

고속 인터넷 통신망을 위한 스위치 설계에 관한 연구

A Study on the Design of Switch for High Speed Internet Communication Network

조 삼 호*
Sam-Ho Cho

요 약

복잡한 통신망과 병렬컴퓨터에서는 효율적인 상호 연결을 위해 스위치가 중요한 영향을 미친다. 이 스위치는 라우팅 정보에 따라서 입력포트와 출력포트 사이에 연결을 해주는 역할을 하게 된다. 따라서 스위치에 성능을 향상시키는 것은 중요한 일이다. 본 논문에서 제안한 스위치는 컴퓨터 시뮬레이션 결과 입력버퍼형 보다 최대처리율이 11%이상 향상되었고, 다른 반안형 스위치들에 비하여 성능과 하드웨어 양을 비교하여 볼 때 좋은 결과를 얻었다. 따라서 이 스위치는 VLSI칩으로 구현될 경우 초고속 ATM/LAN 과 병렬컴퓨터를 개발하는데 유용하게 활용될 수 있을 것이다. 제안된 반안형 스위치는 MAX+PLUS II, VHDL을 이용하여 설계 및 검증을 하고 시뮬레이션을 하였다.

Abstract

A complex network and a parallel computer are made up of interconnected switching units. The role of a switching unit is to set up a connection between an input port and an output port, according to the routing information.

We proposed our switching network with a remodeled architecture is a newly modified Banyan network with eight input and output ports. We have analysed the maximum throughput of the revised switch. Our analyses have shown that under the uniform random traffic load, the FIFO discipline is limited to 70%. The switching system consists of an input control unit, a switch unit and an output control unit. Therefore the result of the analyses shows that the results of the networking simulation with the new switch are feasible and if we adopt the new architecture of the revised model of the Banyan switch, the hardware complexity can be reduced. The FIFO discipline has increased by about 11% when we compare the switching system with the input buffer system. We have designed and verified the switching system in VHDL using Max+pluse II. We also designed our test environment including micro computers, the base station, and the proposed architecture. We proposed a new architecture of the Banyan switch for BSDN networks and parallel computers.

1. 서 론

복잡한 통신망과 병렬컴퓨터에서는 효율적인 상호 연결을 위해 스위치가 중요한 영향을 미친다[1]. 이 스위치는 라우팅 정보에 따라서 입력포트와 출력포트 사이에 연결을 해주는 역할을 하게 된다. 따라서 스위치에 성능을 향상시키는 것은 중요하다.

기존에 사용된 $N \times N$ 크로스바(Crossbar) 스위치는 성능은 좋으나 N^2 의 교차점을 가지기 때문에 하드웨어 양이 많아지는 단점을 가지고 있다[2]. 이러

한 단점을 해결하고 구조가 간단하고 VLSI설계에 적합한 Banyan형 스위치를 사용하고 있으나 이 스위치 역시 동일한 클럭을 갖는 복수의 패킷을 전송할 때 내부 블러킹과 구조적인 블로킹이 발생한다. 이러한 블러킹으로 인한 단점을 해결하여 성능을 향상시키는 방법은 버퍼를 사용하는 방법과, 내부 스위칭 속도를 입력되는 패킷의 속도에 비해 빠르게 하는 방법이 있다.

버퍼형 스위치에는 블러킹을 해결하기 위한 방법으로 입력버퍼형 스위치와 출력버퍼형 스위치가 있는데, 입력버퍼형 스위치는 HOL(head-of-line) 블러킹으로 인하여 시스템의 성능이 낮아지는 것을

* 서남대학교 전자계산소 조교
jsh@tiger.seonam.ac.kr

말하며, HOL은 대기행렬 구조에서 발생하는 블럭킹으로 각 입력포트가 하나의 큐(Queue)를 갖는 대기행렬 구조에서 심각하게 발생한다. 입력버퍼링의 최대처리율은 랜덤 균일 트래픽일 경우 $2 - \sqrt{2} = 0.586$ 이다[3]. 출력버퍼링은 셀이 입력 단으로부터 입력되는 셀 들이 곧바로 스위치로 입력되어 출력 단 측의 출력 버퍼로 라우팅 되는 방식이다. 출력 버퍼형 스위치는 출력된 셀이 출력 트렁크로 바로 전송되지 못하는 경우 셀 들을 일시적으로 저장하는 역할을 하게된다. 따라서 HOL 블럭킹 현상은 발생하지 않으나 출력 단에서 생기는 셀 손실과 블럭킹을 방지하기 위해서는 많은 양의 버퍼가 필요로 한다.

