

論文 2001-38TC-12-4

ATM기반 MPLS망에서 확장성을 고려한 VC-Merge 가능한 LSR 설계에 관한 연구

(A Study on Design and Implementation of a VC-Merge
Capable LSR on MPLS over ATM)

金永哲*, 李泰元*, 李東遠*, 崔德載**, 李貴相**

(Young-Chul Kim, Tae-Won Lee, Dong-Won Lee, Deok-Jae Choi,
and Guee-Sang Lee)

요약

최근 인터넷 사용자와 인터넷을 통한 다양한 서비스들의 증가로 인한 확장성의 문제점을 해결하기 위하여 IETF(Internet Engineering Task Force)는 새로운 스위칭 기술인 MPLS(Multi-Protocol Label Switching) 기술을 소개하였다. 본 논문에서는 ATM 기반 MPLS망에서 라우터의 레이블 공간을 효율적으로 사용하여 망의 확장성을 높이기 위한 방안인 레이블 통합 기능이 탑재된 LSR(Label Switching Router)을 구현한다. 차등서비스(Differentiated Service)를 제공함에 있어서 레이블 통합 기능을 수행할 수 있는 적합한 구조를 제안하며, 망 폭주 발생 가능성이 있을 시 EPD(Early Packet Discard) 알고리즘을 통한 적응적 폭주 제어를 행함으로써 네트워크 자원의 낭비를 막고, VC-merge와 Non VC-merge 기법을 시뮬레이션을 통해 각각 비교 분석하였다. 제안한 스위치를 VHDL 모델링을 통하여 설계하고, 삼성 0.5 μ m SOG 공정을 이용하여 칩을 제작하였다.

Abstract

Recently, as Internet and its services grow rapidly, IETF(Internet Engineering Task Force) introduced a new switching mechanism, MPLS(Multi-Protocol Label Switching), to solve the problem of the scalability in Internet backbone. In this paper, we implemented the LSR loaded with VC-merging function, which causes LSR's management cost to be significantly reduced. We propose a new VC-merge function which supports differentiated services. In case of network congestion in the output buffer of each core LSR, applying link polices to the output modules of the LSR using the EPD algorithm can prevent the buffer from being overflowed. Simulation was performed for Diffserv by using multiple traffic models and investigated the impact of VC-merge method compared to non VC-merge method. The proposed switch is modeled in VHDL and fabricated using the SAMSUNG 0.5 μ m SOG process.

* 正會員, 全南大學校 電子工學科
(Department of Electronics Engineering, Chonnam
National University & RRC HECS)

** 正會員, 全南大學校 電算學科
(Department of Computer Science, Chonnam National

University)

※ 본 논문은 2000년도 두뇌한국21사업 핵심분야의 지원에 의한 연구이며 IDEC 지원의 설계 특을 사용하였음.

接受日字:2001年2月8日, 수정완료일:2001年10月24日

I. 서론

최근 인터넷의 사용자 증가 및 다양화된 응용 서비스의 출현으로 ISP(Internet Service Provider)들은 고속의 데이터 전송과 멀티미디어 서비스를 제공하기 위한 방안인 3계층 스위칭에 관하여 많은 연구를 진행하고 있다. 3계층 스위칭은 layer-2의 고속 스위칭 및 트래픽 관리가 용이하다는 장점과 layer-3의 확장성 및 융통성의 장점을 통합하여 제시된 방안이다. 현재 IETF에서는 통합 라우팅 방식인 3계층 스위칭의 표준화를 정립하기 위해서 MPLS 워킹그룹을 결성하였다^[1]. MPLS는 라우팅과 포워딩을 분리하여 라우팅 파라다임이 발전하여도 포워딩 인프라는 그대로 유지할 수 있다는 측면과 부족한 네트워크 자원을 효율적으로 이용하면서 사용자에게 서비스 품질을 보장해 줄 수 있는 트래픽 엔지니어링이 용이하다는 측면에서 차세대 인터넷 백본망 기술로 각광을 받고있다. 특히 트래픽 엔지니어링은 망 운영자의 입장에서 늘어나는 망 수요에 대한 해결책으로 중요하게 대두되고 있다.

MPLS는 기존의 IP 전송 방식인 LPM(Longest Prefix Match) 방식을 사용하지 않고 MPLS를 인식하는 인접 라우터 사이에 약속된 레이블을 사용하여 IP 트래픽의 고속 전송을 수행한다. 또한 동일 포워딩 등급(FEC : Forwarding Equivalent Class)을 갖는 플로우들에게 동일 레이블을 매핑하는 것으로 연결 관리의 확장성을 제공한다. ATM기반 MPLS망에서 각 LSR은 각 목적지의 LSP(Label Switched Path)에 대한 레이블 정보를 가지고 있다. 만약 목적지가 IP 주소 prefix라면 LSR은 라우터와 같은 크기의 레이블 테이블 정보를 갖게 된다. 그런데 망의 확장성 측면에서 망이 커질수록 LSR이 유지하는 레이블의 수를 감소시켜야 한다. MPLS 영역 내에서 동일한 출구 라우터를 거쳐 전달되는 여러 FEC들을 하나의 FEC로 묶어 전달하면 가상 연결의 수를 줄일 수 있어 확장성이 향상되고 레이블 공간을 효율적으로 사용할 수 있다. 즉 다중점 대 점 트리가 구성되며 망의 설계 시 망의 확장성 측면에서 고려되어야 한다.

