

ATM 기반 MPLS 망에서 확장성을 고려한 고속 크로스바 스위치 설계 및 구현

이 동 원, 홍 현 석, 김 영 철, *최 덕 제, *이 귀 상
전남대학교 전자공학과, *전남대학교 컴퓨터정보학부
전화 : 062-530-0369 / 핸드폰 : 011-699-4516

Design and Implementation of a Scalable Fast Crossbar Switching Fabric on MPLS networks over ATM

Dong Won Lee, Hyun Suk Hong, Young Chul Kim, Deok Jai Choi, Guee Sang Lee
Dept. of Electronic Engineering, Chonnam University
E-mail : dwlee@neuron.chonnam.ac.kr

Abstract

In this paper, we propose VC merge capable hardware architecture for scalability based on ATM switching fabric. We implemented a scheduler for configuring crossbars in input-queued switches which support virtual output queues at the input ports. Also, we implemented VC merge capable scheduler at the output ports. We verified the proposed model by using C language, and designed with VHDL language. Then, we simulated and synthesized it with software of the SYNOPSIS corporation.

I. 서론

최근 인터넷의 급속한 성장과 인터넷을 통한 다양한 서비스의 증가로 인한 문제점 해결방안으로 제시된 MPLS기술과 고속라우터 방식은 서로 다른 방향에서 출발하여 서로 다른 방향으로 연구가 진행되고 있다. 그러나 향후 극복해야 할 망의 확장성, 트래픽 엔지니어링, 고속 IP 전송을 위해서는 서로간의 통합이 요구된다.

본 논문은 망의 확장성을 높이고 고속의 스위칭 기능을 제공하기 위한 고속 크로스바 스위치를 설계하였다. 입력단에서 고속라우터에서 제안된 입력버퍼링 방식의 목적지별 큐를 가지는 스케줄러를 설계하였으며, 출력단에서는 ATM 기반 MPLS망에서 망의 확장성 문제를 해결하기 위한 VC merge 기법을 수행할 수 있는 스케줄러를 설계하였다. 각 기능블록을 통합 설계하여 고속 크로스바 스위칭 패브릭을 구현하였으며, 제안한 스위치를 VHDL로 모델링하여 설계하였다.

II. 확장성을 고려한 스위치 구조

2.1 MPLS와 고속라우터 비교

표 1은 인터넷의 증가에 따른 문제점을 해결하기 위해서 제안된 고속라우터 방식과 MPLS 기술의 차이점을 보이고 있고, 현재 두 기술 모두 연구가 활발히 진행되고 있다.

향후 백본망은 고속라우터에서 채용하는 POS, IP over DWDM은 ATM과 같은 대역폭 낭비를 막아 줄 수 있으며, 인터넷 트래픽 증가로 인해 인터넷 백본이 성장하여 ATM과 같이 대역폭을 잘게 나누어 쓰지 않아도 된다. 기존의 IGP는 네트워크 자원을 효과적으로 사용하기 어려운 반면에 MPLS는 트래픽 엔지니어링을 적용해서 네트워크 자원을 효율적으로 사용할 수

* 본 논문은 두뇌한국 21 핵심분야 연구사업에 의하여 지원되었음.

있기 때문에 향후 인터넷 백분망은 MPLS기능을 탑재한 고속라우터가 중심적인 역할을 수행할 것이다.

표 1. MPLS와 고속라우터의 비교

	고속라우터	MPLS
연결	비연결형	연결형
포워딩	hop-by-hop 포워딩	레이블 스위핑
확장성	Scalable	Scalable
QoS	Not Support	Support
트래픽엔지니어링	Not Support	Support
망 구조 변경	필요 없음	필요
적용 범위	인터넷 백분망	인터넷 백분망

2.2 MPLS에서 망의 확장성을 고려한 VC merge

기존 오버레이 방식은 에지 라우터간 full-meshed 가상연결을 설정해야 하므로 망 규모가 커질 경우 가상 연결 수가 급격하게 증가하여 확장성에 제약을 받게 된다. VC Merge 기술은 MPLS영역 내에서 동일한 출구 라우터를 거쳐 전달되는 FEC들을 하나의 FEC로 묶어 전달하면 LSR에서 관리하는 레이블의 수를 감소시켜 망의 확장성을 높일 수 있다. 그림 1은 VC merge와 non VC merge의 비교를 나타낸다.

