

論文97-34S-8-6

# ATM 스위치에서의 여분 경로 전송 메커니즘

## (Alternate Path Transfer Mechanism on ATM Switch)

李周榮\*, 林寅七\*

(Joo-Young Lee and In-Chil Lim)

### 요 약

고속 패킷 스위칭에 유리한 반얀망(Banyan network)기반의 ATM 스위치 설계를 위해선, 스위치 평면 내부의 충돌로 인한 패킷 손실을 줄이기 위한 입, 출력 포트간 다중 경로 설정과 이들 경로 상으로 전송될 패킷들을 제어하기 위한 적절한 패킷 전송 기법의 적용이 필수적이다. 본 논문은 병렬 반얀망에서의 충돌 문제를 손쉽게 해결할 수 있는 새로운 패킷 전송 기법, 즉 다이너믹 바이패스(Dynamic Bypass) 전송 기법을 이용한 여분 경로 전송 메커니즘을 제안한다. 제안한 ATM 스위치는 반얀망으로 구성된 다단의 스위치 평면과 입, 출력 포트, 바이패스 링크, 바이패스 링크 컨트롤러로 구성한다. SE(Switching Element)내에서 충돌을 유발하는 패킷들은 제안된 메커니즘을 통해 여분 스위치 평면상으로 패킷 재전송의 기회를 갖는다.

### Abstract

To design a ATM Switch which has advantages in high speed packet switching, it is essential to set multiple paths between input ports and output ports and to design a new packet transfer technique on that paths for decreasing Packet Loss by conflicts in internal Switch Plane. We propose new packet transfer method, Alternate Path Transfer Mechanism by Dynamic Bypass Transfer Method which can solve Conflict problem in Banyan network easily. Proposed ATM Switch consists of Banyan networks, Input/Output Port, Bypass Link, and Bypass Link Controller. Packets caused conflicts in SEs have another chances of packet transfer over alternate switching planes by using this mechanism.

### I. 서 론

B-ISDN(Broadband Integrated Services Digital Networks)은 광섬유 기술 및 고속 VLSI 기술의 발전을 바탕으로 등장한 차세대 통신망으로써, 음성, 데이터, 비디오 서비스 등 다양한 종류 데이터들을 전송률이 높은 광섬유 등을 통해 전송하는 통신을 의미한다.<sup>[1]</sup> 이러한 다양한 특성의 데이터 전송을 위하여 B-ISDN은 기존의 STM(Synchronous Transfer Mode)이 아닌 ATM(Asynchronous Transfer Mode) 전송 방식을 채택하고 있다. 기존의 STM

방식이 트래픽 소스들을 일정한 시간 채널을 할당하여 트래픽 발생 여부와 무관하게 반복적으로 패킷을 전송 하던 것에 반해, ATM 방식은 트래픽 소스로부터 패킷이 발생할 시, 트래픽 특성에 맞는 적절한 크기의 시간 채널을 할당한 후 패킷을 전송하는 방식으로써 서로 다른 특성을 갖는 트래픽이 혼재된 B-ISDN에 적합한 전송 방법이다.

광섬유 기술을 기반으로 하는 ATM 스위치는 스위칭 해야 할 패킷 양의 증가와 서비스 종류의 다양화, 실질적 외부 링크에 해당하는 물리 매체 속도의 증가로 말미암아 고속 및 다량으로 유입되는 패킷들을 전체의 정보 전송 품질을 저하시키지 않는 한도 내의 손실만으로 원하는 다음 목적지로 고속 스위칭을 해야만 한다. 이러한 요구 조건들을 충족해야 하는 ATM 스위치의 종류로는 그 기준의 설정 조건에 따라 다양하

\* 正會員, 漢陽大學校 電子工學科

(Dept. of Elec. Eng., Hanyang Univ.)

接受日子: 1996年11月26日, 수정완료일: 1997年7月23日

게 나눌 수 있으나,<sup>[1] [2]</sup> 전송 방식을 기준으로 분류 시 크게 RAM과 같은 메모리 소자를 이용한 시간 분할(Time Division)방식 스위치와 스위치 내부의 하드웨어적 병렬화를 통해 전 입력단의 패킷들을 동시에 전송하는 공간 분할(Space Division)방식 스위치로 나눌 수 있다.<sup>[3]</sup>

시간 분할 방식 스위치의 경우, RAM등 메모리 소자의 집적화 기술의 발전 및 액세스 속도의 증가로 말미암아 공유 메모리 방식(Shared Memory) 스위치나 공유 매체 방식(Shared Medium) 스위치 등 다양한 방식으로 연구가 활발히 진행되어 왔다. 그러나 이러한 구조의 스위치의 경우, 스위치의 성능 평가 기준이 되는 스루풋은 입, 출력 포트 수가 증가된 스위치를 구성시 메모리와 관련된 제반 기술의 발전에 전적으로 의존해야 한다는 단점을 갖고 있다. 즉, 입출력 포트의 수가 점차 증가함에 따라, 진입하는 각 포트에 대한 패킷들을 일시적으로 저장해야 하는 메모리의 용량이 상대적으로 증가해야 하며 또한 스위치에 연결된 외부 전송 매체의 속도에 상응하는 스위칭 속도를 얻기 위해 내부적으로 포트 수 증가, 메모리 읽기, 쓰기 등을 고려한 고속의 처리 속도가 요구된다는 단점을 갖고 있다. 따라서 이러한 문제점을 극복하기 위해 이와는 별도로 스위치의 구조적 병렬화를 이용하여 패킷을 고속으로 전송하기 위한 연구가 활발히 진행되고 있다. 이러한 시도로써의 공간 분할 방식 스위치는 단일 패킷 스트림(Packet stream)으로 스위치로 진입한 패킷들을 디멀티플렉싱과정을 통해 스위치 전 입력단으로 분산한 후, 공간적으로 다중 경로를 갖는 스위치 평면을 통해 전송하고자 하는 방식으로써 목적 포트에 전 입력단의 패킷들이 동시에 전송되는 방식의 스위치이다. 이러한 스위치의 대표적인 예로는 그림 1과 같이 크로스바망(Crossbar network) 스위치와 반얀망 스위치가 있다.

그러나 이러한 방식의 스위치 경우, 공간적으로 한정된 경로를 통해 다수의 패킷들이 동시에 전송을 시도하므로 동일 경로를 요구하는 패킷간에 충돌(Collision) 문제가 발생할 수 있다. 이러한 충돌을 해결하기 위한 방안으로 여러 가지 방식이 제안되고 있으나 이들 방식은 크게 버퍼(Buffer)의 존재 유무에 따라, 버퍼형 스위치와 비 버퍼형 스위치로 나눌 수 있다. 버퍼형 스위치란 충돌이 예상되는 스위치 내부 및 전단에 2개 이상의 패킷들을 저장할 수 있는 버퍼

를 설치하는 방식으로 충돌을 발생한 패킷들은 이 버퍼에 저장되어 충돌 여건이 해소된 후 재전송을 시도한다.<sup>[4] [5]</sup> 이러한 방식의 경우, 더 이상의 버퍼 공간이 없어 패킷이 소실되는 버퍼 오버플로우(overflow)로 인한 경우를 제외하곤 패킷 손실이 없는 장점을 갖고 있으나 연속적으로 내부 충돌이 발생할 경우, 버퍼 내에 저장되어 있는 패킷이 장시간 지연된다는 단점을 갖고 있다. 또한 스위치 내부 구조에 따라선 이러한 버퍼링(Buffering)으로 인해 패킷의 순서가 변경되어 전송되는 문제(Out of order Problem)가 발생할 수 있으므로 이를 제어하기 위한 별도의 부가장치를 마련해야 한다.