본 논문에서는 현재 개발중인 ATM 기반 MPLS망에서 망의 확장성을 고려한 VC-merge가 가능한 LSR 구현에 대해 기술한다. 2장에서는 MPLS망에서 여러 가지 레이블 통합 방법에 대하여 기술하고, 3장에서는 제

안한 VC-merge ATM LSR의 구조와 알고리즘을 설명한다. 4장에서는 본 논문에서 사용하는 우선순위 제어 알고리즘을 설명하고 5장에서는 시뮬레이션 결과를 통하여 VC-merge 기법에 대한 성능 평가를 수행한다. 6장에서는 VC-merge 기법을 수행하기 위한 스위치를 설계하여 시뮬레이션과 합성 결과를 분석하였다. 마지막으로 7장에서는 결론 및 향후 과제에 대해 기술한다.

II. MPLS 망에서 VC-merge 기법

기존 공중망에서 오버레이 방식은 IP 패킷 전송을 위하여 ATM 스위치를 사용하지만, 에지 라우터간 완전 연결(full-meshed)의 가상연결을 설정해야 하므로 망 규모가 커질 경우 가상 연결 수가 급격하게 증가하여 확장성에 제약을 받게 된다. 통합 라우팅 방식인 MPLS는 레이블을 통합(merge & aggregation)하는 기능 및 레이블 스택(stack)을 사용하여 요구되는 가상 연결의 수를 감소시킬 수 있으므로 보다 개선된 확장성을 제공한다. 그림 1은 LSR의 출력단에서 레이블 통합 기능을 가진 VC-merge 기법과 non VC-merge 기법의 비교를 나타낸다. 그림과 같이 VC-merge 기법은 동일한 목적지를 가지고 있는 입력 레이블(VCI)들을 동일한 출력 레이블로 매핑하는 방법이다.

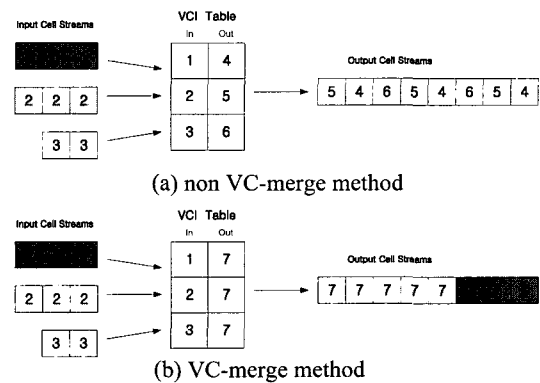


그림 1. VC-merge와 non VC-merge 기법 비교
Fig. 1. Comparison of VC-merge method and non VC-merge method.

ATM 기반 MPLS 망에서 사용하는 통합(merge)방법은 다음과 같다^[4].

가) Non VC-merge 기법

그림 1에서와 같이 스위치는 패킷의 소스와 목적지

별로 각각 다른 하나의 레이블을 할당한다. 이 방법은 full-meshed 연결 $O(n^2)$ 의 VCI 레이블을 가져야 하므로 대규모 백본망에서는 적절하지 못하다.

나) VP-merge 기법

동일한 목적지를 가지는 여러 입력 셀들은 하나의 VPI로 통합하고 입력 LSR에서 할당한 VCI값으로 패킷의 입구 LSR을 구분하는 방법이다. VP-merge 기법은 추가적인 하드웨어의 변경은 필요로 하지 않지만 VPI 공간이 4096개로 제한되어 있기 때문에 사용할 수 있는 레이블 공간의 제약을 받는다. 또한 VCI의 16비트를 사용하여 입구 LSR을 표시하여야 하는 별도의 관리 과제가 부과된다.

다) VC-merge 기법

그림 1에서와 같이 동일한 목적지를 가지고 있는 입력 VCI 레이블들을 동일한 출력 레이블로 매핑하는 방법이다. 이 방법은 VP-merge와 같은 레이블 공간 제약 문제는 없지만, 셀 단위로 나누어진 패킷을 단순히 merge를 수행하게 되면 패킷들이 뒤섞이는 셀 인터리빙이 발생하여, 본래의 패킷으로 복원할 수 없게 된다. 이러한 셀 인터리빙 문제를 해결하기 위해서는 기존의 ATM 교환기에 추가적인 하드웨어가 필요하므로 지연이 증가한다는 단점을 가진다.