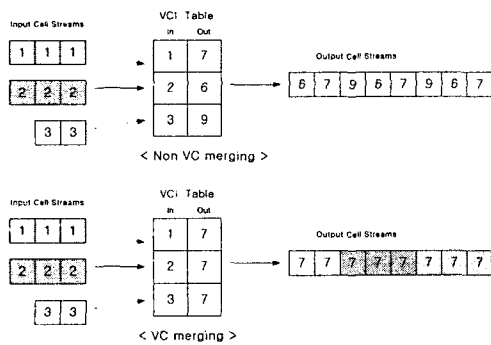


그림 1. Non VC Merge 와 VC merge 비교

그러나, ATM기반 MPLS망에서 VC merge를 적용하기 위해서는 새로운 하드웨어가 필요하다. 즉 출력단에서 셀 인터리빙이 발생하지 않도록 버퍼와 스케줄러가 새로 추가되어야 한다.

2.3 고속라우터에서 목적지별 큐잉방식

입력 버퍼링 스위치는 스위칭 패브릭의 교환 속도와 버퍼 메모리의 동작속도가 스위칭 포트 수에 상관없이

한 포트의 전송 속도와 같으면 된다. 따라서 확장성 측면에서 기존 공유 버퍼링 방식보다 유리하다.

입력 버퍼링 스위치의 세 가지 특징은 한 입력단에서 하나의 데이터만 전송 가능, 한 출력단으로 하나의 데이터만 전송 가능, 각 입력 버퍼의 선두 데이터만 전송 가능으로 인해, 출력단 충돌, 선두차단, 입력단 충돌이 발생하며 이로 인해 처리율이 제한되는 것이다.

첫 번째 이유인 출력단 충돌은 한 출력단으로 복수개의 데이터가 동시에 전달될 수 있도록 출력단을 확장해주면 해결된다. 본 논문에서는 VC merge 기법을 사용하기 위하여 출력단을 N배 확장하여 최대 처리율을 개선하였다.

두 번째 이유인 HOL 차단 현상을 개선하기 위하여 입력단의 버퍼를 출력단별로 관리하는 목적지별 큐잉방식이 제안되어, 이를 스케줄링하기 위한 알고리즘이 다수 제안되었다. 본 논문에서는 입력단 스케줄러로 iSLIP 알고리즘을 이용함으로써, HOL 차단현상을 완전히 제거하여 100%의 최대 처리율을 얻을 수 있다.

III. 제안한 고속 크로스바 스위치

MPLS와 고속라우터의 통합을 위하여 본 논문에서 제안하는 스위치의 구조를 그림 2에 나타내었다. 본 논문에서는 8x8 크로스바 스위치를 사용하였으며, 입력단은 입력버퍼링 방식의 목적지별 큐를 가지고, 이를 스케줄링하는 알고리즘은 iSLIP 알고리즘을 사용하였다. 출력단은 VC merge가 가능한 스위치의 구조로 OM(Ouput Module) 블록에 버퍼와 VC merge 스케줄러를 설계하여 통합함으로써 망의 확장성을 높였다.

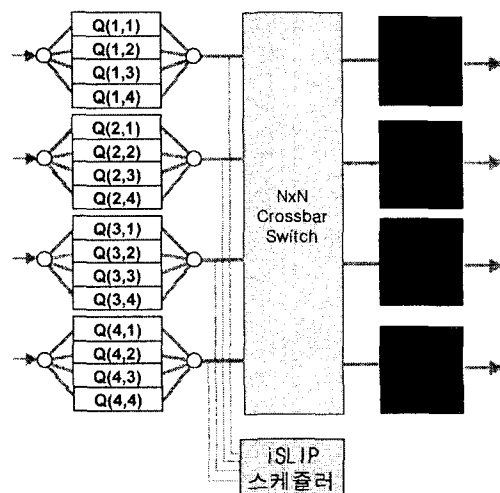


그림 2. 제안한 고속 크로스바 스위치

3.1 입력단 스케줄러

그림 3는 입력단 스케줄러의 상위 블록도를 나타낸다. iSLIP 알고리즘의 request-grant-accept 3단계가 그림 안의 세 블록들에 대응한다. 세 블록들은 각각의 아비터를 가지고 있다. 그림에서 accept 아비터에서 request 아비터로 피드백되는 신호는 이미 입력과 출력에 매치된 것으로부터 request를 mask off시키기 위하여 사용된다.

