

# ATM 멀티캐스트 스위치를 위한 3차원 논블록킹 복사망의 설계 및 성능평가

(Design and Performance Evaluation of a 3-Dimensional  
Nonblocking Copy Network for Multicast ATM Switches)

신재구<sup>\*</sup> 손유익<sup>\*\*</sup>  
(Jae-Gu Shin) (Yoo-Ek Son)

**요약** 본 논문은 멀티캐스트 ATM 스위치를 위한 새로운 복사망에 관하여 언급한다. Lee의 복사망 제안 이후 현재까지 많은 연구가 진행되어 왔으나 스위치 내에서의 오버플로우와 셀 충돌 문제는 여전히 문제점으로 제기되어 왔다. 본 논문에서는 이러한 문제점을 개선하기 위해 공유 버퍼를 갖는 3차원 멀티캐스트 스위치 구조를 제안함으로써 높은 부하에서도 다중경로와 다중출력을 통해 오버플로우와 충돌을 줄일 수 있는 방법에 관하여 언급한다.

이를 위해 큰 팬아웃의 셀을 처리하기 위한 셀 분할 알고리즘(Cell splitting algorithm)과 Lee의 Broadcast Banyan Network를 확장시킨 복사망을 제안하였다. 셀 복사는 셀의 두 비트 멀티캐스트 패턴과 부울 방식의 내부분할 알고리즘(Boolean interval splitting algorithm)을 사용하여 네트워크가 갖는 자가-라우팅의 특성에 따라 이루어지도록 하였다. 제안된 복사망에서는 Lee의 복사망에서 문제가 되었던 오버플로우 문제, 큰 팬아웃의 셀 처리 문제, 셀 충돌 문제 등을 개선시킴으로써 기존의 멀티캐스트 ATM 스위치 성능을 향상시키고자 하였다. 시뮬레이션에 의한 성능 평가 결과 기존의 방법과 비교하여 산출량, 셀 손실률, 셀 지연에 대해 좋은 성능을 보이고 있다.

**키워드** : 다단계상호연결 네트워크, 멀티캐스트, 복사망, ATM 스위치

**Abstract** This paper presents a new copy network for multicast ATM switches. Many studies have been carried out up to date since the proposition of Lee's copy network. However, the overflows and cell conflicts within the switch have still been raised a problem in argument. In order to reduce those problems, we proposed a 3-dimensional multicast switching architecture which has shared buffers in this paper. The proposed architecture can reduce the overflows and cell conflicts through multiple paths and output ports even in the high load environments. Also, we proposed a cell splitting algorithm which handles the cell in the case of large fan-out, and a copy network to increase throughput by expanding the Lee's Broadcast Banyan Network(BBN). Cell copy uses the Boolean interval splitting algorithm and the multicast pattern of the cells according to the self-routing characteristics of the network. In the proposed copy network, we improve the problems such as overflow, cell splitting of large fanout, cell conflicts, etc., which were still existed in the Lee's network.

The results of performance evaluation by computer simulation show that the proposed scheme has better throughput, cell loss rate and cell delay than the conventional method.

**Key words** : Multistage Interconnection Network, Multicast, Copy Network, ATM Switch

## 1. 서론

<sup>\*</sup> 정회원 : 계명대학교 컴퓨터공학과  
shin@tgedu.net

<sup>\*\*</sup> 종신회원 : 계명대학교 컴퓨터공학과 교수

yeson@kmu.ac.kr

논문접수 : 2001년 12월 17일

심사완료 : 2002년 8월 8일

광대역통신망(B-ISDN)에서 멀티캐스트 통신으로 특정 지워지는 주문형 비디오, 화상회의, 분산된 데이터 처리 등과 같은 서비스는 필수적 요구 사항으로 받아들여지고 있다. 이러한 기능을 제공하기 위해서는 광대역 통신망을 형성하는 주요 기능 중의 하나인 ATM 스위치에서 점대점(point-to-point) 연결뿐만 아니라 멀티캐

스트 스위칭 기능이 제공되어야 한다. 이것은 스위치에 입력된 셀을 복사할 수 있는 복사망이 필요하다는 의미이며, 지금까지 멀티캐스트 스위칭을 위한 ATM 스위치 복사망에 관해서는 Huang과 Knauer의 Starlite System[1], Turner의 Broadcast Packet Switch[2], Lee의 Non-blocking copy network[3] 등을 대표적인 구조로 하여 많은 연구가 진행되어 왔다[1-5].

현재까지 제안된 주요 구조는 대체로 Crossbar 네트워크, Banyan 네트워크 및 Clos 네트워크 기반의 복사망 등으로 분류할 수 있다[6]. Crossbar 네트워크 기반의 복사망은 내부 충돌은 발생하지 않지만 하드웨어 복잡도가 높고 출력 충돌 가능성이 있다. 이 부류의 대표적인 스위치로는 Gauss ASE, SCOQ switch, Clos-Knockout switch, Abacus switch 등이 있다.

Banyan 네트워크 기반의 복사망은 단단계 상호연결 네트워크(MINs)의 대표적인 구조로써 자가라우팅(self routing)과 단일경로(unique path) 특성으로 인해 내부 충돌이 발생할 수 있다. 이러한 계열의 대표적인 스위치로는 Starlite switch, Turner의 broadcast switch, Lee의 multicast switch, Sunshine switch, Recursive multistage structure, LGMIN switch, Multinet switch, Prescheduling multicast switch, Dilated network 등이 있다.

Clos 네트워크 기반의 복사망은 입력에서 출력으로 다중경로를 제공하므로 Banyan 네트워크에서 보다 내부 충돌이 적다. 제안된 대표적인 스위치는 Growable packet switch, Ring sandwich network, Clos-Knockout switch, Abacus switch 등이 있다.

현재까지 제안된 멀티캐스트 스위치들을 분석해 보면, 요구된 복사본의 수가 스위치의 크기보다 클 경우 발생하는 오버플로우 문제, 남아 있는 BBN의 포트 수보다 처리해야 할 셀의 복사본 수가 클 경우 셀이 전송되지 못하여 네트워크의 유용성을 떨어뜨리는 큰 팬아웃의 셀 처리 문제, 그리고 같은 셀이 동시에 같은 출력 포트를 요구할 경우 발생하는 셀 충돌 문제 등이 주요 과제가 되어 왔으며, 이들은 멀티캐스트 스위치의 성능에 중요한 영향을 주게 된다.