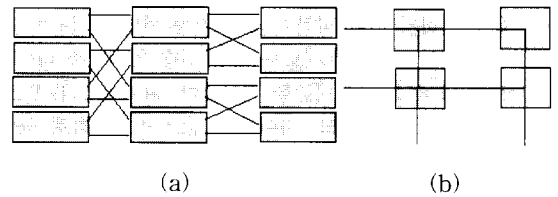


그림 1. 공간 분할 방식 스위치의 예

(a) 반얀망 스위치 (b) 크로스바망 스위치

Fig. 1. The Example of the Space Division Switch. (a) Banyan Network Switch (b) Crossbar Network Switch

비 버퍼형 ATM 스위치는 스위치 내부의 충돌을 최소한으로 허용하면서 고속의 패킷 전송 효과를 갖는다. 반얀망을 기반으로 하는 스위치의 경우, 반얀망이 갖는 고유한 충돌 문제, 즉 입, 출력간의 단일 경로 존재와 내부 링크의 공유로 인해 발생하는 내부 충돌 문제를 피하기 위해 과도한 하드웨어를 부가하고 있는 단점을 가지고 있다. 따라서 비 버퍼형 스위치에 대한 연구는 버퍼형 스위치 연구에 비해 이론적이고 제한적으로 진행되어 왔다. 선샤인(Sunshine) 스위치나 스타릿(Starlit) 스위치와 같은 경우엔, 배치망(Batcher network)<sup>[8]</sup>과 같은 정렬망(sorting network)을 다단으로 배치함으로써 스위치 내부의 패킷 충돌을 최소화하였으나, 복잡한 구조의 정렬망을 다단으로 구성함에 따른 전체 스위치 크기의 기하급수적 증가로 인해 실제의 구현이 어렵다. 그러나, 1991년 발표된 충돌 패킷의 誤-라우팅(Mis-Routing) 기법 및 스위치 평면간의 직렬 연결 방식을 이용한 탠덤 반얀 스위치(Tandem Banyan Switch)<sup>[9]</sup>경우와 1991년, 1995년 발표된 SE간 링크 수의 물리적 확장을 이용한 확

장 링크(Dilated) 상호 접속망 스위치<sup>[10]</sup>와 가변 확장 링크(Fat) 반양 스위치<sup>[2]</sup> 등과 같은 스위치는 스위치 평면 내부에 버퍼를 두지 않고도 간단한 링크의 확장 및 적절한 충돌 회피 제어 메커니즘의 적용만으로 높은 패킷 전송률을 갖는다.

그러나 이상의 스위치들은 앞서 언급한 바와같이 하드웨어 오버헤드를 증가시키는 단점과 더불어 이들 부가된 하드웨어를 충분히 활용하지 못하는 문제를 가지고 있다. 예를 들어 크로스바망 스위치의 경우, 입력 포트수  $N$ 에 대해  $N^2$ 의 SE를, 배쳐-반양 스위치의 경우  $(\log_2 N * (\log_2 N + 1))/2$ 의 SE를 필요로 하며, 탠덤 반양 스위치의 경우,  $k$ 를 부가되는 반양망의 수라고 가정시  $k * M \log_2 N$ 의 SE를 갖는다.

제안하는 전송 방식은 기존 스위치들의 문제점을 극복하기 위해 다음의 몇 가지 사항을 고려한다. 우선 스위치의 기본 형태는 패킷들의 고속 처리를 위하여 셀프 라우팅(Self-routing)이 가능한 반양망 기반 스위치를 대상으로 하며, 버퍼로 인한 불필요한 지연 시간의 단축을 위해 스위치 평면 내부에 버퍼를 두지 않는 것으로 한다. 또한 자원의 활용률을 높이기 위해 다단의 반양망 스위치 평면(Switch plane) 간에 바이패스 링크(Bypass link)를 부가하고, 이 링크 상의 패킷들을 제어하기 위한 다이내믹 바이패스 전송(Dynamic Bypass Transfer) 기법을 제안한다.

제안한 방식의 유효성 입증을 위해 기본적으로 입력단에 도달하는 패킷들은 랜덤 함수를 사용하여 발생하며, 스위치 평면의 수를 점진적으로 늘려 최적의 스위치 구성 조건을 찾아낸다. 2장에서는 반양망 기반 스위치의 개요 및 연구 배경을 기술하고 3장에서는 제안하는 다이내믹 바이패스 전송 기법을, 그리고 4장에서는 바이패스 링크 상의 패킷 제어를 위한 여분 경로 전송 기법을 제시한다. 5장에서는 제안한 방식의 시뮬레이션 결과를, 그리고 6장에서 결론 및 향후 연구 과제를 기술한다.

## II. 반양망 기반 ATM 스위치의 개요 및 분석

전 절에서 언급한 바와 같이 공간 분할형 ATM 스위치는 주로 반양망 스위치를 중심으로 설계된다. 반양망 스위치는 크로스바 망 스위치와는 달리 적은 수의 SE만으로도 전 입력 단으로부터 전 출력 단까지

패킷 전송이 가능하며, 각 입력으로부터 각 출력으로 가는 단일 경로를 갖는 특성을 갖는다. 이러한 반양망은 다중 스테이지 상호 접속망 (Multistage Interconnection Network ; MIN)의 대표적인 종류로써 L. R. Goke 와 G. J. Lipovski에 의해 1973년 발표되었다.<sup>[11]</sup> 각 SE는 기본적으로  $2 \times 2$  크로스바 망 (Crossbar network)으로 구성되며 구조적 모듈화로 인해 적은 포트 수의 스위치로부터 다수의 포트 수를 갖는 스위치로의 확장이 용이하다. 스위치는 내부적으로  $\log_2 N$ 개 스테이지와  $(N/2) * \log_2 N$ 개의 SE로 구성되어 있으며, 입력 포트로부터 출력 포트로의 전송은 스위치 전단에 부가된 입력 포트 단에 의해 별도의 라우팅 정보를 부가하여 스위치 내부 및 출력 포트단으로 전송된다. 각 스테이지의 SE는 라우팅 정보 내의 해당 비트만을 검색하여 그 정보 값에 따라 패킷의 흐름을 SE내의 상위 또는 하위 출력 단자로 연결하는 간단한 구조를 갖고 있다. 따라서 이상의 구조에서의 패킷들은 임의의 입력단으로부터 임의의 출력단까지 총 스테이지 수만큼의 단계를 거쳐 단일 경로를 형성하며 전송하게 되는데, 내부 구조적으로 스테이지간의 링크 수가 크로스바망에 비해 감소된 만큼 각 링크는 서로 다른 두 입력단과 공유되는 특징을 갖는다. 따라서 두 입력 단이 다음 스테이지로 전송되기 위해 SE의 공통 출력 단자로의 출력을 요구할 경우 이들 두 패킷간에 충돌이 불가피하며, 이러한 특성으로 인해 반양망 스위치는 크로스바망 스위치에 비해 낮은 스루풋을 갖는다. 따라서 이러한 특성을 갖는 반양망 계열 스위치 설계는 반양망 내부의 패킷 충돌을 해결하기 위한 구조 개발에 초점을 맞춰 다음과 같은 연구가 진행되어 왔다.<sup>[11]</sup>