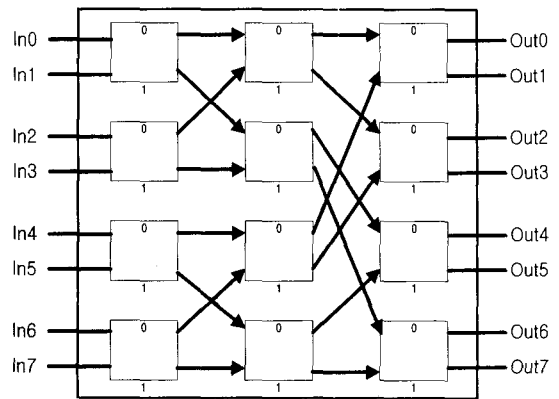
본 논문은 입력버퍼링 방법과 출력버퍼링 방식의 장점을 이용하여 하드웨어 양과 셀 손실을 효과적으로 줄일 수 있는 역 반안형 스위치에 한 행을 추가한 수정 반안망형 스위치의 구조를 제안한다. 역 반안형 스위치는 반안형 스위치를 역으로 배치한 스위치로써 입력 데이터의 도착지 주소가 같지 않은 경우에는 블럭킹이 발생하지 않는 원리를 이용하여 설계한다[4].

본 논문에서 제안한 반안 스위치는 크로스바 스위치의 $O(N^2)$ 나 배제반안 $O(\log_2 N)^2$ 보다 작은 $\log_2 N + N/2$ 개의 스위치 구조를 가진다. 여기서 N은 입력포트 수이다. 또한 처리율은 N이 무한대 일 때 반안형 0.58보다 더 좋은 효율을 갖도록 설계한다. 스위치의 구성은 각 포트마다 2개의 입력버퍼와 2개의 출력 버퍼를 사용하여 하드웨어 양과 셀 손실을 효과적으로 줄일 수 있도록 제안하였고, 제안된 반안형 스위치는 MAX+PLUS II, VHDL을 이용하여 설계 및 검증을 하고 시뮬레이션 한다.

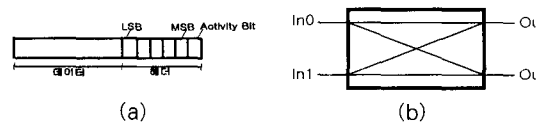
본 논문의 구성은 다음과 같다. I장은 서론, II장 반안형 스위치 구조와 동작, III장 반안 스위치 망의 성능분석, 그리고 IV장에서 결론을 맺었다.

2. 수정 반안망의 구조와 특성

위치의 구조는 그림 1과 같이 2×2스위치 소



(그림 1) 반안 스위치의 구조



(그림 2) 헤더와 반안형 스위칭 소자

자를 트리 구조로 배열하여 $\log_2 N$ 단으로 구성된 망이다. 즉 $N \times N$ 의 상호 연결 망을 $\log_2 N$ 으로 줄인 것이다[4].