VC-merge를 수행할 때 셀 인터리빙 문제에 대한 해결방안으로는 AAL5대신에 AAL3/4의 MID(Message Identifier)필드를 이용하는 것이다. 그러나 MID의 10비트 크기가 모든 송신자를 나타낼 수 있을 만큼 충분하지 않기 때문에 인터넷 서비스에서는 AAL5가 널리 쓰이고 있다. 다른 방법은 하나의 패킷에 속한 전체 셀들을 버퍼에 저장하였다가 연속적으로 전달함으로써 패킷단위의 셀 인터리빙이 이루어지도록 하는 방법이다. 이외에도 AAL5 extension과 같은 프로토콜을 수정하여 셀 인터리빙을 방지하는 방법 등이 제안되고 있다.

III. 제안한 ATM LSR의 구조 및 알고리즘

VC-merge 기법을 수행할 수 있는 ATM LSR의 구조로 성능평가를 위하여 제안한 구조는 그림 2와 같다.

그림 2의 구조는 스위치에서 블록킹이 발생하지 않는 공유버퍼형 셀 스위칭 구조이며, 셀 단위 스위칭을 이용하기 위해서 출력측에 출력 포트수와 동일한 개수의 OMs(Output Modules)로 구성되어 있다.

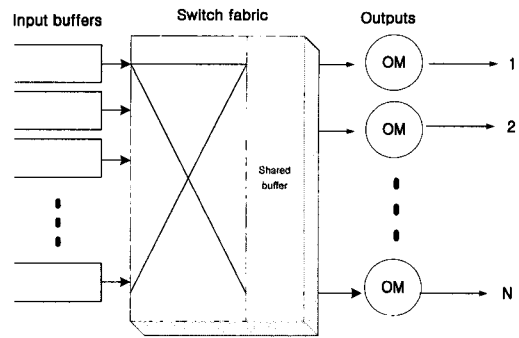


그림 2. VC-merge 가능한 ATM LSR
Fig. 2. VC-merge capable ATM LSR.

그림 3은 제안한 OM의 구조를 보여주고 있다. 여기서 하나의 OM 블록은 셀 인터리빙이 발생하지 않도록 각각의 입력포트별로 셀을 저장하는 RBs(Reassembly Buffer)와 트래픽 특성에 따라 스케줄링을 수행하기 위한 실시간/비실시간 FIFO로 구성된 OB(Output Buffer)로 이루어져 있다. 만약 RBs가 없다면 스위치 구조는 VC-merge를 수행하지 않는 ATM LSR의 구조와 동일하다. 손실 우선순위 제어를 위해서 동적 임계치를 갖는 부분 버퍼공유기법을 사용하였고, 지연 우선순위 제어를 위해서 QLT(Queue Length Threshold) 제어 알고리즘을 사용하였다. 또한 전단의 RB는 패킷 단위의 셀을 저장하는 역할을 수행한다.

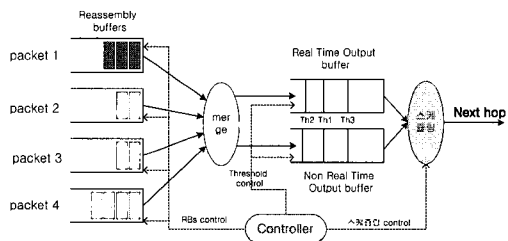


그림 3. 제안한 OM(Output Module)의 구조
Fig. 3. The structure of the proposed OM.

OM은 두 개의 동적임계치를 갖는 버퍼 두 개로 구현하였으며 하나는 실시간 트래픽을 처리하기 위한 버퍼이며, 다른 하나는 비실시간 트래픽을 처리하기 위한 버퍼이다. 폭주된 망을 통과한 셀(FCI bit=1)들이 입력되면 망의 폭주가 예상되므로 가변폭(m)만큼 Th1, Th2를 FCI가 1인 셀이 하나 도착할 때마다 하나씩 감소시켜서 미리 예방하는 효과를 갖고, 만약 FCI가 0인 셀들이 M개 만큼 계속 입력될 경우 Th를 하나씩 다시 증

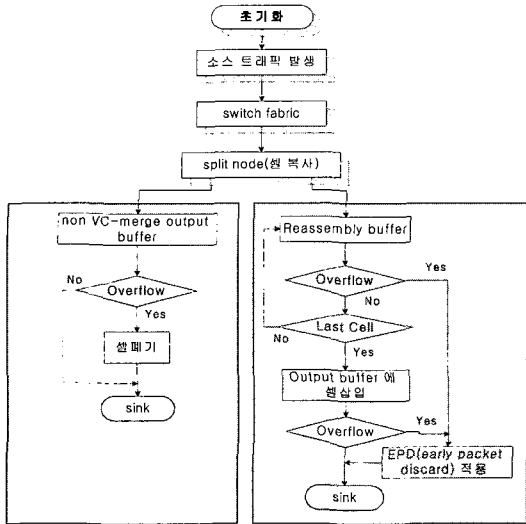


그림 4. OM 블록에서 트래픽 처리 절차
Fig. 4. Procedure of traffic processing in OM.