- 외부 링크 속도에 대한 내부 링크 속도 고속화 방식
- 충돌이 예상되는 모든 SE에 버퍼 부가 방식
- 충돌 패킷을 지연시키기 위한 스테이지간의 상호 메커니즘 및 Back-Pressure 메커니즘 적용 방식
- 다중 스위치 평면 적용에 따른 입, 출력 포트간 다중 경로 설정 및 다중 링크 설정 방식
- 부하를 공평히 분배하기 위한 반양망 전단의 분산 망(Distributed network) 부가 방식
- 비 충돌 패킷 패턴 발생을 위한 정렬 망 부가

방식

이러한 스위치들은 그림 2와 같이 충돌 패킷의 흐름을 기준으로 분류시, 크게 충돌된 패킷을 스위치의 입력단으로 재 진입시켜 재전송의 기회를 부여하는 피드-백(Feed-back) 방식과 충돌 패킷들을 임의의 여유 출력 포트에 전송한 후 다음 스위치 평면으로 다시 진입을 시켜 재 라우팅을 시도하는 誤-라우팅(Mis-Routing) 방식 그리고 충돌이 발생한 패킷을 충돌 여건이 해소될 때까지 버퍼에 저장한 후 재전송을 시도하는 저장-전 방향(Store and Forward) 전송 방식, 스테이지의 SE간에 여분의 물리적 경로를 설정하여 같은 목적지로 향하는 패킷들을 복수 링크를 통해 다음 스테이지의 SE로 전송하는 확장(Dilated) 전송 방식 등으로 나눌 수 있다.

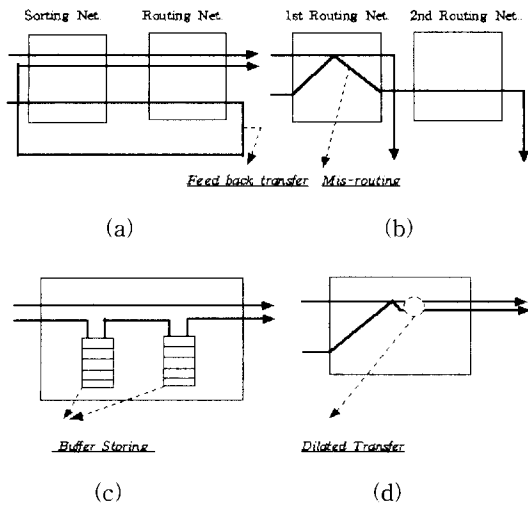


그림 2. 충돌 패킷 흐름을 기준으로한 분류의 예 (a) 피드-백 전송 방식 (b) 誤 라우팅 전송 방식 (c) 저장-전 방향 전송 방식 (d) 확장 전송 방식

Fig. 2. The Example of Classification by Conflict Packet Flow. (a) Feedback Transfer Method (b) Mis-Routing Transfer Method (c) Store and Forward Transfer Method (d) Dilated Transfer Method

그러나 선샤인(Sunshine)<sup>[6]</sup> 스위치나 스타릿(Starlit)<sup>[7]</sup> 스위치 및 개량된 탄뎀(Tandem) 반안망 스위치의 일부에 적용된 피드-백 전송 방식의 경우, 충돌을 발생한 패킷들은 다음 패킷 슬롯에 스위치에 진입한 패킷들과 혼재되어 스위치 내의 복잡도를 증가

시키므로 이를 해소시키기 위한 부가적인 하드웨어를 필요로 하게 된다. 이에 반해 탄뎀 반안망 스위치에서 적용된 誤 라우팅 전송 방식의 경우나 가변 확장(Fat) 반안망 스위치등 확장 전송 방식을 적용한 스위치의 경우, 충돌을 발생한 패킷들이 스위치 평면내 불필요한 경로를 점유하는 문제 및 부가 스위치 평면의 양적 증가에 비해 패킷 전송 활용도가 떨어진다는 단점을 갖고 있다. 마지막으로 버퍼형 반안망등에서 적용하고 있는 저장-전 방향 전송 방식의 경우 단순한 반안망에 적용시 버퍼의 용량이 패킷 양에 비례하여 증가해야 하며 다중 경로를 갖는 스위치 평면에 저장-전 방향 전송 방식을 적용할 경우, 동일 입력단으로 진입하는 패킷들의 순서를 보장하기 위한 별도의 부가 장치를 필요로 하는 단점을 가지고 있다.

따라서 이상적인 ATM 스위치 구성을 위해선 기본적으로 적은 수의 SE로 구성이 되어야 하고 충돌 패킷들의 재전송 기회 부여를 위해 다중 경로가 설정되어야 하며, 충돌을 발생한 패킷들의 불필요한 스위치 평면 내 경로 점유를 막아야 하고 전 스위치 평면을 충분히 활용할 수 있어야 한다.

Ⅲ. 다이내믹 바이패스 전송 기법

스위치로 진입하는 패킷들을 고속으로 전송하고 패킷간의 충돌이 발생할 경우, 즉각적인 충돌 회피 과정을 수행하기 위하여, 본 논문은 그림 3과 같은 다이내믹 바이패스 전송 기법을 제안한다.

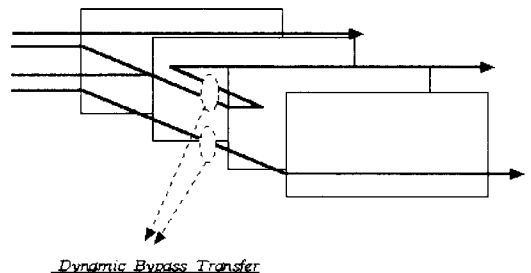


그림 3. 다이내믹 바이패스 전송 Fig. 3. Dynamic Bypass Transfer.