이를 해결하기 위한 기존 연구를 보면, Lee의 복사망의 특성을 기반으로 BBN의 구조 변경을 통해 문제점을 해결하고자 한 Turner의 N개 출력포트보다 크게 확장한 복사망[7], Tagle과 Sharma의 dilated banyan network[8], Alimuddin, Alnuweiri과 Donaldson의 Fat-Banyan(FAB) network[9] 등의 복사망이 있다. 그리고 요구된 복사본 수를 사전에 셀을 분할하여 조정

을 통해 BBN에서 복사과정이 이루어지게 하는 X. Lju, H. T. Mouftah의 복사망[10], 입력셀에 대해 우선 순위를 부여하여 셀 손실을 줄이려고 한 Tarnng, Meditch와 Somani의 복사망[11] 등이 있다.

Turner의 경우 큰 팬아웃을 가진 셀을 처리하기 위해 확장된 복사망을 사용했지만 N을 초과하는 복사본의 처리 문제와 과도한 메모리 사용의 문제가 남아 있었다. Sharma는 N을 초과하는 복사본을 처리하기 위한 그룹 전송에서 다중 네트워크 구조를 요구하게 되며, Donaldson의 경우는 링크 확장을 통해 셀 손실률은 줄일 수 있었지만 복사본의 합이 N을 초과하여 발생하는 오버플로우 문제는 여전히 남아 있었다. X. Lju, H. T. Mouftah의 경우는 셀 분할을 통해 그룹 전송이 가능하지만 동시 처리할 수 있는 복사본은 N을 초과할 수 없었으며, Tarnng, Meditch와 Somani는 단지 하위 입력에 도착한 셀의 입력 공정성을 해결을 통해 성능을 향상시키려고 했다.

따라서 본 논문에서는 기존 해결 방법에서 문제가 된 오버플로우, 큰 팬아웃의 셀 처리, 셀 충돌 등을 개선하기 위해 이진트리 네트워크 구조를 확장시킨 3차원의 확장된 복사망 구조를 제안한다. EBN 복사망은 다중입력 및 입력부와 출력부 사이에 다중경로와 다중출력 포트를 제공하고 스위치 내부의 SE에 버퍼를 사용함으로써 오버플로우, 충돌을 줄여 복사망의 성능을 높이고자 하였다. 또한 제안된 복사망에는 셀 분할 알고리즘을 적용하여 큰 팬아웃의 셀을 처리함으로써 망의 유용성을 높였으며, BBN에서 요구된 셀을 복사하기 위해서는 부울 방식 내부분할 알고리즘(Boolean interval splitting Algorithm)을 적용하였다.

## 2. 관련 연구

### 2.1 Lee의 복사망

Lee는 고정된 지연을 갖는 논블록킹 자가 라우팅 복사망을 제안하였다[3]. 이 복사망은 그림 1과 같이 Running adder network(RAN), Dummy address encoder(DAE), Broadcast banyan network(BBN), Trunk number translator(TNT)의 네 부분으로 구성되어 있다. RAN은 입력 패킷의 헤드에 명시된 복사본 수를 합하여 running sum을 생성하며, DAE는 RAN에서 생성된 running sum을 근거로 [최대값, 최소값]을 갖는 Dummy address interval을 생성하여 새로운 셀의 헤드를 형성한다. BBN은 셀을 복사할 수 있는 브로드캐스트 스위치 노드를 가진 반안망으로 구성되어 있으며 DAE에서 생성된 Dummy address interval를 근거로

부울 방식 내부분할 알고리즘에 기초하여 셀의 복사 과정을 수행한다. 그리고 TNT는 출력포트의 주소와 Index reference(IR)에 기초하여 실제 목적지 주소를 생성하여 셀을 전송한다.

셀의 복사는 인코딩과 디코딩 과정을 거치며 이루어진다. 인코딩은 입력된 셀의 헤드에 명시된 CN(copy number)의 집합을 새로운 패킷 헤드를 형성하는 Monotone address interval의 집합으로 변환시키는 과정으로 RAN과 DAE에서 이루어진다. 패킷이 RAN에 도착하게 되면, 패킷 헤더 속에 있는 CN을 순환적으로 합하여 복사본 합을 구한다. DAE에서는 Dummy address interval과 IR의 두 필드를 갖는 새로운 헤더를 구성하게 된다. Dummy address interval은 복사본 합의 값에 따라 최대값(MAX)과 최소값(MIN)의 두 이진수로 표현된다. IR은 Dummy address interval의 최소값이 되며 TNT에서 Copy index(CI)를 결정하는데 사용된다.

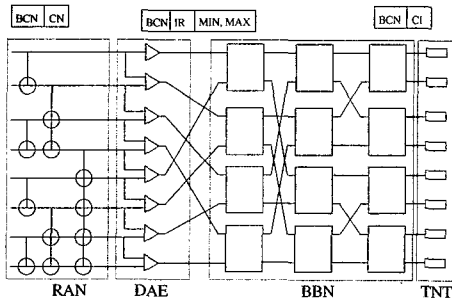


그림 1 Lee의 언블러킹 자가-라우팅 복사망

디코딩 과정은 BBN과 TNT에서 이루어진다. BBN에서는 인코딩 과정에서 생성된 Dummy address interval에 기초하여 부울 방식 내부분할 알고리즘에 의해 패킷을 복사한다. 그리고 BBN의 출력부에 패킷이 도착하면 TNT는 출력주소와 IR로부터 각 복사본의 CI를 계산하게 된다. BCN(Broadcast channel number)과 CI의 값에 따라 실제 목적지주소를 할당하고, 전송을 완료하게 된다.

**2.2 이진트리 베이스라인 네트워크(BTBN)**

이진트리 베이스라인 네트워크는 그림 2와 같이 계층화된 구조를 가지며, 상위층에서 하위층으로 갈수록 한 스테이지가 감소되는 형태로 구성되어 있다. 이진트리 형태의 입출력 경로는 기본 베이스라인 네트워크의 경로와 같으며 임의의 입력에서 임의의 출력으로 다중 경로 및 다중출력을 제공하게 된다[12, 13].