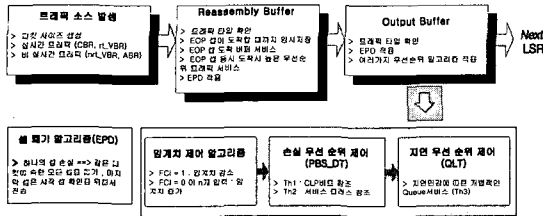


그림 5. 제안한 OM에서 데이터 흐름도
Fig. 5. Data flow in the proposed OM.

표 1. 입력 트래픽 파라미터

Table 1. Input traffic parameter.

구분	내용	값
L	평균 서비스율 (cells/sec)	355,000(약 150Mbps)
A	평균 셀 도착률 (cells/sec)	CBR(110,000), rt-VBR(80,000), nrt-VBR(125,000), ABR(40,000)
p	부하(p = A/L)	0.4 ~ 1.0
c	폭주발생률 (FCI=1인 셀수/전체 도착 셀수)	$10^{-3} \sim 10^{-2}$

가시킨다.

본 논문에서 제안한 시뮬레이션 구조의 데이터 처리 절차는 그림 4와 같다. 제안한 구조는 VC-merge 기법과 non VC-merge 기법을 비교 분석하기 위해서 크게

VC-merge 모듈과 Non VC-merge 모듈로 구성하였다. 또한 MPLS 망에서 VC-merge 기법의 수행은 망의 중간 라우터인 코어 LSR에서 이루어지기 때문에 단순한 스케줄링 알고리즘을 사용하여 최대의 효과를 얻을 수 있도록 하였다.

그림 5는 OM의 데이터 흐름을 구체적으로 설명하고 있다. 입력 트래픽으로는 실시간 트래픽(CBR, rt-VBR) 및 비실시간 트래픽(nrt-VBR, ABR)을 구분하여 발생시켰으며 트래픽 파라미터는 표 1과 같다.

트래픽의 패킷 사이즈는 그림 6에서와 같이 실제 WAN에서 패킷 사이즈 분포를 고려하여 평균 패킷 길이를 10셀로 하였다^[4]. 셀 인터리빙이 발생하지 않도록 주어진 패킷에 대한 각 입력 셀들은 패킷의 마지막 셀이 도착할 때까지 RB에 저장되어야 한다. 셀 헤더의 EOP(End of Packet)에 의하여 마지막 셀이 도착하였을 때, 비로소 RB에 있는 모든 셀들은 다음 흐름으로 전달되기 위하여 출력 버퍼로 전달되고 동시에 입력 VCI를 출력 VCI로 변환된다. 출력 버퍼는 망의 폭주 경험 여부를 나타내는 셀의 FCI(Flow Congestion Indicator) 비트에 의해서 출력버퍼의 임계치를 조정하게 된다. 출력버퍼는 손실 우선순위 제어를 수행하기 위한 두개의 임계치인 Th1, Th2를 가지고 있으며 셀이 도착하였을 때 임계치 Th1 이하라면 모든 셀들을 받아들이고 버퍼에 셀 수가 Th1을 넘게 되면 CLP(Cell Loss Priority) 비트가 0인 셀만 받아들인다. 계속 셀이 증가하여 Th2를 초과하게 되면 실시간 셀들만 받아들이고 비실시간 셀들은 폐기한다. 만약에 버퍼가 가득 차게 되면 도착하는 모든 셀들은 폐기한다. 지연 우선순위 제어는 Th3에 의해서 수행하는데 두개의 출력 버퍼가 같은 상태이면 번갈아 가면서 서비스를 하고 한쪽이 Th3보다 크다면 그 쪽 버퍼를 서비스한다. 여기에 사용된 셀 폐기 알고리즘으로는 하나의 셀이 폐기되었을 경우 잇따르는 패킷 전체를 폐기하는 EPD알고리즘을 사용한다. AAL5를 이용한 데이터는 셀 손실이 발생하였을 경우 그 셀이 속한 패킷 전체를 사용할 수 없기 때문이다. 반면, PPD알고리즘은 이미 전송된 패킷은 목적지에서 폐기되고 나머지 패킷들은 버퍼에 입력되기 전에 폐기하는 방법이다. 본 논문에서 사용된 EPD 방법은 PPD 방법과는 다른 방법으로 다음과 같은 처리과정을 수행한다. 먼저 AAL5에서 AUU=0는 상위계층 데이터 단위(CPCS-PDU)가 SAR 계층에서 셀로 나뉘어질 때 첫 번째 혹은 중간 셀을 나타내고 AUU=1은 그 프레임의

마지막 셀을 나타내는데 AAL은 하나의 VC에서 동시에 여러 패킷의 다중화를 지원하지 않기 때문에 AUU 파라미터는 프레임의 경계를 판별하는데 사용될 수 있다. ATM 셀의 첫 번째 셀을 구분하기 위해서 AAU=1인 셀을 폐기하지 않는다.