이 스위치는 임의의 입력포트와 출력포트 사이에 유일한 경로를 가지고 있으므로 복수의 패킷을 전송할 때, 중간 링크를 여러 개의 입출력이 공유하므로 스위치 내부에서 셀 손실과 부하가 많이 걸린다. 각 단계는 $N/2$ 노드들이 있고 각 노드들은 2×2스위치 소자가 있다. 그림 1은 점대점 스위치에서 자기경로선택(self-routing) 알고리즘에 따라 동작하고 스위치 소자들이 독립되어 있어 확장이 간단하다

2.1 반안망 스위치의 소자

반안 스위치의 구성은 2×2기본 스위칭 소자를 다단으로 연결하여 구성하며 전체 스위치가 동기를 맞추어 동작한다. 셀 헤더와 스위칭 소자의 구조는 그림 2와 같다. 스위칭소자의 동작은 자기경로 제어방식에 의하여 이루어진다. 즉, N번째 단의 각 스위칭 소자는 셀 헤더 부분의 출력단 주소

중 최상위 비트(MSB: Most Significant Bit)로부터 N번째 비트를 조사하여 N번째 비트가 “0”이면 스위칭소자가 위쪽포트로, “1”이면 아래쪽 포트로 경로 설정하도록 구성한다.

2.2 수정 반안망의 구조

제안한 수정 반안망 구조는 그림 3과 같다. 이 방법은 역 반안망의 뒤 단에 N/2개의 스위치를 추가하는 방법으로 내부 블러킹을 해결하였다. 이 스위치의 동작은 입력 제어부에서 발생가능성을 조사하여 같은 출력 주소로 가진 2개의 셀 중에서 한 개의 셀에 1비트를 추가하여 통과시키므로 블로킹을 줄이는 방법으로 설계한다.

2.3 수정 반안망의 동작

반안 네트워크는 두 가지 블러킹 요인을 가지고 있다. 네트워크 내부의 특정 링크에서 두 개 이상의 패킷들의 충돌로 인한 패킷 분실을 일으키는 내부 블로킹과 특정 출력포트에서 두 개 이상의 패킷의 충돌로 인한 출력 블로킹이 있다.

본 논문에서는 그림 3과 같이 역 반안망(Reverse Banyan)의 뒤에 하나의 단을 추가하여 수정된 반안망을 설계하였다(그림의 Over-Dimensioned Switch를 제안하였다.). 이 방법은 역 반안망의 뒷단에 N개의 스위치를 추가하는 방법으로 내부블러킹을 해결하였다. 이 스위치의 동작은 입력 제어부에서 블로킹이 발생할 셀에 1비트를 추가하여 블로킹

을 제거하는 방법을 사용하였다. 그림 3은 수정된 반안망의 구조를 보여준다. 동작을 살펴보면 다음과 같다. N개의 셀이 입력큐에 동시에 도착하면, 입력제어 부에서는 주소 생성과, 큐에 동시 액세스 두 단계의 입력이 이루어진다. 각 셀들은 입력 큐에서 큐안에 두 개 이하의 주소만 있어야 한다. N개의 입력과 N개의 출력은 0에서 N-1까지 주소가 만들어지고 반안에는 $\log_2 N + N/2$ 2×2 스위칭소자가 있다. 역반안망에서 이루어지고 입력데이터에 관해 출력 주소가 순차적이면 비블러킹(Non Blocking)이다. 때문에 큐는 순차적이다.

스위치의 추가된 스위치들의 구조는 기존의 반안망의 스위치들과 다르게 블로킹이 발생하지 않도록 특별히 설계하였다. 즉 내부 블로킹 해결 방안은 다음과 같다.

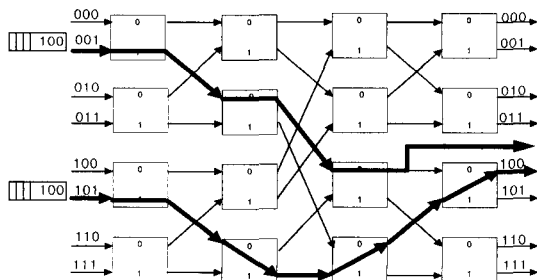
- ① 입력 조정회로에서 도착지 주소가 2개 이상이면 대기시킨다.
- ② 입력버퍼에 데이터 비트가 3비트 입력되면 4비트로 변환한다. 이때 각 포트에 “100”이나 “011”의 같은 데이터가 들어올 경우 두 개중 1개의 데이터에 데이터 앞부분에 “1”을 추가하거나 “0”을 추가한다. “010”이나 “101” 경우 두 번째 column에 “0”을 추가하고, “001” 일 경우 맨 앞에 0을 추가한다.
- ③ 입력포트의 값과 도착지 주소의 값이 같으면, 출력포트에 직접 통과하고, 같지 않으면 스위치 망을 통과한다.
- ④ 출력포트에 두 개의 패킷이 도착할 경우를 대비하여 출력버퍼의 크기를 2배로 한다.