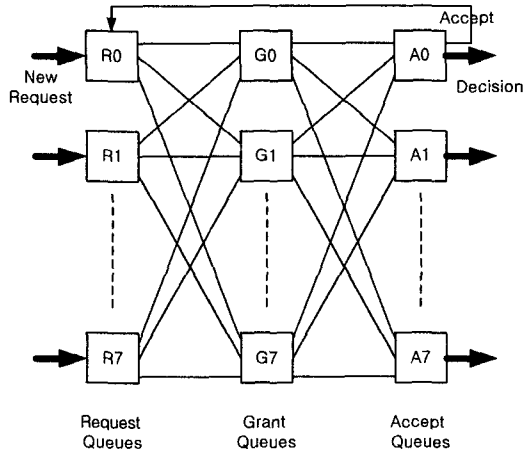


그림 3. 스케줄러의 상위 레벨 블록도

입력 아비터가 출력 아비터로부터 최소 1개의 grant를 받으면 그 입력포트는 반드시 매치됨을 알 수 있다. 따라서 accept를 받기 전에 출력 아비터들의 grant를 이용해서 request를 mask off 할 수 있다. 이 점을 이용하면 한 슬롯의 복수 iteration을 파이프라인 형식으로 구현 가능하며, 본 논문에서는 고속의 스위칭이 가능하도록 파이프 라인 구조를 이용하여 설계하였다.

스케줄러의 핵심요소 아비터는 라운드로빈 방식으로 구성하였으며, PPE(programmable priority encoder)와 상태를 저장하기 위한 라운드로빈 포인터로 구성된다. 여기서 PPE는 외부 입력으로 가장 높은 priority를 지정 가능해야 한다. PPE의 programmability를 제거하기 위하여 thermometer encoding과 simple_PE를 사용하였다. 매 중재 사이클마다 라운드로빈 포인터는 PPE를 통해 N개 입력 request에 대하여 하나를 선택하기 위하여 priority encoding을 사용한다. 포인터를 업데이트하는 메커니즘은 일반적으로 단순하고 병렬적으로 처리된다. 본 논문에서는 PPE를 따라 request 아비터에서 grant 아비터로 가는 조합회로로 구성된 path, 즉 combinational feedback loop를 제거하였다.

3.2 VC merge가 가능한 출력단

그림 2의 입력단에서 서비스를 받아 크로스바 스위치를 지난 셀들은 OM블록에서 실시간 셀과 비실시간 셀로 구분되며 비실시간 셀들은 OM블록 내부에 존재하는 VC merge 스케줄러에 의해 셀 인터리빙이 발생하지 않도록 서비스를 받는다. OM블록의 내부 구조는 그림 4와 같다. 그림 5는 제안한 VC merge 알고리즘의 순서도이며, VC merge 스케줄러의 처리 절차이다.

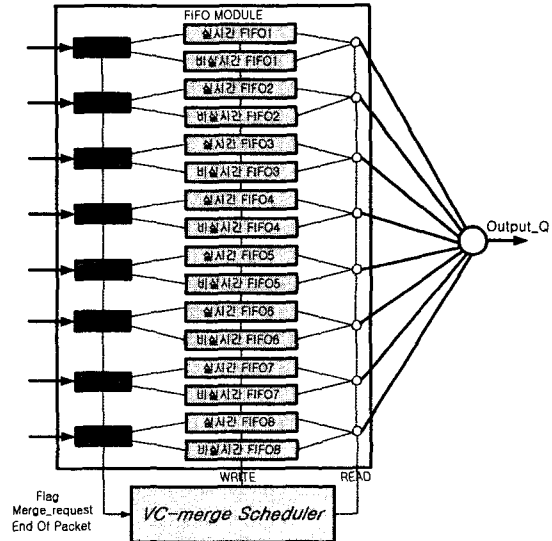


그림 4. OM(Output Module)블록의 내부구조

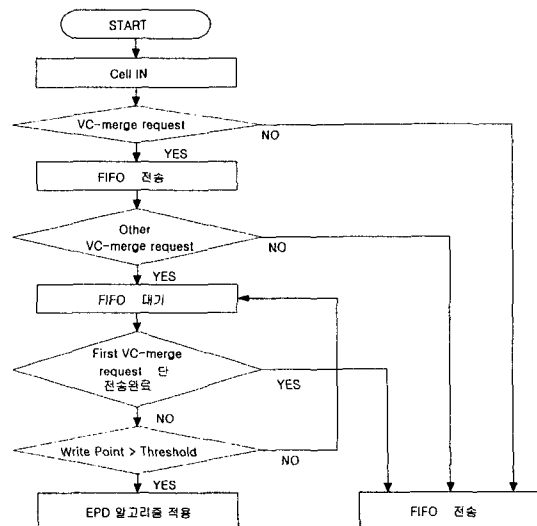


그림 5. 제안한 VC merge 알고리즘

셀의 인터리빙 문제를 해결하기 위하여 셀 헤더에서 패킷의 마지막 셀 정보가 필요하며, 이 정보를 VC merge 스케줄러에 보낸다. VC merge 스케줄러는 셀의 인터리빙이 발생하지 않도록 한다. 셀의 대기 시량의 폭주가 발생을 예방하기 위하여 EPD 방식에 의해 폭주 발생 셀들은 폐기하게 된다.