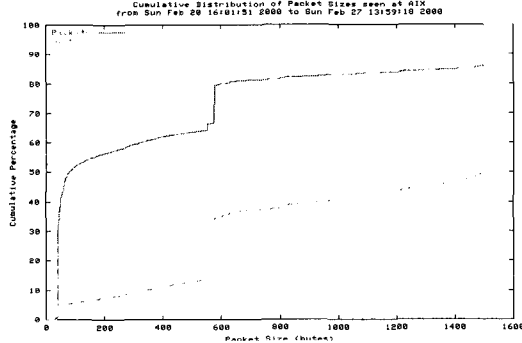


그림 6. IP 패킷 길이 분포도
Fig. 6. IP packet length distribution.

IV. 우선순위 제어 알고리즘

스케줄링 방법으로써 우선 순위 제어방법은 크게 손실 우선순위 제어방법과 지연 우선순위 제어방법으로 구분할 수 있다. 본 논문에서는 출력버퍼에 QLT 지연 우선순위 제어방법과 동적 임계치를 갖는 PBS 손실 우선순위 제어알고리즘을 제안하였다.

지연 우선순위 제어방법은 우선 순위에 따라 실시간 전송을 요구하는 트래픽을 먼저 전송함으로써 큐잉지연을 줄이는 방법이다. HOL(Head Of Line), HOL-PJ(HOL with Priority Jump), MLT(Minimum Laxity Threshold), QLT(Queue Length Threshold) 알고리즘이 있지만 큐내의 비실시간 셀의 개수가 임계치를 넘으면 비실시간 셀에 우선 순위를 부여하고 지연 민감 정도에 따라 2개의 클래스로 나누고 각각의 큐는 개별적인 임계값을 둘 수 있다는 특징 때문에 QLT 알고리즘을 사용하였다.

또한 손실 우선순위 제어방법은 셀을 폐기 해야할 경우 손실 우선 순위가 낮은 셀을 선택적으로 폐기함으로써 손실 우선 순위가 높은 셀손실을 줄여 큐를 효율적으로 이용하는 방법으로 다음과 같은 방법들이 있다.

가) CLP 만을 고려한 Push-out 기법

이 모델은 셀 헤더 내의 CLP 비트를 참조하여 손실 우선 순위 제어를 하는 모델로 하나의 임계치를 갖는

단일 버퍼를 가진다. 셀이 도착하였을 때 현재 버퍼 내의 셀 수가 임계치를 넘게 되면 그 이후에는 CLP 비트가 1인 손실 우선순위가 낮은 셀을 폐기하고 CLP 비트가 0인 셀만을 버퍼에 저장하게 된다. 만약에 버퍼가 가득 차게되면 도착하는 모든 셀들은 폐기된다.

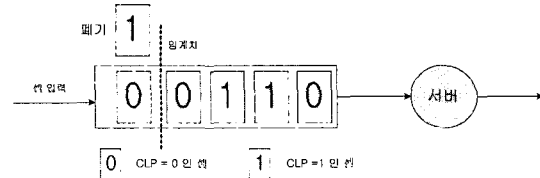


그림 7. Push-out 우선순위 제어
Fig. 7. Push-out priority control.

CLP 모델은 단순한 제어 구조를 가지므로 구현이 비교적 용이하지만, CLP 비트만을 이용한 2단계의 우선 순위 제어를 수행함으로써 손실 및 지연에 대한 다양한 QoS 등급의 트래픽을 고려할 경우 각 등급의 트래픽에 대해서 적절한 서비스를 해줄 수 없으므로 QoS를 보장하지 못하는 문제점이 있다.

나) PBS(Partial Buffer Sharing)

클래스별 입력 트래픽이 구분이 되어 버퍼에 도착하고 임계치를 넘으면 높은 우선 순위의 셀만 받아들인다.

V. 시뮬레이션 결과

시뮬레이션에 사용된 입력 트래픽으로 CBR은 일정한 비율로 셀을 발생시키고 VBR과 ABR트래픽은 버스티한 특성을 가장 잘 나타내는 On-Off모델을 사용하였다. 시뮬레이션은 C++를 사용하였고 5초 동안의 과도기 상태를 거친 후 100초 동안 수행하였다.

셀 손실율은 하나 이상의 소스에서 발생한 전체 셀 수에 대한 손실된 셀 수의 비율로서 다음과 같이 정의할 수 있다.

$$\text{셀손실율}(\%) = \frac{\text{손실셀수}}{\text{전체전송셀수}} \times 100 \quad (1)$$

셀 지연시간은 RB에 입력한 시간부터 출력버퍼를 떠날 때까지의 시간으로 정의한다. 단, 버퍼에서 하나의 셀이 서비스되는 시간은 약 $0.2817e \times 10^{-5}$ 초를 기본단위로 측정하였다.

그림 8과 9는 실시간 트래픽과 비 실시간 트래픽에

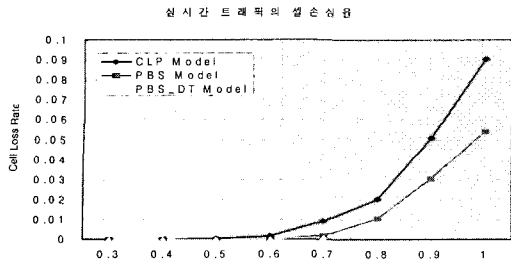


그림 8. 우선순위 제어별 실시간 트래픽의 셀 손실율
Fig. 8. Cell loss rates of real-time traffics with different priority control.