3. 수정 반안망의 설계

수정 반안망은 VHDL코드를 Max II⁺Plus를 이용하여 설계한다.

반안 네트워크의 특성을 살펴보면 다음과 같다.

- ① 스위치 소자의 크기를 a, 스위치의 크기를 N



(그림 3) 수정 반안 스위치망

이라 할 때 스테이지 수는 $k = \log_b N$ 이 된다.

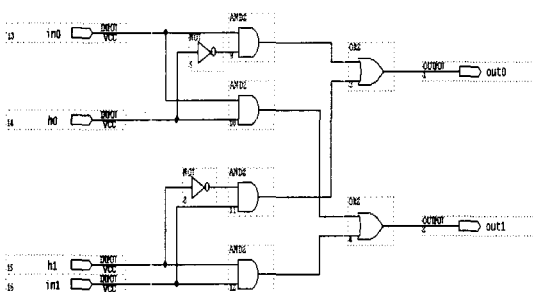
- ② 단일한 k 개의 목적지 어드레스 비트에 의한 셀프 라우팅 특성을 갖는다. 즉 입력되는 패킷에 대해서 i 번째 단의 스위치 소자는 목적지 주소의 i 번째 bit를 보고 그 값이 "0"이면 위쪽의 출력포트로, "1"이면 아래쪽의 출력 포트로 스위칭 한다.
- ③ 입력과 출력사이에 N 개의 경로를 동시에 결정할 수 있다. N 개의 입력단 모두가 N 개의 출력단 모두로 동시에, 병렬적으로 스위칭될 수 있다.
- ④ 스위치 구조자체가 모듈 형식을 취하고, 규칙적이므로 VLSI로 구현하는데 용이하다.
- ⑤ 특정입력과 출력사이에 오직 하나의 경로가 제공된다. 이러한 단일 경로 특성으로 인하여 중간단의 링크를 여러개의 입/출력이 공유하므로 중간 링크에 많은 부하가 걸리게 된다.

3.1 스위치 소자 설계

스위칭 소자는 입력포트의 내용(In0, In1)과 헤더 주소(H0,H1)와 비교하여 출력큐는 입력큐와 직접 연결되고 포트의 번호와 출력주소가 같을 때 직접 전송된다.

맨 마지막 단 스위치는 큐가 부착되어 3비트의 주소를 가지고 스위칭 했을 때 버퍼로 사용한다.

각 셀은 헤더에 라우팅 비트를 가지며 1과 0으로 이루어진 라우팅 비트를 순서 대로 읽어 나가며 스위치 된다. 즉 스위치는 셀 헤더의 일부를 보고 라우팅하고 각 단계마다 다른 비트를 보게 된다.



(그림 4) 스위치 소자 디자인

3.2 입력부 설계

(1) Input Control

- ① 포트에 입력된 출력 주소 같은 셀이 2개 초과면 대기 시킨다.
- ② 2×2 스위치 모듈에 들어오는 주소의 MSB의 주소가 서로 다른 경우만 통과한다.
- ③ 출력 주소가 2개 일 경우 두 개중 하나의 셀 주소에 왼쪽 비트와 MSB가 0일 경우 주소앞에 "0"을 추가하고, MSB가 1일 경우 주소앞에 "1"을 추가한다.

입력제어부의 전체 디자인은 그림 5에서 보는 바와 같고 같은셀 도착시 대기시키는 부분의 세부 디자인은 다음에 보이도록 한다.