각 단계는 4개의 입력( $I_0, I_1, I_2, I_3$ )과 2개의 정상출력

( $f_0, f_1$ )과 2개의 여분출력( $r_0, r_1$ )을 갖는 4x4 스위치들로 구성된다. 각 SE에 도착한 셀은 정상적으로는 정상출력으로 라우팅 되지만, 정상출력에서 충돌이 발생하면 여분출력으로 라우팅 된다. 네트워크의 각 단계는  $[i, j]$  ( $0 \leq i \leq n-1, 0 \leq j \leq n-1$ )로 표현되고 단계마다  $N/2$ 개의 4x4 SE로 구성되며, 단계  $[i, j]$ 에서 각 스위칭 소자는  $SE[i, j, k]$ , ( $k=d_{n-2}d_{n-3}...d_0$ )로 표현된다. 단계  $[S_i, S_j]$ 와  $[T_i, T_j]$  사이의 상호연결 패턴은 베이스라인 네트워크에서 단계  $[S_i+S_j]$ 와  $[T_i+T_j]$ 의 상호연결 패턴과 같다 ( $|T_i-S_i| \leq 1, T_j=S_j$  또는  $T_i=S_i, |T_j-S_j| \leq 1$ ).

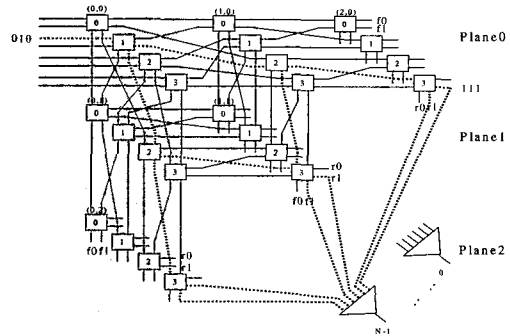


그림 2 이진트리 베이스라인 네트워크의 구조

이진트리 베이스라인 네트워크에서의 정상출력 및 여분출력 연결 패턴은,

정상출력 연결 :  $SE[i, 0, d_{n-1}d_{n-2}...d_1] \Rightarrow$

$$SE[i+1, 0, d_{n-1}d_{n-2}...d_{n-i}dd_{n-i+1}...d_2], 0 \leq i \leq n-2$$

여분출력 연결 :  $SE[i, 0, d_{n-1}d_{n-2}...d_1] \Rightarrow$

$$SE[i, 1, d_{n-1}d_{n-2}...d_{n-(i+j)}dd_{n-(i+j)-1}...d_2], 0 \leq i \leq n-2$$

와 같으며, 이때  $d$  값이 '1'이면  $f_1(r_1)$ 로 연결되고, '0'이면  $f_0(r_0)$ 로 연결된다. 같은 방법으로

정상출력 연결 :  $SE[i, j, d_{n-1}d_{n-2}...d_1] \Rightarrow$

$$SE[i, j+1, d_{n-1}d_{n-2}...d_{n-(i+j)}dd_{n-(i+j)-1}...d_2], 1 \leq i+j \leq n-2$$

여분출력 연결 :  $SE[i, j, d_{n-1}d_{n-2}...d_1] \Rightarrow$

$$SE[i+1, j, d_{n-1}d_{n-2}...d_{n-(i+j)}dd_{n-(i+j)-1}...d_2], 1 \leq i+j \leq n-2$$

가 된다. 이러한 이진트리 베이스라인 네트워크 구조는 트리 형태의 충돌 이루고 있으므로 셀을 출력부에 전송시키기 위해  $\log_2 N$ 개의 단계만을 거치게 되므로 전송이 빠르고, 또한  $N$ 개의 다중 경로에 의해 내부충돌을 줄일 수 있으며,  $2(\log_2 N)$ 개의 접근점(access point)을 가지므로 출력충돌을 줄일 수 있다. 이로 인해 산출량과 셀 손실, 셀 지연에 대해 높은 성능을 가지게 된다.

**2.3 셀 분할 알고리즘**

셀 분할은 한 번에 복사할 수 없는 패킷을 여러 번에

나누어 복사하는 것으로 망의 대역폭을 최대한 활용하게 하여 망의 성능을 높이기 위한 것이다. Lee의 복사망에서 DAE는 Dummy address interval([max,min])과 IR을 생성하여 새로운 패킷 헤더를 구성하게 된다. 이 경우 단순히 복사본 합에 기초하여 새로운 헤더를 구성하기 때문에 패킷을 복사하기 위한 BBN의 상황을 고려하지 않았다. 따라서 남은 출력포트보다 큰 팬아웃을 가진 패킷은 처리할 수 없게 되어 망의 성능을 떨어뜨리게 된다.

셀 분할 구조[10]는 입력 셀 분할 단계와 공유 메모리 버퍼 단계로 나뉘어 진다. 입력 셀 분할 단계는 가산기에서 생성된 복사본 합을 기초로 하여 셀 분할이 이루어지며, 이때 각 패킷의 최대 CN은 N이고, 최대 복사본의 합은  $N^2$ 이 된다. 그래서 가산기의 비트 수는  $L + \log_2 N$ 으로 확장된다. 상위 비트  $L$ 은 패킷의 DT (Departure time) 스케줄링을 위한 그룹으로 사용되며 하위 비트  $\log_2 N$ 은 Dummy address interval([max, min]) 생성을 위해 사용된다.

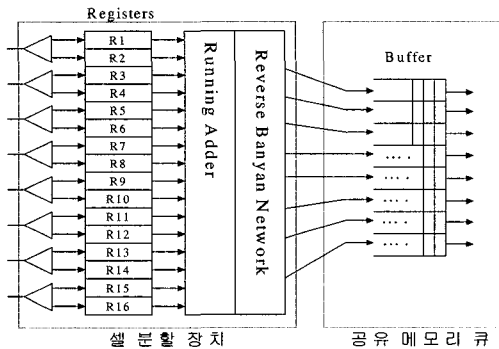


그림 3 셀 분할 구조

CN과 DT의 값에 따라 이루어지는 셀 분할 방법은,  $[DT_i = DT_{i-1}]$  경우, 셀 분할은 이루어지지 않으며 출력 셀의 헤더 구조는 아래와 같다.

$R_i$	BCN	$CN_i$	$DT_i$	Active ID(=1)
$R_{i+1}$	-	-	-	Inactive ID(=0)

$[DT_i > DT_{i-1}]$  경우, 셀 분할이 이루어지고 출력 셀의 헤더 구조는 아래와 같으며

$R_i$	BCN	$CN_{i1}$	$DT_{i1}$	Active ID(=1)
$R_{i+1}$	BCN	$CN_{i2}$	$DT_{i2}$	Active ID(=1)

CN과 DT는 (1)식과 같이 나타난다.

$$CN_{i1} = N - CN_i, \quad CN_{i2} = CN_i - CN_{i1},$$

$$DT_{i1} = DT_{i-1}, \quad DT_{i2} = DT_i \quad (1)$$

$[DT_i < DT_{i-1}]$  경우,  $S_i$ 는 오버플로우 발생으로 셀 손실이 일어난다. 그리고 각 셀의 Dummy address interval의 최대값, 최소값은 (2)식과 같다.

$$min_i = CN_{i-1}, \quad max_i = CN_i - 1 \quad (2)$$

공유 메모리 버퍼링 단계는  $2N \times 2N$  가산기, 역 반안 네트워크와  $N$  인터리빙 버퍼로 구성되어 있으며, 모든 활성 패킷을 선택하여 주소별로 연속적이고 순환적으로 정렬하여  $N$ 개의 공유 메모리 버퍼에 저장하게 된다. 저장된 패킷은 출발 시간 스케줄링 알고리즘에 따라 BBN으로 보내져 복사하게 된다.