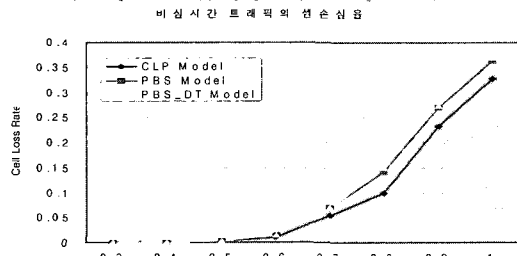


그림 9. 우선순위 제어별 비실시간 트래픽의 셀 손실율
Fig. 9. Cell loss rates of non real-time traffics with different priority control.

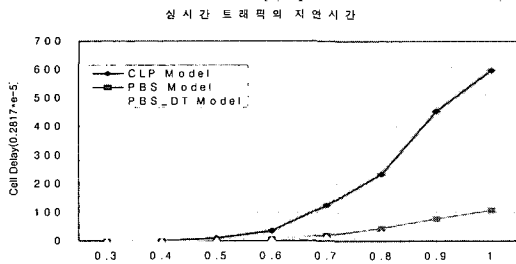


그림 10. 우선순위 제어별 실시간 트래픽의 지연시간
Fig. 10. Delays of real-time traffics with different priority control.

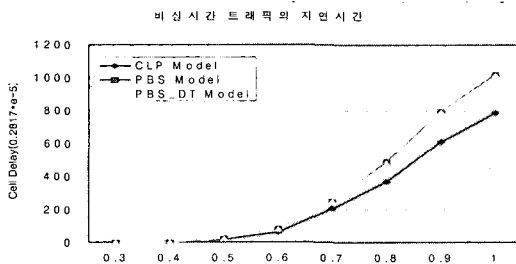


그림 11. 우선순위 제어별 비실시간 트래픽의 지연시간
Fig. 11. Delays of non real-time traffics with different priority control.

대한 셀 손실율을 보여주는 그래프이다. 실시간 트래픽은 CLP만을 고려한 모델의 경우 최악의 성능을 보여주었고, PBS_DT 모델의 경우 동적 임계치를 조정함으로써 실시간 트래픽의 성능이 개선됨을 보여주고 있다. 그러나 임계치 조정으로 인해서 비 실시간 트래픽의 셀 손실율은 CLP만을 적용한 것보다 약간 증가하였지만 PBS 모델과 비슷한 성능을 보였다.

그림 10, 11은 지연 우선순위 제어방법의 종류에 따른 지연시간을 비교한 그래프이다. 지연 우선순위 제어를 수행한 PBS_DT 모델의 시뮬레이션 결과, 다른 모델에 비해서 실시간 트래픽의 성능이 크게 향상되었고 비 실시간 트래픽의 성능은 감소하였다. 따라서 추가적인 지연시간이 요구되는 VC-merge 기법의 경우 PBS_DT 모델을 사용하여 실시간 트래픽의 지연시간을 줄일 수 있도록 설계하는게 바람직하다.

그림 12는 VC-merge에 따른 트래픽별 지연시간의 영향을 나타낸 그래프이다. 지연 우선순위 제어로 인한 실시간 트래픽의 성능이 비 실시간 트래픽의 성능에 비교해서 성능이 향상됨을 볼 수 있다. 그러나 VC-merge는 추가적인 버퍼링을 필요로 하기 때문에 VC-merge를 사용하지 않는 경우보다 실시간 트래픽은 평균 15셀, 비실시간 트래픽은 평균 20셀의 추가지연시간이 요구된다. OC-3의 속도를 가진 ATM 링크에서 20셀 지연은 약 60 μ sec와 비슷하므로 추가 지연시간은 대부분의 어플리케이션에서는 그렇게 중요하게 작용하지 않을 것으로 예상된다.

그림 13과 14는 VC-merge에 따른 트래픽별 셀 손실율을 보여주고 있다. 부하가 작을 때는 셀 손실율이 거의 발생하지 않지만 부하가 증가함에 따라서 추가 버퍼링에 의해서 셀 손실율이 감소하게 된다. 실시간 트래픽의 경우 부하가 0.9일 때 약 20%의 성능향상을 보

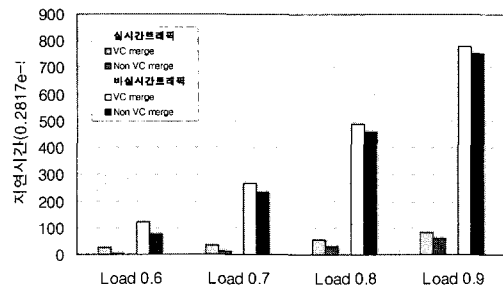


그림 12. VC-merge에 따른 트래픽 지연시간
Fig. 12. Average packet delays with VC-merge.