제안하는 전송 기법은 한 스위치 평면에서 충돌을 발생한 패킷을 다른 스위치 평면에서의 동일 스테이지, 동일 SE간의 여유 경로의 존재 여부를 확인한 후 전송을 시도하는 방식으로써, 기존의 충돌 패킷 제어 방

식이 단 방향적인데에 반해 제안하는 기법은 각 스위치 평면 내의 동일 위치의 SE간에 교번식 패킷 전송을 유도함으로써 스위치 내의 모든 자원을 충분히 활용할 수 있는 장점을 가지고 있다. 또한 충돌이 발생한 링크는 기존의 방식에선 경로로써의 기능을 상실하는 데에 반해 제안하는 전송 기법을 적용 시엔 충돌을 경험한 경로도 이후 다른 충돌 패킷의 여유 경로로써 활용될 수 있는 장점을 갖는다. 이러한 전송 방식에서의 충돌 패킷 흐름은 그림 4와 같이 다음의 세 가지 상태를 갖는다.

- 상태 i : 일반 전송 상태
- 상태 ii : 바이패스 전송 상태
- 상태 iii : 바이패스 충돌 상태

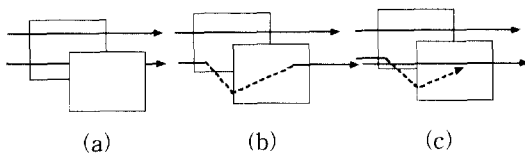


그림 4. 패킷 흐름 상태  
 (a) 일반 전송 상태 (b) 바이패스 전송 상태  
 (c) 바이패스 충돌 상태  
 Fig. 4. Packet Flow States.  
 (a) Normal Transfer State (b) Bypass Transfer State (c) Bypass Conflict State

일반 전송 상태는 SE내에 주 입력 단으로 진입한 패킷들을 충돌 없이 SE내 주 출력단으로 전송하는 단계를 의미하며, 현재 패킷이 진행되어 온 스위치 평면 상으로 정상적인 전송이 이루어진다. 바이패스 전송 상태는 일반 전송 상태에서 충돌을 발생한 패킷들이 다른 스위치 평면 SE의 여유 경로로 전송되는 단계를 의미하는 것으로 패킷이 진행되어 온 스위치 평면이 아닌 다른 스위치 평면으로 정상적인 전송이 이루어진다. 마지막으로 바이패스 충돌 상태란 일반 전송 상태에서 충돌을 발생한 패킷이 여유 경로를 찾지 못하고 폐기되는 단계로써 다른 스위치 평면을 통해 전송을 시도하는 패킷이 이미 그 스위치 평면의 경로를 점유하고 있는 패킷으로 인하여 충돌이 발생하는 상태를 의미하며, 이러한 경우 같은 스위치 평면으로 진행하고 있는 패킷은 우선 순위를 갖고 정상적인 전송이 이루어지며 다른 스위치 평면으로부터 진행되어 온 패킷은 폐기된다.

이상의 세 가지 상태를 갖는 다이너믹 바이패스 전송 기법을 적용함으로써 패킷들은 스위치 내의 모든

링크를 능률적으로 이용하며 전송되며 바이패스 충돌을 겪은 패킷들은 스위치내 패킷 전송에 부담을 주는 패킷으로 간주되어 폐기되므로, 정상적인 전송이 가능한 패킷들에 아무런 영향을 미치지 않게 된다.

#### IV. 여분 경로 전송 메커니즘이 적용된 ATM 스위치

제안하는 스위치는 그림 5의 구조를 갖는다. 기본적으로 입,출력 포트, 다단의 스위치 평면 그리고 다이너믹 바이패스 전송 방식을 적용하기 위한 바이패스 라인(Bypass Line), 바이패스 컨트롤러(Bypass Controller)로 구성된다. 실제 라우팅이 이루어지는 공간인 병렬 구조의 스위치 평면 각각은 기본적인  $N \times N$  반안망으로써 전체 스위치 구성은 이 스위치 평면을 다수개 병렬로 연결함으로써 구성된다.

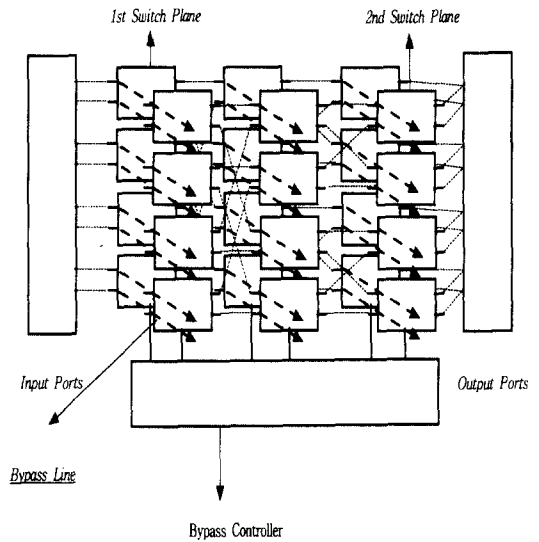


그림 5. 바이패스 라인을 적용한 ATM스위치  
 Fig. 5. ATM Switch with Bypass Lines.

바이패스 라인은 주 스위치 평면으로부터 병렬로 부가된 여분의 스위치 평면들 간을 연결하는 물리적 링크로써, 스위치 평면간의 동일 스테이지, 동일 SE간을 연결한다. 바이패스 라인 상에서의 충돌 패킷 전송은 바이패스 컨트롤러에 의해 제어된다. 바이패스 컨트롤러는 충돌 패킷을 발견할 경우, 다른 스위치 평면상의 여유 공간의 존재 유무를 나타내는 동일 스테이지 동일 위치의 SE들에 대한 ACK(ACKnowledge) 신호를 검색한 후, 여유 공간이 존재함을 의미하는 ACK

= 0일 때에 한해 해당 스위치 평면으로 충돌이 예상되는 패킷을 전송하고, 만일 ACK = 1이라면 전송을 위한 여유 공간이 확보되어 있지 않다는 것을 의미하므로 충돌 패킷은 현재 버퍼 상에서 제거된다. 예를 들어 그림 6 과 같이 충돌을 발생하는 출력 패턴을 갖는 패킷이 스위치 내로 진입한다고 가정하자.

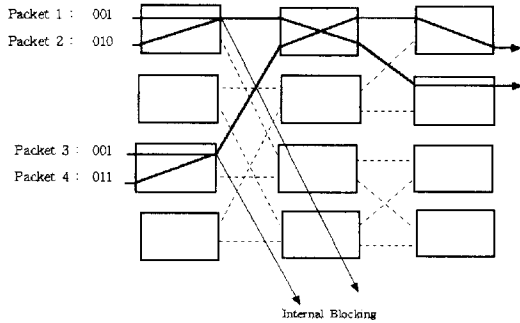


그림 6. 8 x 8 반얀망에서의 충돌 예  
Fig. 6. The conflict example on 8 x 8 Banyan network.

그림에서와 같이 (0,0,1)의 목적 포트를 갖는 패킷 1 과 목적 포트 (0,1,0)의 패킷 2는 첫 스테이지의 라우팅 비트인 0 (0,0,1)과 0 (0,1,0)이 동일 링크로의 전송을 요구하므로 내부 충돌이 발생하게 된다. 그러나 제안한 다이내믹 바이패스 전송 기법을 적용하면, 동일 링크를 요구하는 패킷들을 여분의 전송 경로를 통해 충돌 없이 출력 포트로 전송할 수 있다. 그림 7은 제안된 메커니즘이 적용된 2단 병렬 반얀망의 예이다.