### 3. 제안된 3차원 복사망

#### 3.1 시스템 구성

EBN 복사망은 기본적으로 Lee의 복사망에 기초하고 있다. Lee의 복사망은 요구된 복사본 합의 값이 N을 초과할 경우 오버플로우가 발생하여 해당 패킷은 폐기된다. 그리고 입력에서 다음의 처리해야할 패킷의 팬아웃 값이 남아 있는 BBN의 출력포트보다 클 경우 패킷은 복사될 수 없게 된다.

제안된 복사망은 큰 팬아웃 처리를 위해 DAE 부분에 셀분할 알고리즘을 적용하여 해결하고, 오버플로우 문제는 다중경로와 다중출력 및 내부버퍼를 가진 EBN을 사용하여 해결한다. 제안된 복사망의 구조는 그림 4와 같이 4개의 구성요소로 되어 있다.

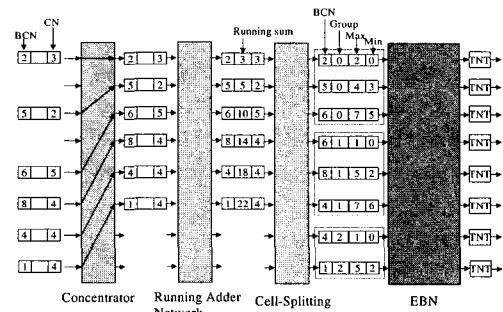


그림 4 제안된 EBN 복사망 구조

Concentrator는 입력된 패킷이 RAN 입력부에 연속적으로 놓이게 정렬을 해주며, RAN은 입력된 패킷의 CN필드 값에서 복사본 합을 구한다. 이 단계에서 구해진 복사본 합을 기초하여 셀 분할 단계에서는 BBN의

크기에 따라 셀 분할이 이루어지고 Dummy address interval([max,min])을 할당한다. EBN에서는 셀 헤더의 정보에 따라 부울 방식 내부분할 알고리즘에 의해 셀 복사가 이루어지며, 복사된 셀은 TNT에서 실제 주소를 할당받아 멀티캐스트 전송을 완료하게 된다.

**3.2 확장된 네트워크 구조(EBN)**

BBN에 적용시키기 위해 3차원의 EBN 구조를 제안하였다. 제안된 네트워크에서 멀티캐스트의 특성을 위한 오버플로우와 셀 충돌 문제를 개선하고, 산출량과 셀 손실에 대한 성능을 높이기 위해 확장된 입력부를 제공한다. 제안된 구조는 두 가지의 설계 개념을 갖는다.

첫 번째 개념은 입력부에서 발생하는 오버플로우를 줄이고 그룹별 전송을 하기 위한 네트워크 입력의 확장이다. 제안된 EBN은 BTBN에 4x4 SE의 k개 네트워크 단계를 추가하였으며, 입력부에서는 디멀티플렉서를 통해 입력 셀을 각 층으로 분산시키므로 BTBN을 사용할 경우 동시 최대수 N개의 복사본을 처리할 수 있지만, EBN은 동시에 (k+1)N개의 복사본을 처리할 수 있고 각 입력과 출력 사이에 kN개의 다중경로를 가지게 된다.

두 번째 적용된 개념은 버퍼 정책으로써 네트워크에서의 내부 충돌을 방지하여 셀 손실을 줄임으로써 성능을 향상시키기 위한 것이다. 버퍼 정책은 SE 내부에서 버퍼의 위치에 따라 입력 버퍼(input buffer), 출력 버퍼(output buffer), 공유 버퍼(shared buffer) 등으로 나눌 수 있다. 입력 버퍼는 버퍼가 SE의 입력부에 위치한 것으로 두 개 이상의 패킷이 SE의 같은 출력포트를 두고 경쟁할 때 한 개의 패킷만 전송되고 다른 패킷은 저장된다. 구현하기는 쉽지만 입력버퍼에 저장된 선두 셀이 블록킹 되어 같은 버퍼 내에 대기하는 다음 셀이 블록킹이 없음에도 불구하고 출력되지 못하는 HOL(Head-of-Line) 블록킹이 발생하여 성능향상에는 제한된다. 출력 버퍼는 버퍼가 출력부에 위치한 것으로 처리량과 지연에 대해 좋은 성능을 보이지만 주어진 클럭 내에 k개의 입력이 같은 출력 버퍼를 요구할 경우 다중 포트 버퍼와 같은 추가적인 기능을 요구하게 된다. 공유 버퍼는 입력과 출력이 같은 버퍼를 공유하여 사용하는 것으로 다른 출력으로 가는 패킷이 같은 버퍼에 저장되며, 다른 라우팅 주소를 가지고 공유 버퍼에 대기할 경우 라우팅 비트를 두 번 이상 검사하는 현상이 발생하게 된다.

일반적인 공유 버퍼 구조의 스위치에서는 저장된 셀을 전송하고자 할 때마다 라우팅 비트를 조사하게 되나, 그림 5와 같이 두 개의 공유버퍼를 사용하면 스위치의

각 입력은 라우팅 태그에 따라 두 개의 버퍼에 저장되고 출력시 라우팅 태그의 조사 없이 출력포트로 전송된다.