이고 있으며, 비 실시간 트래픽의 경우 약 12%의 손실률이 개선됨을 볼 수 있다. 그러나 셀 손실은 감소하지만 추가버퍼로 인한 지연이 증가하게 된다. 이러한 추가 지연은 대부분의 어플리케이션 성능에서 영향을 끼치는 않으며, 따라서 VC merge 방식은 label의 수를 줄임으로써 망의 확장성을 확보할 수 있다.

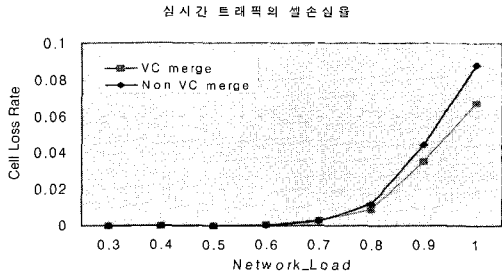


그림 13. VC merge에 따른 실시간 트래픽의 손실률
Fig. 13. Cell loss rates of real-time traffics with VC-merge.

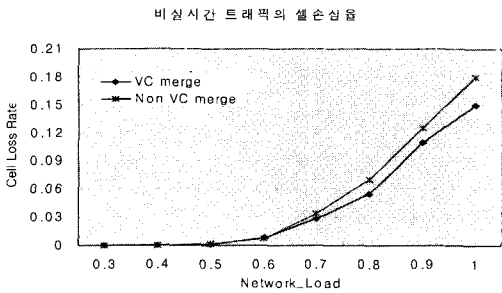


그림 14. VC merge에 따른 비실시간 트래픽의 손실률
Fig. 14. Cell loss rates of non real-time traffics with VC-merge.

VI. VC merge 가능한 스위치 설계

VC-merge 기법을 적용하기 위해서는 셀 인터리빙 문제를 해결해야 한다. 따라서 ATM 스위치에 새로운 하드웨어가 추가되어야 한다. C++을 이용하여 검증된 결과를 기반으로 VC-merge 기법을 지원하는 스위치를 구현하기 위하여, 하드웨어 기술 언어인 VHDL을 이용하여 모델링하였으며, Synopsys 툴을 이용하여 시뮬레이션 및 합성하였다[8]. VC-merge 기능을 가지는 크로스바 스위치의 구조는 그림 15와 같다.

본 논문에서 제안한 스위치는 8×8 크로스바 스위칭 패브릭을 사용하였으며, 입력단에서는 스위치의 확장성

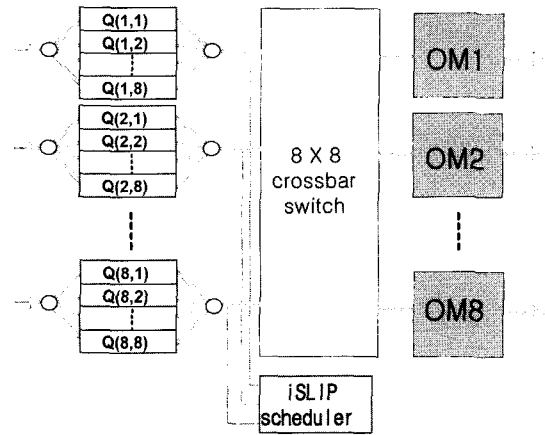


그림 15. VC-merge 가능한 고속 크로스바 스위치
Fig. 15. VC-merge capable a fast crossbar switch.

을 고려하여 입력버퍼링 방식의 목적지별 큐를 두어 HOL블록킹을 방지하고, 이를 스케줄링하는 알고리즘은 iSLIP 알고리즘을 사용하였다. 출력단에서는 OM에 RB와 VC-merge 스케줄러를 설계하여 입출력단을 통합하였다.

입력 버퍼링 스위치는 스위칭 패브릭의 교환 속도와 버퍼 메모리의 동작속도가 스위칭 포트수에 상관없이 한 포트의 전송속도와 같으면 되므로, 확장성 측면에서 장점을 지니고 있다. 그러나 입력 버퍼링 스위치는 HOL차단 현상으로 인해 최대 처리율이 58%로 저하된다는 단점을 가지고 있다. 이러한 HOL차단 현상을 제거하는 방법으로 입력단의 버퍼를 출력단별로 관리하는 목적지별 큐잉 방식이 제안되었다. 그림 16은 입력단에서 목적지별 큐잉 방식 버퍼의 구성도를 나타낸다 여기서 i 의 값은 입력포트의 번호이며, j 는 출력포트 번호를 나타낸다.

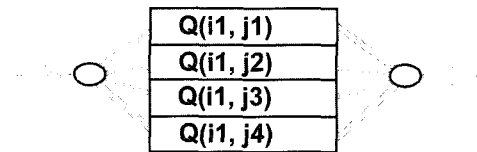


그림 16. 목적지별 큐의 구성도
Fig. 16. Diagram of virtual output queue.