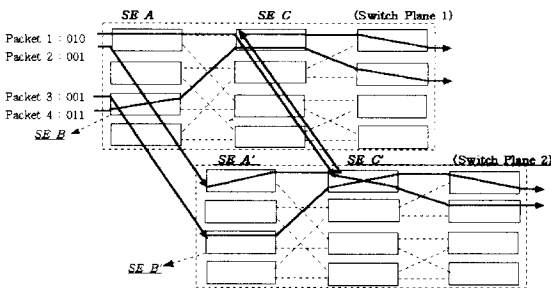


그림 7. 여분 경로 전송 메커니즘이 부가된 병렬 반얀망  
Fig. 7. The Parallel Banyan Networks with Alternate Path Transfer Mechanism.

앞서 그림 6에서 충돌을 발생했던 패킷 1과 패킷 2는 그림 7과 같이 SE A에서 동일 목적지로의 전송을 위해 경쟁을 한후, 이중 패킷 1은 충돌 없이 주 스

위치 평면으로 전송된다. 기존의 반얀망에선 충돌할 경우 폐기될 패킷 2는 여분 전송 경로인 다른 스위치 평면의 목적 링크로의 전송 가능 여부를 ACK신호 참조를 통해 확인한 후, 만일 ACK = 0 (즉, 여유 공간이 확보된 경우)인 경우엔 패킷 2를 여분 전송 경로의 전송을 위해 SE A'로 전송한다. 이후, 스위치 평면 2로 전송된 패킷 2는 해당 라우팅 비트에 따라 전송되어야 하는 다음 스테이지의 SE C'로 전송된다. 마찬가지로, 패킷 3과 패킷 4의 경우 역시 첫 스테이지에서 동일 출력 링크로의 전송을 요구하므로 SE B에서 경쟁을 한후, 패킷 4가 경쟁에 이겼을 땐 패킷 3은 바이패스 링크를 통해 스위치 평면 2의 SE B'를 거쳐 다음 전송 경로인 SE C'로 전송을 실행한다. 한편, SE C내에선 패킷 1의 두 번째 스테이지의 라우팅 태그 1 (0,1,0) 과 패킷 4의 라우팅 태그 1 (0,1,1)이 같은 경로로의 출력을 요구하므로 한 패킷은 진행되어 온 스위치 평면의 목적 링크로, 다른 한 패킷은 여분 전송 경로를 통해 전송을 시도하기 위해 다른 스위치 평면으로 전송된다. 또 SE C'에선 앞선 스테이지에서 충돌을 겪어 스위치 평면으로 진행된 패킷 2 와 패킷 3이 다시 동일 목적 링크로의 전송을 요구하므로 이상과 같은 방법으로 전송이 진행된다. SE내에서의 경쟁은 기존의 반얀망 계열 스위치와 마찬가지로 결정 로직(Decision Logic)을 통해 선택되나, 시뮬레이션에선 간단한 프로그램을 통해 랜덤 하게 선택되도록 하였다.

이상의 예에서는 네 개의 패킷 모두가 원하는 목적지로 전송됨을 보였으나, 실제의 경우 그림 7에서와 같이 모든 패킷들이 전송되는 것은 아니다. 그림 8은 여분 경로 전송 메커니즘을 적용 시에 나타나는 충돌의 예이다.

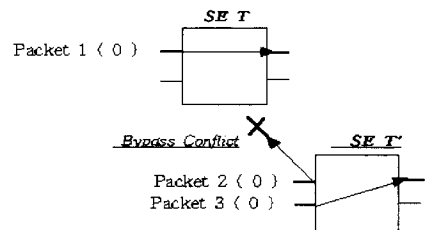


그림 8. 여분 경로 전송에서의 충돌의 예  
Fig. 8. The Example of Conflicts during Alternate Path Transfer.

두 스위치 평면 SE T와 SE T'을 가정시, SE

T'에서 그림과 같이 패킷 2와 패킷 3이 모두 0번 출력을 요구하여 경합을 하였다고 할 때, 패킷 3이 정상 전송을, 패킷 2가 SE T'내에서 충돌했다고 가정할 경우, 패킷 2는 여분 경로인 SE T에서의 0번 포트의 전송 가능 여부를 확인하게 된다. 그러나 SE T에서의 0번 포트는 이미 앞선 스테이지의 SE로 부터 SE T로 진입해 있는 패킷 1에 의해 그 경로가 점유되고 있는 상태이므로, 우선 순위에 따라 패킷 2는 더 이상의 여분 경로 전송을 실행 할 수 없게 된다. 따라서 패킷 2는 SE T'내에서 제거되며, 이러한 바이패스 충돌 상태는 제안하는 스위치에서 스루풋을 떨어뜨리는 유일한 요인이 된다. 이상의 메커니즘을 구동시키기 위해 각 SE는 그림 9와 같이 구성한다. SE는 IC(Input Controller), OC(Output Controller), BCL(Bypass Control Line), BDL(Bypass Data Line) 으로 구성된다.

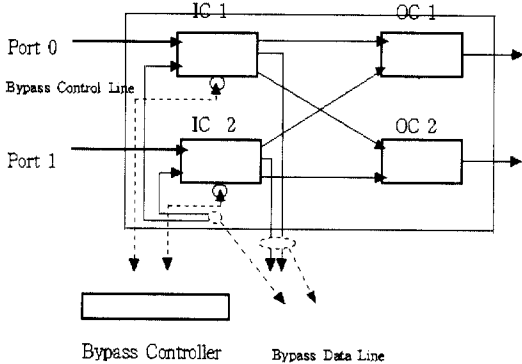


그림 9. SE의 내부 구조  
Fig. 9. The Internal Structure of single SE.

BCL은 바이패스 컨트롤러와 SE내의 상위 IC 및 하위 IC와 신호를 주고받는 연결선으로써 바이패스 컨트롤러는 이 라인을 통해 각 IC내의 버퍼 사용 여부 및 경합 결과에 대한 정보를 받으며 여분 경로 전송이 가능할 경우 해당 IC로 ACK 신호를 보내 바이패스 전송을 승인한다. BDL은 실제 바이패스되는 데이터가 전송되는 라인으로서 IC내 버퍼로의 패킷 송신이나 다른 스위치 평면에 있는 SE로의 패킷 전송시 데이터를 주고받는 라인이다. IC는 스위치 내의 일반 전송시 출력 포트를 결정하는 결정 로직(Decision logic)과 라우팅 태그 비교를 위한 패킷 하나를 저장할 수 있는 크기의 버퍼 및 다른 스위치 평면으로부터 입력되는 패킷이 일시적으로 저장될 수 있는 버퍼로 구성된다.

SE내에서의 패킷 전송은 다음의 단계를 거친다.  
단계 1 : ① SE로 패킷들이 전송되면 비어 있는 버퍼는 갖는 IC가 있는지 확인한 후, 비어 있는 버퍼를 갖는 IC가 있는 경우엔 바이패스 컨트롤러로 Empty신호를 보낸다.