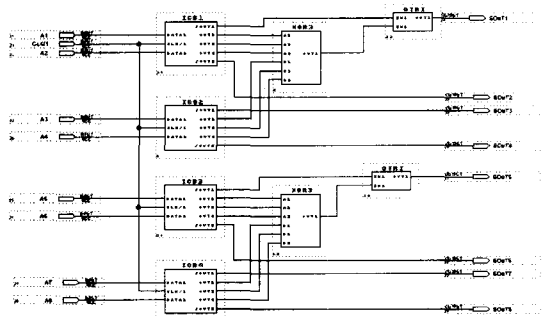
3.3 출력부 설계

(1) Output Control

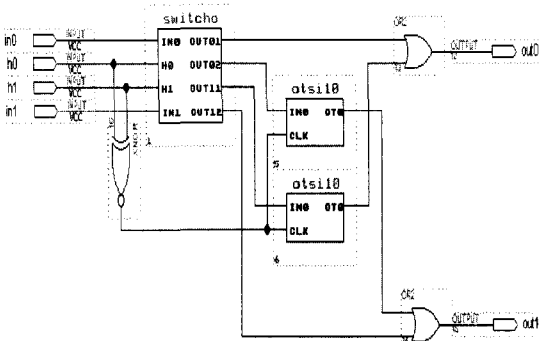
- ① 두개의 셀이 동시에 도착한 경우 1개의 셀을 FIFO 버퍼에 저장한다.
- ② 두개의 셀이 동시에 도착할 경우는 출력주소와 입력 포트의 번호와 같을 경우와 입력 포트에 목적지 주소가 같은 셀이 동시에 도착할 경우이다.

(2) Self routing 동작

- ① 입력 제어에서 생성된 주소를 이용하여 각



(그림 5) 입력제어부분 디자인



(그림 6) 출력제어 구조

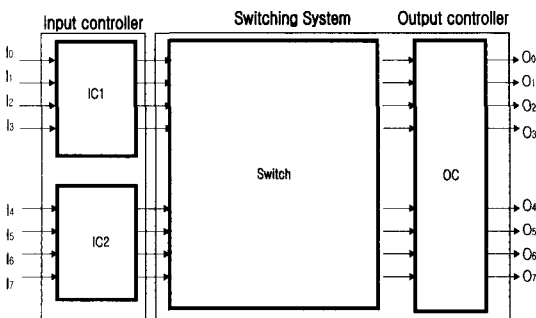
단의 스위치에 연결한다. 그림 2와 같은 스위칭소자의 선택비트에 주소의 내용을 연결하여 사용한다.

- ② 생성된 주소의 비트의 내용이 “0”이면 윗쪽으로, “1”이면 아래쪽으로 동작하여 셀을 라우팅 해준다.
- ③ 2×2 스위칭 소자에는 비블록킹으로 동작해야 하기 때문에 같은 값이 들어오지 않는다.

내부디자인 시에 비블럭킹 부분을 고려하여, 내부 디자인을 하여야 한다.

3.4 전체회로 설계

이 스위치의 전체구조는 그림과 같이 구성할 수 있다.



(그림 7) 스위치의 전체구조

내부구조를 위 그림처럼 입력제어부와 출력제

어부로 나누어 내부블럭킹을 해소하였다.

4. 반양 스위치망의 성능분석

반양형 스위치는 구성이 단순하나 동일한 출력 포트에 복수개의 패킷을 전송할 때 블럭킹이 발생한다. 스위치에 블럭킹이 발생하면 셀을 잃어버리거나 폐기시키는 문제가 발생한다. 이러한 블럭킹을 줄이기 위해서 여러 가지 방법들이 사용되고 있다[5].

이장에서는 블럭킹을 해결하기 위한 여러 가지 방법들의 성능을 컴퓨터 시뮬레이션을 통하여 고찰하고 본 논문에서 제안한 스위치 구조와 비교해본다.

4.1 버퍼형 스위치의 성능

이 방식은 입력 단에 버퍼를 두는 방법으로 스위칭 시스템 내의 연결 경로가 할당되지 않거나 일시적으로 차단되었을 경우 셀을 일시적으로 저장하기 위해서 사용한다.