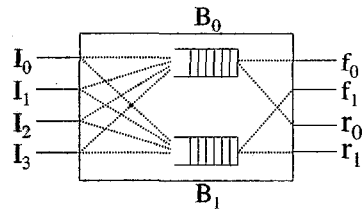


그림 5 공유 버퍼를 가지는 4x4 스위치

각 SE에는 두 개의 버퍼를 가지게 되며, 버퍼 B<sub>0</sub>는 f<sub>0</sub> 또는 r<sub>0</sub> 출력포트로 연결되고, 버퍼 B<sub>1</sub>은 f<sub>1</sub> 또는 r<sub>1</sub> 출력포트와 연결되어 패킷을 전송하게 된다. 패킷이 입력부에 도착하면 라우팅 비트를 조사하여 정상출력(f<sub>0</sub> 또는 f<sub>1</sub>)으로 전송하게 된다. 만약 두 개의 패킷이 같은 라우팅 비트를 가지고 들어오면 한 개는 정상출력으로 다른 한 개는 여분출력으로 전송되지만, 두 개 이상의 패킷이 들어오게 되면 두 개는 위와 같은 방법으로 전송되고 나머지는 라우팅 비트의 값에 따라 버퍼에 저장되며, 저장된 패킷은 다음 주기에 라우팅 비트의 검사 없이 지정된 출력포트로 전송되게 된다. 각 SE에서 처리되는 셀 처리 알고리즘은 그림 6과 같으며, 각 입력에

```

While(1)
for( i=0 ; i<4 ; i++ )
    If[i]=Packet[Stage_NOI] '해당 스테이지의 라우팅 비트를
                           '조사하여
if Packet.eof then '배열에 저장.
    return
else
    if ( B0 or B1 is full ) then '버퍼가 full인 경우 입력 패킷은
                                '폐기되고
                                '각 버퍼속의 두 패킷은 출력경로로
                                '전송.
        {
            f0=B0[0] ; f1=B1[0] ; r0=B0[1] ; r1=B1[1]
        }
    else
    {
        for ( i=0 ; i<4 ; i++ ) '버퍼가 full이 아닌 경우 라우팅
                                '비트를
                                '조사하여 해당 버퍼에 저장.
            {if ( If[i]=0 ) then
                B0=If[i]
            else
                B1=If[i]
            end if
        }
        f0=B0[0] ; f1=B1[0] ; r0=B0[1] ; r1=B1[1] '버퍼 속에 저장된
                                                '패킷은
                                                '출력 경로에 따라
                                                '전송.
    }
end if
do
    
```

그림 6 SE에서 셀 처리 알고리즘

걸리는 부하를  $\lambda$ 라 할 때 네트워크에서의 SE별 트래픽 부하[12]는

$$stage[0, 0],[i, 0] : \lambda_{T0}=\lambda_{T1}=\lambda, \lambda_{T2}=\lambda_{T3}=0 \quad (3)$$

이므로 BTBN의  $SE(i, j, k)$ 에서  $i=0$ 이거나  $j=0$ 인 SE에서는 충돌이 발생하지 않기 때문에 BTBN에서는  $i \neq 0$ 이고  $j \neq 0$ 인  $SE(i, j, k)$ 에서만 공유 버퍼를 가진 SE로 구성된다.

### 3.3 EBN 복사망에서 셀 복사

RAN에서 계산된 CN의 값에 따라 셀 분할 단계에서는 EBN 복사망을 통해 전달될 셀을 복사하기 위해 새로운 헤더를 구성하게 된다.  $L + \log_2 N$ 으로 확장된 라우팅 비트는 L의 전송 그룹 필드(DT)와  $\log_2 N$ 의 주소 필드로 나누어진다. 주소 필드 값은 Lee의 설계와 같이 셀을 복사하기 위한 Dummy address interval([max, min])을 생성하게 되고 그룹별 전송을 위한 그룹 필드가 추가되어 셀 헤더를 구성하게 된다. 이때 셀은 CN과 DT의 값에 따라 그림 7과 같이 분할되어 각 레지스터로 입력된다. 그리고 공유 메모리 큐에 저장된 셀은 다음 단계인 EBN으로 입력된다.

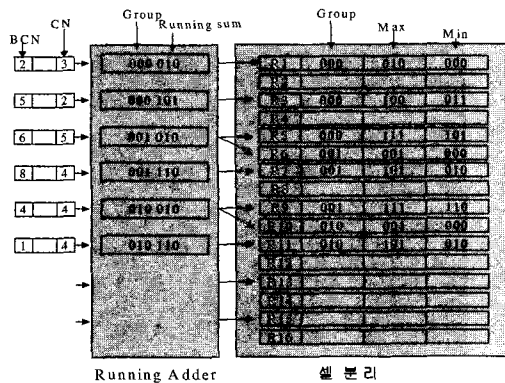


그림 7 셀 분할

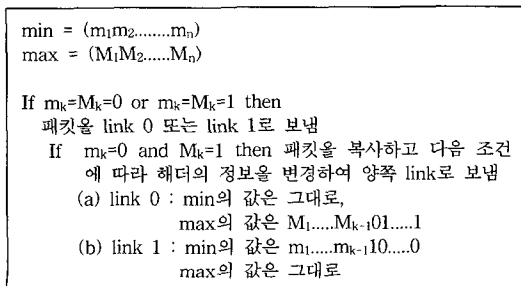


그림 8 부울 방식 내부분할 알고리즘

EBN에 입력된 셀은 그룹 필드의 값에 따라 EBN 복사망의 입력부에서 입력 경로가 결정된다. 경로가 결정된 그룹은 주소 필드의 값에 기초하여 그림 8과 같은 부울 방식 내부분할 알고리즘에 따라 그림 9와 같이 셀을 복사하게 된다.

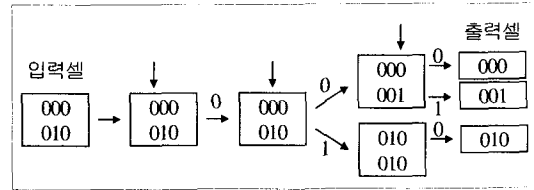


그림 9 부울 방식 내부분할 알고리즘에 따른 셀 복사

그림 10은 EBN에서 그룹별 셀이 전송되는 예를 보여준다. 그룹 필드의 값에 따라 000, 001, 010의 세 개의 그룹으로 나뉘어져 복사망에 입력되며 입력된 셀들은 주소 필드의 최대값[010], 최소값[000]에 기초하여 셀의 복사가 이루어진다. 입력부의 크기를 초과한 그룹은 다음 주기에 전송이 이루어진다.