② 바이패스 컨트롤러는 각 IC로부터 Empty 신호를 받아 참조 테이블을 작성 한다.

단계 2 : 각 IC의 패킷들은 결정 로직을 통해 해당 비트를 비교한다. 만일 동일 목적지를 갖지 않는 경우엔 해당 OC로 정상 전송을 하며, 동일 목적지를 가질 땐 바이패스 컨트롤러에 바이패스 요청 신호 ACQ (Acquire)를 보낸다.

단계 3 : 바이패스 컨트롤러는 ACQ신호를 받으면 테이블을 참조하며, 전송 가능한 경로가 존재시 ACK신호를 해당 IC로 보낸다.

단계 4 : ACQ 신호를 발생한 IC들은 ACK신호를 받으면 바이패스 링크를 통해 해당 IC를 거쳐 OC로 패킷을 전송하고, NACK 신호를 받으면 자신의 패킷들을 폐기한다.

단계 5 : 다음 스테이지 슬롯에 OC의 패킷들을 다음 스테이지의 IC로 전송한다.

### V. 실험 결과 및 검토

제안한 메커니즘을 적용한 스위치의 성능을 평가하기 위해 임의의 입력 단에서 임의의 출력 단으로의 패킷 전송 확률이 균일한 분포를 갖는 패킷들을 발생시켰다. 또한, 제안한 모델은 C언어를 사용하여 볼랜드 C++ 컴파일러 환경 하에서 구현하였다. 스위치 입력 단은 각 노드에 도착한 ATM 패킷들을 다음 목적 노드로 전송하기 위해 스위치 내의 해당 경로를 지정하는 ATM 스위칭 헤더를 ATM 패킷 전단에 부가한다. 8 x 8 스위치 내에서 전송되는 패킷의 구조는 그림 10과 같다.



Pri. bit : (Priority bit)      Conf. bit : (Conflict bit)  
Act. bit : (Activity bit)    D3, D2, D1 : (Routing Tag)

그림 10. ATM 스위칭 헤더의 구조  
Fig. 10. The Structure of ATM Switching Header.

ATM 유료 부하(Payload) 부분은 실제 스위치에 도달한 53 바이트의 ATM 패킷을 의미하며 ATM 스위칭 헤더 부분은 실제 스위치내 전송을 위한 정보를 담고 있는 공간이다. 그림 11에서와 같이 ATM 스위칭 헤더는 D3, D2, D1 등 스테이지 수와 같은 비트 수를 갖는 라우팅 정보 영역과 우선 순위를 위한 우선 순위 비트(Priority bit), 충돌의 경험 여부를 나타내는 충돌 비트(Conflict bit), 패킷내 ATM 정보 유무를 나타내는 활동 비트(Activity bit)로 구성된다. 실제 SE 내에서 충돌이 발생하면 여분 경로 전송이 되어야 할 패킷들은 충돌 비트로 '1' 값을 가지며, 만일 여분 경로 전송이 성공하였으면 이 비트는 다시 '0'으로 변환되어 다음 스테이지로의 전송을 완료하며, 만일 실패하였다면 SE에서 제거된다.

우선 순위 비트는 예비 비트로써 출력 포트 단에서의 출력 버퍼에서의 우선 순위 기반 전송 등 장차 있을지 모를 스위치내 우선 순위 제어에 대비하여 확보된 공간이나 본 시뮬레이션에선 사용되지 않았으며, 이상과 같은 ATM 스위칭 헤더는 총  $(3 + \log_2 N)$  bit만큼의 크기를 갖는다. 성능 평가를 위한 시뮬레이션은 다음의 기본 절차를 따른다.

- i) ATM 스위치 입력 단으로부터 출력 단까지 패킷이 충돌 없이 전송되는 시간을 패킷슬럿으로 정의하고 한 패킷슬럿은 스테이지 수만큼의 스테이지 슬럿으로 나눠 전 스위치 평면의 동일 스테이지는 동일 스테이지 슬럿으로 동기적으로 동작시킨다.
- ii) ATM 스위치로 진입한 패킷들은 일반 전송 단계 ↔ 바이패스 전송 단계 ↔ 스테이지 전송 단계를 반복적으로 진행하며 이 세 단계 과정은 최종 스테이지에 도달할 때까지 반복한다.
- iii) 공간 분할형 스위치의 입력 패킷 발생은 출력 포트 번호가 균일한 분포를 갖는 경우나 Hot spot 패턴과 같이 다수의 충돌 가능 패턴을 갖는 입력 패킷을 발생 하여 성능을 평가하는 방법이 있으나 본 시뮬레이션에선 출력 포트 번호가 균일 분포를 갖도록 제한하였다.

또한 올바른 시뮬레이션 수행을 위한 기본 전제는 다음과 같다.

- i) 외부 링크 및 내부 링크는 동일한 전송 속도를 갖는 것으로 가정한다.
- ii) 바이패스 전송 단계에서의 여유 공간 검색 시간

은 무시 할 수 있을 만큼 작은것으로 가정한다.

- iii) 최종 출력 단은 동일 목적지로 향하는  $k$ 개의 패킷들을 일시적으로 저장해야 할 한 패킷을 저장할 수 있는 크기의  $k$ 개의 병렬 버퍼 및 이들을 고속을 멀티 플렉싱하기 위한 멀티 플렉서, 그리고 멀티플렉싱된 패킷들을 저장하여 외부 링크로 직렬 전송할  $k$ 개 크기의 FIFO 버퍼가 필요하다. 본 시뮬레이션에선  $k$ 개의 FIFO 버퍼를 무한으로 가정하여 목적지가 같은 패킷들은 최종 출력단에서 충돌 없이 전송되는 것으로 제한하였다.

먼저 제안된 스위치 구조 및 메커니즘을 적용한 경우, 여분 경로 전송 메커니즘이 갖는 효과를 평가하기 위하여 2단 스위치 평면을 갖는 스위치(X2)와 일반 반양망 스위치(X1)의 각 부하당 스루풋을 비교하였다.

표 1. 단일 반양망과 여분 경로 전송 기법이 적용된 2단 병렬망의 스루풋 비교

Table 1. Throughput Comparison between single Banyan Network and Banyan Network with 2 Switch Planes.