N	최대 처리율
1	1.0000
2	0.7500
3	0.6825
4	0.6553
5	0.6399
6	0.6302
7	0.6234
:	:
∞	0.5858

입출력 포트 수(N)에 따른 입력 버퍼형 시스템의 최대 처리율이다. 입력버퍼형 스위치는 HOL(head-of-line)블럭킹으로 인하여 시스템의 성능이 낮아진다. HOL은 대기행렬 구조에서 발생하는 블럭킹으로 각 입력포트가 하나의 큐(Queue)를 갖는 대기행렬구조

에서 심각하게 발생한다. 입력버퍼링의 최대처리율은 랜덤 균일 트래픽일 경우 $2 - \sqrt{2} = 0.586$ 이다.

HOL블러킹을 방지하기 위해서는 스위치내의 각 입력포트에 다중큐(multiple-queue)를 두어 여러 개의 입력데이터가 하나의 출력포트로 동시에 진입하는 경우에 이를 완충(buffering)하는 방법을 이용한다. 그러나 이 방법은 하드웨어 용량이 많이 든다.

입력버퍼형 ATM교환 시스템은 조정회로가 존재해야 함에도 불구하고 다른 구조의 교환 시스템에 비해 구조가 간단하다. 또 다른 방식의 하나인 출력 버퍼형 방식은 셀이 입력 단으로부터 입력되는 셀 들이 곧바로 스위치로 입력되어 출력 단 측의 출력 버퍼로 라우팅 되는 방식이다. 출력 버퍼형 스위치는 출력된 셀이 출력 트렁크로 바로 전송되지 못하는 경우 셀 들을 일시적으로 저장하는 역할을 하게된다. 따라서 HOL블러킹 현상은 발생하지 않으나 출력 단에서 생기는 셀 손실과 블러킹을 방지하기 위해서는 많은 양의 버퍼가 필요하다.

그 반면에 본 논문에서 제안한 방법은 입력포트 셀을 동시에 2개까지 처리하고, N개의 외부 입출력 포트를 갖는 교환 시스템의 내부 교환망은 $2N \times 2N$ 으로 커진다.

하나의 외부 입출력 포트에는 b개의 내부 입출력 포트를 가지므로 같은 출력포트 주소를 갖는 셀이 한 타임 슬롯에 도착하는 경우에 k-b개의 셀은 손실된다. 유류부하 p라 할 때, 교환 시스템

에 입력되는 셀 수가 k일 확률 $\Pr[A=k]$ 은

$$\Pr[A = k] = \binom{bN}{k} \left(\frac{p}{bN}\right)^k \left(1 - \frac{p}{bN}\right)^{bN-k}$$

단, $k=0, 1, 2, \dots, 2N$ 이다.

따라서, 포트수가 N인 입력 버퍼형 교환 시스템의 셀 손실 확률 $\Pr[\text{Cell loss}] =$ 는

$$\begin{aligned} \Pr[\text{Cellloss}] &= \frac{1}{p} \sum_{k=2}^{bN} (k-1) \binom{bN}{k} \left(\frac{p}{bN}\right)^k \left(1 - \frac{p}{bN}\right)^{bN-k} \\ &= 1 - \frac{1}{p} \sum_{k=0}^{bN} (k-1) \binom{bN}{k} \left(\frac{p}{bN}\right)^k \left(1 - \frac{p}{bN}\right)^{bN-k} \end{aligned}$$

이고,

$N \rightarrow \infty$ 일 때

$$\Pr[\text{Cellloss}] = 1 - \frac{1}{p} + \frac{e^{-bp}}{bp} \sum_{k=0}^{b-1} (b-k) \frac{(bp)^k}{k!}$$

이다. 아래 표는 입출력 포트 수 (N)일 때 내부 포트 b개 일 경우와 2개 일 때의 최대 처리율을 비교하였다.