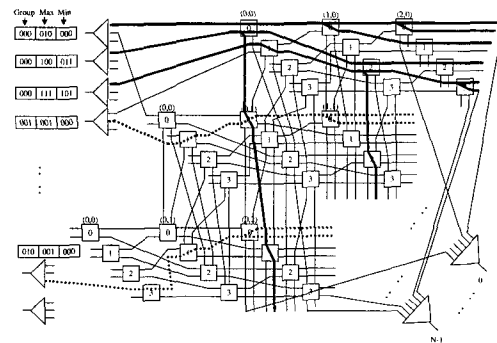


그림 10 그룹별 셀 전송 예

제안된 복사망에서는 셀의 복사 기능을 수행하는 복사망에 EBN을 적용하였다. EBN은 베이스라인 네트워크를 기초하고 자가-라우팅의 특성을 가지므로 이러한 구조적 특성에 따라 셀의 헤더의 두 비트 정보에 따라 셀을 복사하는 부울 방식 내부분할 알고리즘을 적용하였다. EBN은 다중입력 및 입력부에서 출력부에 이르는  $kN$ 개의 다중경로 제공과 각 SE에 공유 버퍼를 가지므로 오버플로우 문제와 내부 충돌 발생을 해결할 수 있다.

## 4. 시뮬레이션 및 성능평가

4.1 시뮬레이션 환경 및 파라미터

여기서는 제안된 복사망의 성능을 평가하기 위해 기본 베이스라인 네트워크 및 2-dilated 반안 네트워크를 기반으로 하는 복사망과 비교, 분석한다.

먼저 EBN의 SE에서 입력 부하에 따른 출력 부하의 산출량을 분석해보고, 시뮬레이션을 통해 EBN 복사망의 성능을 평가한다. 시뮬레이션은 N=8과 N=16인 네트워크에서 다양한 트래픽 부하(0.2, 0.5, 1.0, 1.5, 2.0, 2.5, 3.0)를 적용하였으며, 또한 제안된 복사망에서 버퍼의 수에 따라 부하별 산출량도 비교하였다. 제안된 복사망의 성능평가를 위해 프로세스 중심(process-oriented) 방식의 이산적인 모델링을 사용하며, 처리량(throughput)과 셀 손실(cell loss), 그리고 셀 지연(cell delay)을 평가 기준으로 한다. 처리량은

$$Throughput, T = \frac{total\ output\ cells}{total\ input\ cells} (\%) \quad (4)$$

으로서 매 주기당 출력포트에 네트워크에 의해 전달된 셀의 평균수로 정의하며, 셀 손실율은

$$Celllossprobability, P_{[cell\ loss]} = \frac{total\ deleted\ cells}{total\ input\ cells} \quad (5)$$

으로 네트워크를 통해 출력포트에 전달되지 못한 셀의 비율로 정의한다. 그리고 셀 지연은

Average waiting time in a buffer,

$$D = \frac{\sum_{i=0}^{n-1} \sum_{j=0}^{N-1} buf_{ij}}{N \times \log_2 N} (\mu s) \quad (6)$$

으로서 각 스위치 내부의 버퍼  $buf_{ij}$ 에 대한 각 셀들의 평균 버퍼 지연으로 정의한다. 즉,  $i$ 는 스테이지 번호,  $j$ 는 입력 포트 수이며, 전체 셀들의 평균 버퍼 지연의 합을 전체 스위치 소자의 버퍼 수로 나눈 값이 된다.

4.2 EBN의 성능

EBN의 SE구조에 따른 산출량을 분석하기 위하여 각 SE의 4개의 입력( $I_0, I_1, I_2, I_3$ )에 걸리는 평균 부하를  $\lambda_{I_0}, \lambda_{I_1}, \lambda_{I_2}, \lambda_{I_3}$ 로 표현하고, 정상출력  $f_0, f_1$ 과 여분출력  $r_0, r_1$ 에 걸리는 평균 부하를  $\lambda_{r_0}, \lambda_{r_1}$ 로 표현한다. 각 SE의 출력부하는 입력부하에 따라 유지되며, 4개의 입력이 독립적이라 가정할 때 입력부하에 따른 출력 부하[13]는

$$\lambda_{r_0} = \lambda_{r_1} = 1 - (1 - \lambda_{I_0}/2)(1 - \lambda_{I_1}/2)(1 - \lambda_{I_2}/2)(1 - \lambda_{I_3}/2) \quad (7)$$

$$\lambda_{r_0} = \lambda_{r_1} = \lambda_{f_0} - (\lambda_{r_0}/2)(1 - \lambda_{I_1}/2)(1 - \lambda_{I_2}/2)(1 - \lambda_{I_3}/2) - (1 - \lambda_{r_0}/2)(\lambda_{r_1}/2)(1 - \lambda_{I_2}/2)(1 - \lambda_{I_3}/2) - (1 - \lambda_{r_0}/2)(1 - \lambda_{I_1}/2)(\lambda_{r_2}/2)(1 - \lambda_{I_3}/2) - (1 - \lambda_{r_0}/2)(1 - \lambda_{I_1}/2)(1 - \lambda_{I_2}/2)(\lambda_{r_3}/2) \quad (8)$$

와 같다. 입력 부하에 따른 출력부하는 표 1과 같다.

표 1 EBN의 SE에서 입력 부하에 따른 출력 부하

입력부하	0.2	0.4	0.6	0.8	1.0
정상출력( $\lambda_{r_0}, \lambda_{r_1}$ )	.3439	.5904	.7599	.8704	.9375
여분출력( $\lambda_{r_0}, \lambda_{r_1}$ )	.0523	.1808	.3483	.5248	.6875

또한, SE의 단계별로 걸리는 트래픽 부하를 보면 식(1)과 같다. 왜냐하면, 이 단계의 SE에서  $I_2, I_3$ 는 사용하지 않기 때문이다. 이상에서 알 수 있듯이 단계[0, 0], [i, 0]내의 SE에서는 패킷 손실이 일어나지 않는다. 그 이유는 항상 다음 단계로 전송되기 위한 입력 패킷의 경로가 충분하기 때문이다. 따라서 EBN의 SE에서는 완전부하( $\lambda=1.0$ )가 주어져도 높은 산출량을 유지할 수 있다.

