} > B_{size}) = \frac{p(2-p)}{2(1-p)}\tilde{\omega}_p^{B_{size}} \qquad (13)$$

(그림 4)는 식 (13)에 따라 패킷 손실율에 대한 상위 한계점을 나타낸다. 이것으로 스위치의 최대 부하률인 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 낮은것으로 나타난다. 또한 45% 정도면 버퍼 사이즈, $B_{size} = 10$이라도 무방함을 알 수 있다.



(그림 3) 처리율의 평균 큐의 길이 및 대기시간



(그림 4) 패킷 손실율과 버퍼의 크기

5. 결 론

본 논문에서는 shuffle-exchange, 순위 그리고 밴얀 네트워크로 구성된 패킷 스위치를 제안하였다. 우선, 하

나의 스테이지로 구성되어 하드웨어의 복잡도가 낮은 shuffle/Exchange 네트워크를 구성하고, 이를 통하여 얻어진 결과에 따라 트리 구조의 순위 네트워크에서는 서로 동일한 목적지 주소를 갖는 패킷들에 우선순위를 부여하기위한 알고리즘 적용하였다. 우선순위가 정해지면, 그 순위가 가장 높은 패킷을 제외한 나머지 모든 패킷은 재전송되기 위해 shuffle-exchange 네트워크로 재입력된다 우선순위가 가장 높은 패킷은 다음 단계인 밴얀 네트워크로 입력되어 서로 구별되는 목적지 주소를 갖게된다 이 때 역밴얀 자체 라우팅 네트워크가 되도록 밴얀 네트워크에서 스테이지마다 목적지 주소의 기준 비트를 정하여 두개의 데이터 집합 형태로 반복 분리함으로서 다음 스테이지의 정보 없이도 스위치의 연결이 가능하도록 알고리즘을 설계하였다. 또한 역밴얀 네트워크에서는 앞의 기준비트의 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, 즉 55%인 경우 대략 40이며 45%의 경우는 버퍼 사이즈가 10정도면 수용할 만한 패킷의 손실율을 갖는다.

참 고 문 헌

[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 Jan. 1999, pp.1-13.

[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, Feb 1992, pp.423-430

[5] J. N Giacopelli, W D. Sinkoskie, and M Littlewood, "Sunshine : A high performance self-routing broadband packet switch architecture," *in Proc. ISS'90*, 1990, pp 123-129.

[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, October 1987, pp.1264-1273.

[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, June 1997, pp. 710-722

[9] J. S. Turner, "Design of a Broadcast packet switching network," *in Proc. INFOCOM'86*, pp.667-675. Also, Dep. Comput, Sci., Washomgton 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 Nov 1986, pp. 2.1.1.-2.1.6.

박 병 수

e-mail : bpark@smuc.ac.kr

1986년 한양대학교 전자공학과 졸업

1989년 한양대학교 대학원 전자공학과 졸업

1994년 Texas A&M 대학교 Dept of Electrical Eng. 박사졸업

1994년~1995년 현대전자(주) 소프트웨어 연구소 선임연구원

1995년~현재 상명대학교 컴퓨터정보통신학부 조교수

관심분야 parallel interconnection network, parallel computer, computer network, ATM