4.2 스위치의 하드웨어 양

각 스위치별 하드웨어 양을 비교해 보면 크로스바 스위치 $N \times N$, 반안 스위치 $\log_2 N$, 수정 반안 스위치는 $\log_2 N + N/2$ 으로 구성되며, 표를 비교해 보면 다음과 같다.

(표 1) 입출력 포트 수(N)와 내부포트 (b)개에 따른 최대 처리율

N	b							
	1	2	3	4	5	6	7	8
2	0.75	0.84	0.89	0.92	0.93	0.94	0.95	0.96
4	0.66	0.76	0.81	0.85	0.87	0.89	0.91	0.92
8	0.62	0.72	0.78	0.82	0.85	0.87	0.88	0.89
16	0.60	0.71	0.77	0.81	0.84	0.86	0.87	0.88
32	0.59	0.70	0.76	0.80	0.83	0.85	0.87	0.88
64	0.59	0.70	0.76	0.80	0.83	0.85	0.86	0.88
128	0.59	0.70	0.76	0.80	0.83	0.85	0.86	0.88

(표 2) 각 스위치별 하드웨어 양 비교

종류	크로스바 스위치 (N^2)	반안스위치 ($\log_2 N$)	수정 반안스위치 ($\log_2 N + N/2$)
2	4	1	2
4	16	2	4
16	256	4	12
32	1024	5	21
64	4096	6	38
128	16384	7	71

5. 결 론

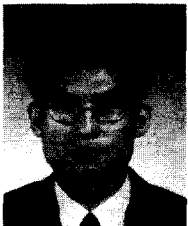
초고속 통신망이나 병렬 컴퓨터의 성능을 향상시키기 위해서는 하드웨어 양과 셀 손실이 적은 스위치의 설계가 중요한 역할을 한다.

본 논문은 기존 반안형 스위치의 성능을 향상시키기 위하여 기존 반안형 스위치의 출력부에 1 단의 특별한 스위치를 추가하였고, 입력 버퍼링 방법과 출력 버퍼링 방식의 장점을 이용하였다. 즉, 기존에 사용된 스위치를 개선하여 하드웨어 양을 줄이면서 성능을 향상시킨 스위치를 설계하였다. 본 논문에서 제안한 스위치는 컴퓨터 시뮬레이션 결과 입력버퍼형 보다 최대처리율이 11%이상 향상되었고, 다른 반안형 스위치들에 비하여 성능과 하드웨어 양을 비교하여 볼 때 좋은 결과를 얻었다. 따라서 이 스위치는 VLSI 칩으로 구현될 경우 초고속 ATM-LAN 과 병렬 컴퓨터를 개발하는데 유용하게 활용될 수 있을 것이다.

참 고 문 헌

- [1] Raif O. Onvural, Asynchronous Transfer Mode Networks: Performance Issues, 2nd Edition, Artech House, 1995.
- [2] M.J.Karol, M.G. hluchyj, and S.P. Morgan, "Input Versus Output Queing on a Space-division packet switch", IEEE Trans. commun., vol. COM.35, pp. 1347~1356, Dec. 1987.
- [3] Michael G.Hluchyj, Mark J.Kayol, "Queueing in High-Performance Packet Switching", IEEE J. select. Areas commun., vol. 6. No,9 pp. 1587~1597. Dec. 1988.
- [4] J.M.Pitts and J.A.Schormans, "Introduction to ATM Design and Performance", John Wiley & Sons, pp. 189, 1996.
- [5] Marc Boisseau, Michel Demangge and Jean-Marie Munier, "ATM Technology" : An Introduction, An International Thomson Publishing, pp. 122, 1996.

◎ 저 자 소개 ◎



조 삼 호

1995년 서남대학교 전자계산학과 졸업(학사)
 1997년 전북대학교 대학원 전산통계학과 졸업(석사)
 2002년 서남대학교 대학원 컴퓨터정보통신학과수료(박사)
 1995년~현재 : 서남대학 전자계산소 조교
 관심분야 : 네트워크설계, 초고속 통신, 무선인터넷통신, etc
 E-mail : jsh@tiger.seonam.ac.kr