		8 x 8	16 x 16	32 x 32	64 x 64	128 x 128	256 x 256
X 1	1.0	0.515350	0.449056	0.399438	0.360159	0.327027	0.300218
	0.9	0.534994	0.466350	0.413667	0.369417	0.334957	0.307114
	0.8	0.543335	0.461433	0.416239	0.372320	0.337166	0.308726
	0.7	0.546938	0.473407	0.417569	0.374053	0.338680	0.310474
X 2	1.0	0.892200	0.845513	0.804669	0.769028	0.736955	0.70797
	0.9	0.897779	0.850822	0.810254	0.773585	0.740932	0.711476
	0.8	0.896311	0.851425	0.810240	0.773977	0.740942	0.711538
	0.7	0.900563	0.851936	0.811888	0.774578	0.741480	0.712147

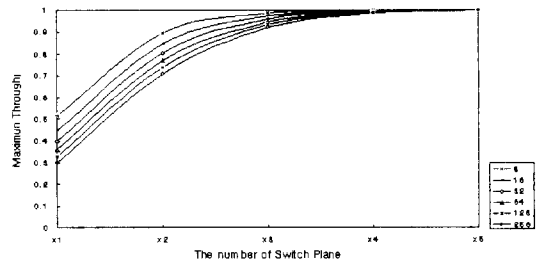


그림 11. 제안된 방식이 적용된 스위치의 최대 스루풋 (부하 = 1)

Fig. 11. Maximum Throughput of ATM Switch with proposed method. (Load = 1)

표 1과 같이, 부하가 0.8인 경우의 단일 반양망의



스루풋이 0.543335인데 에 반해 두 개의 반얀망을 제안된 기법으로 연결한 스위치의 경우 0.896311로 크게 향상됨을 알 수 있었다. 이것은 단일 반얀망에선 내부 링크 공유로 인한 다수의 충돌이 발생하는데 반해, 제안한 방식의 스위치의 경우 충돌이 가능한 패킷 모두가 여분 경로를 통한 재전송의 기회를 가지므로 스루풋에 상당한 향상을 가져왔음을 알 수 있었다. 그림 11은 8 x 8 부터 256 x 256 스위치까지의 스루풋을 나타낸다. X1, X2, X3, X4, X5는 스위치를 구성하는 스위치 평면의 갯수를 의미한다.

평가 기준이 되는 최대 스루풋(Maximum Throughput)은 입력 단에 도달하는 패킷의 부하가 1일 때의 스루풋으로써 정의는 다음과 같다.

(Load가 1인 경우)

$$\text{최대 스루풋} = \frac{\text{출력 단에서의 패킷 전송률}}{\text{입력 단으로의 패킷 도달률}} \quad (1)$$

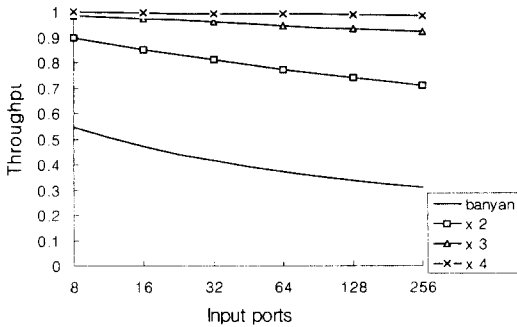


그림 12. 입력 포트수 증가에 따른 스루풋 (부하 = 0.8)

Fig. 12. Throughput for Increased Input Ports. (Load = 0.8)

64 x 64의 경우 4단의 병렬 반얀망의 경우 거의 스루풋 1에 근사하는 성능을 보이며, 256 x 256의 경우 5단의 병렬 반얀망의 경우 스루풋 1에 근사함을 보여준다. 모든 스위치 경우 X1에서 X3사이에서 급격한 스루풋 변화를 보이다가 X3에서 X5까지는 안정된 그래프 모양을 보여주고 있다. 이는 제안된 방식을 병렬망에 적용할 경우, 병렬 반얀망의 수가 일정한 갯수 이상이 되면 스루풋이 거의 일정하게 유지되는 것을 알 수 있으며, 스루풋 향상을 위해 무한정 스위치 평면을 늘리는 것이 아닌 3단부터 5단 사이 정도의 스위치 평면으로 최상의 스루풋을 얻을 수 있음을 의미

한다. 그림 12는 입력 포트 수에 대한 스루풋을 나타낸 그래프이다.

기존의 한 단만으로 구성된 반얀망 스위치의 경우에 비해 제안한 방식의 스위치의 경우, 포트수 증가에도 완만한 스루풋 감소 곡선을 보여주고 있다.

스위치 평면 증가에 따른 스루풋의 향상률을 알기 위해, 그림 13은 병렬 구조를 갖는 다단 스위치 평면에 따른 다이너믹 바이패스 전송 기법 적용 시의 스루풋 향상률을 나타내었다.

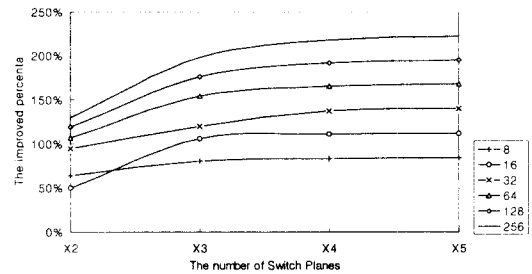


그림 13. 병렬망 증가에 따른 스루풋 향상률 (부하 = 0.8)

Fig. 13. The improved percentage for Throughput. (Load = 0.8)

X2의 경우, 기존 반얀망의 스루풋에 비해 64%부터 130%까지의 성능 향상을 보여주고 있으며, X3스위치의 경우 81%부터 198%까지, X4 스위치의 경우 83%부터 218%까지의 스루풋 향상을 보여주고 있다. 또한 제안한 메커니즘이 적용된 스위치의 경우 적은 포트수의 스위치 경우보다 다수의 포트를 갖는 경우에서 나은 성능을 보여줌을 알 수 있었다. 그림 14는 부하 (입력 부하)에 대한 출력 부하에 대한 그래프를 나타내었다.

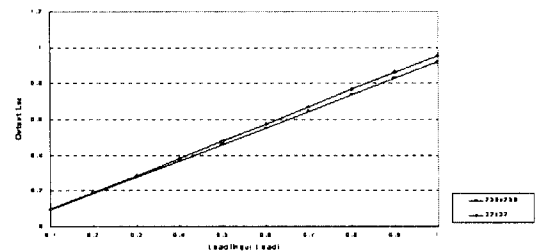


그림 14. 입력 부하에 따른 출력 부하 (X = 3)

그림 14에 나타난 것과 같이 출력 부하는 입력 부하에 거의 정비례하여 증가함을 알 수 있었다. 따라서

제한한 스위치의 경우, 부하가 증가할수록 다이나믹 바이패스 전송 기법의 효과가 두드러짐을 알 수 있었다.

그림 15는 여분 전송 경로를 갖는 32 x 32 ATM 스위치의 패킷 손실률을 나타내었다. 기존의 탄뎀 반안 스위치의 경우에 비해 제안한 전송 방식을 이용한 병렬 구조 반안망 스위치가 보다 나은 성능을 보여주고 있다. 포트 수를 증가시킨 시뮬레이션 결과에선 두 스위치의 격차가 32 x 32 입,출력 경우보다 보다 큰 것으로 나타났다. 이상의 결과에서 제안한 방식의 ATM 스위치는 탄뎀 반안망 스위치에 비해 적은 수의 스위치 평면만으로도 높은 스루풋을 나타내었으며 포트 수가 증가함에 따라 보다 나은 성능을 나타냄을 알 수 있었다.