그러나 식(9)와 식(10)에서 보는 바와 같이 하드웨어 복잡도 면에서는 제안된 구조가 높게 나타난다. 참고문헌 [7-11]의 제안된 구조에서 보듯이 성능향상을 위해 다르게 되는 복잡도는 VLSI 기술로 해결될 수 있을 것으로 여겨진다.

$$SE_{Baseline=2-dilated} = \frac{N}{2} \cdot n$$

$$SE_{B-tree} = N \cdot n \quad (9)$$

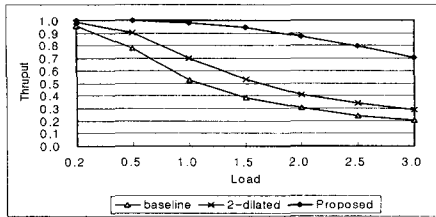
$$Link_{Baseline} = 2N(n-1)$$

$$Link_{2-dilated} = 2Nn$$

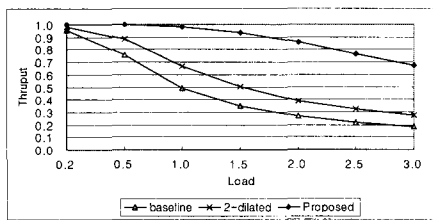
$$Link_{B-tree} = nN(n+1)+N \quad (10)$$

4.3 시뮬레이션 결과 분석

그림 11은 N=8, 16인 복사망에서 부하에 따른 산출량을 보여 주고 있다. 베이스라인 네트워크를 기반의 복사망과 2-dilated 반안 네트워크 기반의 복사망은 낮은 부하에서는 비슷한 산출량을 보이지만 부하가 높을수록 산출량이 떨어지는 것을 볼 수 있다. 특히, 부하가 1.0이상에서는 산출량의 감소가 더 커진다. 반면 제안된 복사망은 부하 0.8까지는 거의 완전한 산출량을 보이며 부하 3.0에서도 높은 산출량(a:0.71, b:0.67)을 보여 주고 있다. 따라서 DAE에서 그룹화된 셀 중 동시에 3개의 그룹(3N개, k=2에서)이 입력되어도 높은 산출량을 보이게 된다. 이는 EBN 복사망에서 입력 확장을 통해 동시에 입력받을 수 있는 셀의 그룹이 확장되었으며, 이로 인해 kN개의 다중 경로가 생기고,  $2(\log_2 N)$ 개의 접근점을 가지며, 내부 SE에 버퍼를 가지므로 다른 복사망에 비해 높은 산출량을 보이고 있다.



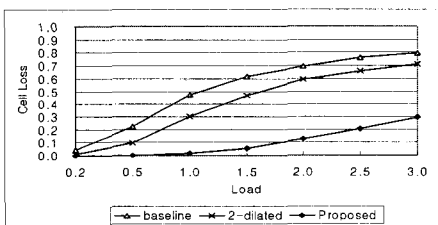
(a) N=8



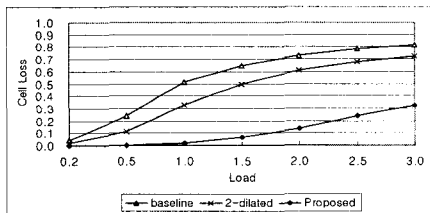
(b) N=16

그림 11 부하에 따른 산출량

그림 12는  $N=8, 16$ 인 복사망에서 부하에 따른 셀 손실율을 보여 주고 있다. 제안된 복사망은 비교 대상의 복사망과 큰 차이를 보여 주고 있다. 제안된 복사망은 부하가 높아져도 낮은 셀 손실율을 보이며, 1.0이상의 높은 부하에서는 약간의 손실을 보이고 있다. 반면 베이스라인 네트워크 기반의 복사망과 2-dilated 반안 네트워크 기반의 복사망은 부하가 높을수록 높은 셀 손실율을 보이고 있다.



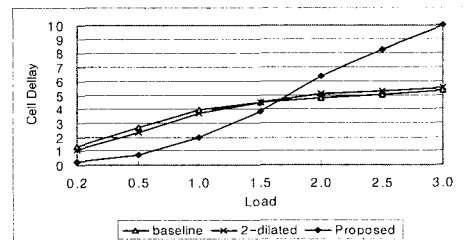
(a) N=8



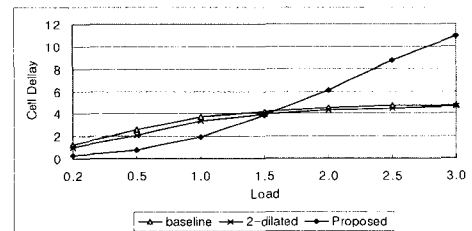
(b) N=16

그림 12 부하에 따른 셀 손실률

셀 지연에서도 제안된 복사망이 우수한 성능을 보이고 있다[그림 13]. 제안된 복사망이 다중 경로와 다중 출력을 제공함으로써 빠른 전송을 통해 셀 지연을 줄였기 때문이다. 그러나 제안된 복사망이 다른 복사망에 비해 우수한 지연 성능을 보이고 있지만 부하 1.0에서부터 셀 지연이 급속도로 증가하여 부하 1.8이상에서는 다른 복사망보다 높은 셀 지연을 보이고 있다. 이는 높은 부하로 인해 버퍼에서 지연되는 셀이 증가하기 때문이다. 베이스라인 네트워크 기반의 복사망과 2-dilated 반안 네트워크 기반의 복사망은 대체로 높은 셀 지연을 보이며, 부하가 증가함에 따라 선형적으로 지연이 높아지고 있다.



(a) N=8

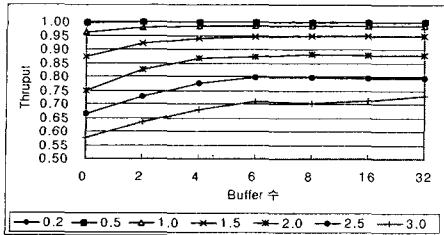


(b) N=16

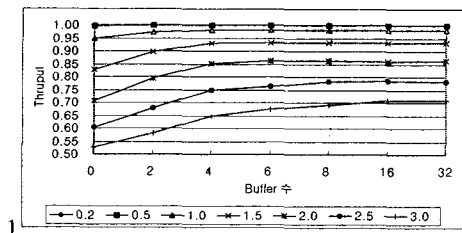
그림 13 부하에 따른 셀 지연

그림 14는 제안된 복사망에서 내부 SE의 버퍼 크기에 따른 부하별 산출량을 보여 주고 있다. 낮은 부하에서는 버퍼의 크기가 산출량에 큰 영향을 주지 못하지만 높은 부하에서는 (a)인 경우 26%이상, (b)인 경우 34% 이상의 산출량 증대 효과를 보여 주고 있다. 앞의 두 경우에서 볼 수 있듯이 (a)보다 (b)가 조금 낮은 성능을 보이는 것은 네트워크 크기가 커짐에 따라 셀의 충돌 발생률과 SE에서 지연이 높아지기 때문이다. 따라서 버퍼의 크기를 크게 할 경우 충돌에 의한 셀 손실율을 줄여 줌으로 (a)보다 (b) 경우의 산출량의 증가 비율이 높아지게 된다. 그리고 (a), (b) 경우 모두 버퍼 크기가 6가지는 버퍼 수가 커짐에 따라 산출량의 상승이 크지만 그 이상에서는 버퍼가 커져도 산출량은 크게 늘지 않는다.