IV. 시뮬레이션 결과

iSLIP 알고리즘을 이용한 입력단 스케줄러를 하드웨어 언어기술인 VHDL 모델링을 통해 설계하였다. 그림 6은 입력단 스케줄러의 Synopsys VSS analyzer를 이용한 시뮬레이션 결과이며, 그림 7은 design analyzer를 이용하여 합성한 결과이다.

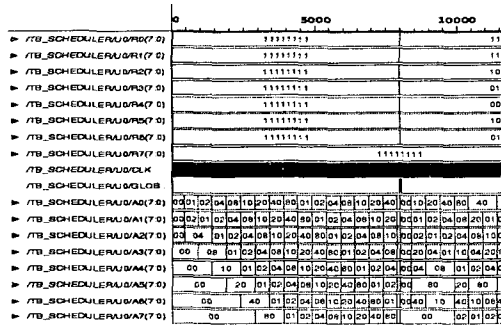


그림 6. 입력단 스케줄러의 시뮬레이션 결과

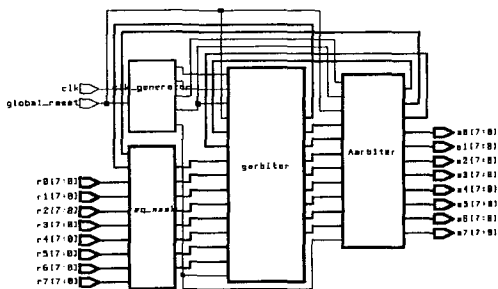


그림 7. 입력단 스케줄러의 합성결과

그림 8은 C를 통해 OM블록을 구현하여 VC merge를 사용할 경우 추가 지연시간에 대한 시뮬레이션 결과이다. 패킷의 평균 길이는 10셀로 하였고 출력링크는 ATM망의 기본 전송속도인 155 Mbps를 사용하였다. VC merge를 사용할 경우 추가지연 시간은 약 $60\mu\text{sec}$ 로 대부분의 어플리케이션에 영향을 미치지 않음을 성능 평가하였다. 그림 9는 OM블록을 VHDL로 모델링하여 시뮬레이션한 결과이다.

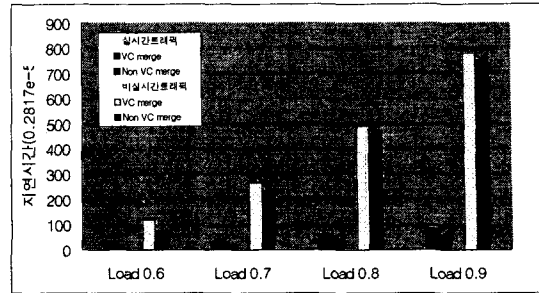


그림 8. 트래픽별 지연시간 비교

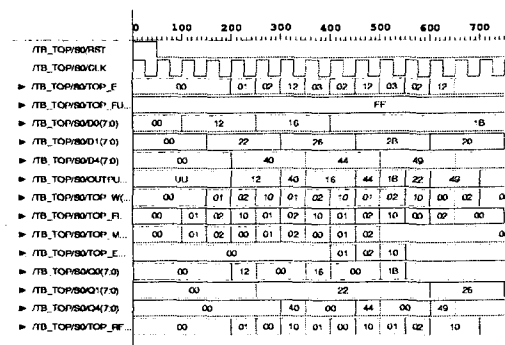


그림 9. Output Module의 시뮬레이션 결과

V. 결론

인터넷의 급속한 성장으로 인한 망의 확장성 문제를 해결하기 위하여 VC merge가 가능한 스위치를 설계하였다. 제안한 스위치는 기존의 공통 버퍼링 스위치에 비해 확장성이 우수하며, 파이프라인 구조에 따른 고속 스위칭이 가능하다. 제안된 스위치는 EPD 알고리즘을 이용하여 망의 폭주에 대비하여 백본망의 안정적 전송을 수행해 나갈 수 있을 것이다. 본 논문에서 제안한 스위치 구조는 향후 인터넷의 성장에 따른 망의 확장성 측면에서 우수한 역할이 기대된다.

참고문헌

- [1] HungKei Keith Chow, A. Leon-Garcia, "VC Merge Capable Scheduler Design" IEEE ATM'99, Japan, May 1999.
- [2] Pankaj Gupta and Nick McKeown, "Design and Implementation of a Fast Crossbar Scheduler," Hot Interconnects VI, August 1998.
- [3] 한국 전자 통신 연구원, "ATM상의 인터넷 서비스 기술개론", 진한도서, 1999년 7월