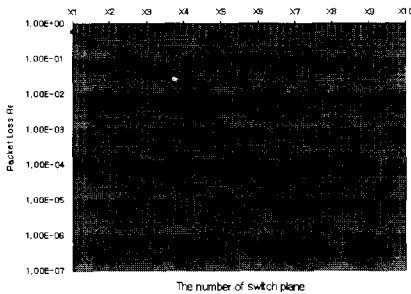


그림 15. 제안한 방식의 32 x 32 ATM 스위치와 탄뎀 반안 스위치의 패킷 손실률 (부하 = 0.8)  
Fig. 15. Packet Loss Rate of proposed 32 x 32 ATM Switch and Tandem Banyan Switch. (Load = 0.8)

## VI. 결론 및 향후 연구 과제

본 논문에서는 충돌 패킷들의 재전송 기회 부여 및 충분한 스위치 자원 활용이 가능한 새로운 여분 경로 전송 메커니즘을 제안하였다. 기존의 ATM 스위치가 충돌 패킷들에게 경로 점유를 허용한 채, 피드-백 방식이나 黽-라우팅 기법 등을 통해 재전송을 시도하거나 라우팅 태그의 변경 등으로 인해 셀프 라우팅이 불가능했던 것에 반해 제안한 여분 경로 전송 메커니즘은 기존의 단 방향성 병렬 구조 반안망 스위치에 다이나믹 바이패스 전송 기법을 적용하여, 충돌이 발생시 충돌한 패킷들에 대한 즉각적인 여분 경로 검색을 실행한 후 여유 경로로 충돌 패킷들을 전송하였다. 또한 여분 경로가 발견되지 않을 시엔 불필요한 버퍼 및 경

로 점유로 간주하여 해당 패킷을 제거했다. 또한 입력 단에 도착한 패킷들은 기존의 순수 반안망 스위치와 마찬가지로 라우팅 태그 변환 없이 최종 목적지로 전송 가능하도록 셀프 라우팅 되었으므로 별도의 복잡한 제어 및 지연 시간이 없는 즉각적인 패킷 전송이 가능하였다. 시뮬레이션 결과, 기존의 충돌 패킷 전송 메커니즘과는 달리 제안된 메커니즘을 적용한 스위치는 소수의 반안망만으로도 충분히 낮은 패킷 손실률을 보장함을 알 수 있었다.

제한한 방식의 경우, 스위치 입력 포트 수의 증가 및 스위치 평면 수의 증가로 말미암은 바이패스 링크 복잡도가 증가하는 문제가 발생할 수 있으나, 시뮬레이션 결과 스테이지에 도달한 패킷당 평균 바이패스 전송 패킷의 수가 최대 0.15를 넘지 않으므로 각 SE로의 상호 연결이 아닌 바이패스 데이터 버스를 이용으로 그 복잡도를 감소할 수 있다. 최종 출력 단에서의 구조는 그림 16과 같이 구성할 수 있으며, 패킷 슬롯의 스위치 평면 갯수 배 만큼 속도의 멀티플렉싱으로 출력 충돌을 막을 수 있다. 그러나 실제 구현시 외부 링크에 비해 내부 버퍼링 속도가 고속화됨에 따른 출력 버퍼 오버플로우 현상이 발생할 수 있으며, 향후 연구 과제로는 이러한 오버플로우 현상을 제거할 수 있는 고속 멀티플렉싱 기술을 개발하는 것이다.

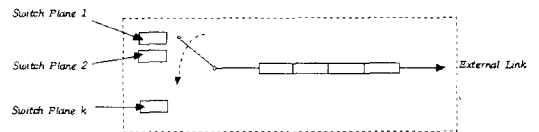


그림 16. 최종 출력 단의 출력 포트  
Fig. 16. Output Port in Final Output Stage.

## 참고 문헌

- [1] Hamid. Ahuai and Wolfgang E. Denzel, "A Survey of Modern High-Performance Switching Technique," IEEE J. Select. Areas. Commun., vol. 7, No. 7, Sep., 1989.
- [2] M. Alimuddin *et al.*, "The Fat Banyan ATM Switch," INFORCOM., pp. 659-666, 1995.
- [3] Thomas M. Chen and Stephen S. Liu, "ATM Switching Systems," Artech House, pp. 6-13, 1995.

- [4] J. S. Turner, "Design of a Broadcast Packet Switching Network," IEEE Trans. Commun. vol. 36, pp. 734-743, June 1988.
- [5] Yu-shuan Yeh *et al.*, "The Knockout Switch : A Simple, Modular Architecture for High-Performance Packet Switching," IEEE J. Select. Areas. Commun., vol. SAC-5, No. 8, Oct., 1987.
- [6] James N. Giacomelli *et al.*, "Sunshine: A High-Performance Self-Routing Broadband Packet Switch Architecture," IEEE J. Select. Areas. Commun., vol 9, No. 8, Oct., 1991.
- [7] Huang A. and S. Knauer, "Starlit: A Wideband Digital Switch," GLOBECOM' 84, pp. 121-125 Atlanta, Dec., 1984.
- [8] K. E. Batcher, "Sorting networks and their application," IEEE J. Select. Areas. Commun., vol. 7, No. 7, Sep., 1989.
- [9] Fouad A. Tobagi *et al.*, "Architecture, Performance, and Implementation of The Tandem Banyan Fast Packet Switch," IEEE J. Select. Areas. Commun., vol. 9, No. 8, Oct., 1991.
- [10] E. T. Bushnell and T. S. Meditch, "Dilated Multistage Interconnection networks for Fast Packet Switching," IEEE INFOCOM pp. 1264-1273, 1991.
- [11] L. R. Goke and G. J. Lipovski, "Banyan Networks for Partitioning Multiprocessor Systems," in Proc. First Ann. Symp. on Computer Architecture, pp.21-28, IEEE, 1973.
- [12] Fabio M. Chiussi, "Design Performance and Implementation of a Three-stage Banyan-based Architecture with Input and Output Buffers for Large Fast Packet Switches," Partial Fulfillment of the Requirements for Ph. D. Stanford Univ., July, 1993.
- [13] Joo-Young Lee, Byung-Chun Lee, In-Chil Lim, "Design of ATM Switch for Redundant Path Transfer," Proceedings of KITE Summer Conference 96, vol 19, No. 1, pp. 77-80, Jun., 1996.
- [14] Shigeo Urushidani, "Rerouting Network : A High-Performance Self-Routing Switch for B-ISDN," IEEE J. Select. Areas. Commun., vol. 9, No. 8, pp. 1194-1204, Oct., 1991.

---

 저 자 소 개
 

---



李周榮(正會員)

1968년 2월 20일생. 1990년 2월 한양대학교 전자공학과 졸업(공학사). 1992년 2월 한양대학교 대학원 전자공학과 졸업(공학석사). 1996년 2월 한양대학교 대학원 전자공학과 박사과정 수료. 관심분야

는 ATM망에서의 스위치 구조 설계, 신경회로망.

林寅七(正會員)

第 30卷 B編 第 2號 參照

현재 한양대학교 전자공학과 교수