(a) N=8



(b) N=16

그림 14 부하별 버퍼 크기에 따른 산출량

## 5. 결론

ATM 멀티캐스트 스위치 설계 시 고려해야 할 사항으로는 스위치 구조, 복사 알고리즘, 내부 및 출력 충돌 문제, 멀티캐스트 패턴 등이 있으며, 이러한 요소들의 특성이 충분히 검토·고려되어야 한다.

이를 위해 본 논문에서는 Lee의 복사망의 특성을 기반으로 Lee의 복사망과 기존의 여러 복사망에서 문제가 되었던 오버플로우 문제, 큰 팬아웃의 셀 처리 문제, 셀 충돌 문제 등을 검토하고 이를 위한 멀티캐스트 스위치를 제안하였다. 세부적으로는 실제 셀의 복사가 이루어지는 네트워크에서 높은 부하에서도 셀 손실과 충돌을 줄이기 위해 기존의 이진트리 네트워크를 확장시킨 확장된 스위치 구조를 제안하였다. EBN은 베이스라인 네트워크가 트리 형태의 층으로 구성되며, 또한 자가-라우팅의 특성과  $\log_2 N$ 개의 단계만에 목적지로 셀을 전송하는 특성을 가지고 있다. 이러한 구조적 특성을 이용하여 셀 헤드에 2비트( $\lceil \max, \min \rceil$ ) 정보를 사용한 멀티캐스트 패턴과, 부울 방식 내부분할 알고리즘에 의한 셀 복사 방식과, 그리고 큰 팬아웃의 셀을 처리하기 위한 셀 분할 알고리즘이 제안되었다.

제안된 복사망은 기존의 복사망에서 문제가 되었던 오버플로우 문제, 큰 팬아웃의 셀 처리 문제, 충돌 문제 등을 셀분할 알고리즘과 확장된 네트워크를 사용함으로써 멀티캐스트 ATM 스위치의 성능을 향상 시켰다. 성능 평가 결과 베이스라인 네트워크 기반의 복사망과

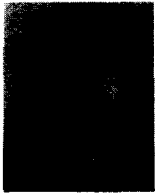
2-dilated 반안 네트워크 기반의 복사망과 비교하여 산출량, 셀 손실률, 셀 지연에 대해 높은 성능을 보였다.

향후 연구과제로서는 멀티캐스트 셀과 유니캐스트 셀이 같은 복사망을 통과함으로써 발생하는 문제와 복사망에서 복사된 셀이 라우팅 망에서 손실이 발생하지 않고 최종 목적지까지 전송될 수 있는 방법에 대한 연구가 필요한 것으로 사료된다.

## 참고 문헌

- [1] A. Huang and S. Knauer, "Starlite : A wideband digital switch," in Proc. IEEE Globecom'84, pp.121-125, 1984.
- [2] J. S. Turner, "Design of a Broadcast Packet Switching Network," IEEE Trans. Comm., pp.734-743, June 1988.
- [3] Tony T. LEE, "Nonblocking Copy Networks for Multicast Packet Switching," IEEE Journal, Selected Areas in Communications. Vol.6, No.9, p.1455-1467, December 1988.
- [4] Chen, D.X.; Mark, J.W, "Multicasting in the SCOQ switch," INFOCOM '94. Networking for Global Communications., 13th Proceedings IEEE, pp.290-297. vol.1, 1994.
- [5] Julian D. Ho, Samar Sing, Neeraj K. Sharma, "Modeling multicasting in shared-memory ATM switches," Communications, 2000. ICC 2000. 2000 IEEE International Conference on, Vol.3, pp.1728-1732, 2000.
- [6] M.H. Guo, R.S. Chang, "Design issues for multicast ATM switches," computer communications 22, pp.771-777, 1999.
- [7] J. S. Turner, "A Practical Version of Lee's Multicast Switch Architecture," IEEE Trans. on Comm., Vol.41, No.8, pp.1166-1169, August 1993.
- [8] Pierre U. Tagle, Neeraj K. Sharma, "Multicast Packet Switch based on Dilated Network," IEEE Global Telecommunications Conf. Vol.2, pp.849-853, November 1996.
- [9] M. Alimuddin, H. M. Alnuweiri and R. W. Donaldson, "Efficient Multicast Copy Network," Broadband Switching Systems Proceedings, 1997. IEEE BSS '97, 2nd IEEE Int'l Workshop, pp.169-172, 1997.
- [10] Xinyi Lju and H. T. Mouftah, "A Dynamic Cell-Splitting Copy Network Design for ATM Multicast Switching," Global Telecommunications Conf., 1994. GLOBECOM '94. Comm.: The Global Bridge, IEEE, pp.458-462 Vol.1, 1994.
- [11] Tanrng, C. L. Meditch, J. S. Somani, A. K, "Fairness and priority implementation in non-blocking copy network," Communication, ICC '91,

- Conf. Record. IEEE International Conference on, Vol.2, pp.1002-1006, 1991.
- [12] Chris Plate, Jack Tan, "Performance Analysis of a Fault-Tolerant B-Tree ATM Switch," IEEE 22nd Int'l Conference on Local Computer Networks, pp.295-304, October 1996.
- [13] J. J. Li, C. M. Weng, "B-tree : A high-performance fault-tolerant ATM switch," IEEE Proc. Comm., Vol.141, No.1, pp.20-28, February 1994.



신재구

1995년 계명대학교 교육대학원 전산교육 석사. 현재 계명대학교 대학원 컴퓨터공학과 박사과정. 관심분야는 MIN, ATM 스위치, 멀티캐스트, 복사망



손유익

1976년 경북대학교 전자공학과 학사  
 1979년 경북대학교 대학원 전자공학과 석사. 1990년 경북대학교 대학원 전자공학과(컴퓨터공학전공), 공학박사. 1979~1984년 한국전자기술연구소 컴퓨터연구부 선임연구원. 1984~현재 계명대학교 컴퓨터공학과 교수. 1994~1995년 한국정보과학회 영남지부장. 1999~2001년 컴퓨터·전자공학부 학부장. 2002~현재 계명대학교 전산원장. 관심분야는 병렬 알고리즘 및 구조 MIN, ATM 스위칭 네트워크 등