이러한 목적지별 큐를 스케줄링하기 위한 알고리즘으로는 PIM(Parallel Iterative Matching) 알고리즘과 RRM(Round-Robin Matching) 알고리즘 등 다수 제안되었다[6]. PIM 알고리즘은 랜덤 선택으로 인한 회로의

복잡성과 공정성이 결여되었으며, RRM은 회로의 동기화 문제점을 가지고 있다. 본 논문에서는 PIM과 RRM의 단점을 극복한 iSLIP 알고리즘을 이용한다. iSLIP 알고리즘은 PIM과 마찬가지로 request, grant, accept의 세 단계를 반복적으로 행함으로써 입력과 출력 사이의 최대 매칭을 찾아내는 방법이다. 그림 17은 iSLIP 스케줄러의 구성도를 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. iSLIP 알고리즘은 PIM의 랜덤선택과 달리 반복 라운드로빈 방식으로 포인터를 하나씩 증가시켜 최대매칭을 찾아낸다. 이러한 방식을 통하여 대역폭을 동등하고 공평하게 사용할 수 있다. 또한 랜덤으로 도착하는 트래픽에 대하여 선택된 grant에 해당하는 포인터들만 라운드로빈 방식으로 하나씩 증가시킴으로써 RRM이 가지고 있던 포인터들 사이의 동기화 문제를 해결하였다. 이러한 iSLIP 알고리즘은 HOL차단 현상을 완전히 제거하여 100%의 최대 처리율을 얻을 수 있다.

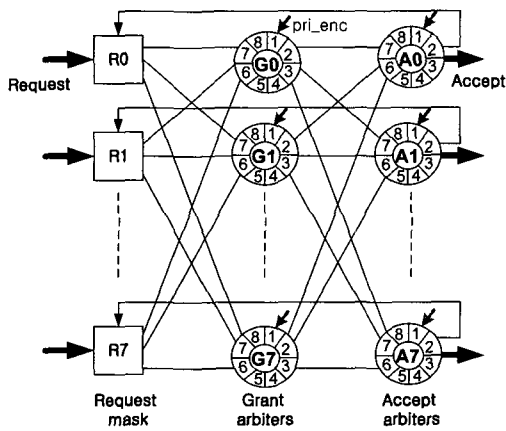


그림 17. iSLIP 스케줄러의 다이어그램
Fig. 17. Diagram of iSLIP scheduler.

그림 18은 iSLIP 입력단 스케줄러의 시뮬레이션 결과이다. 입력 값으로 모든 포트에서 request를 보냈을 때, 출력 값으로 accept된 A0 ~ A7의 값을 보면 HOL 차단 현상이 발생하지 않고 출력됨을 볼 수 있다.

그림 19는 제안한 OM을 VHDL로 기술한 후 시뮬레이션한 결과이다. 포트 D0에서 merge를 먼저 요청하고

포트 D1에서는 merge를 포트 D0의 요청 후에 요구하고 있으며, 포트 D4는 merge를 요구하지 않는 패킷이 들어왔을 때의 시뮬레이션 결과를 보이고 있다. 시뮬레이션 결과를 살펴보면 포트 D0의 패킷의 마지막 셀이 들어 올 때까지 포트 D1의 패킷은 버퍼에 대기하게 된다. Merge를 요구하지 않는 포트 D4의 패킷은 셀 인터리빙 문제가 발생하지 않으므로 버퍼에 대기하지 않고 서비스됨을 볼 수 있다.

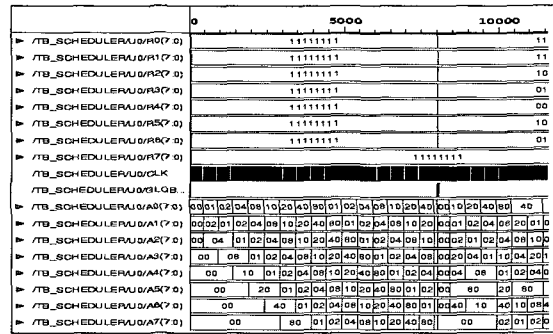


그림 18. iSLIP 스케줄러의 시뮬레이션 결과
Fig. 18. Simulation result of iSLIP scheduler.

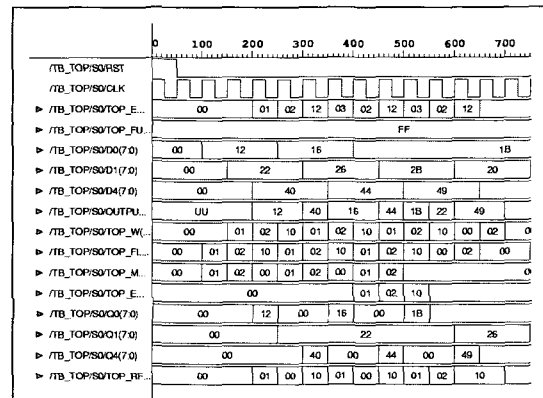


그림 19. 제안한 OM의 시뮬레이션 결과
Fig. 19. Simulation result of the proposed OM.

그림 20은 시뮬레이션을 통하여 검증된 스위치를 삼성 0.5 μ m SOG공정 KG-80 라이브러리를 이용하여 합성한 결과이다. 칩은 약 170,000 게이트 수로 합성되었으며, 현재 반도체설계교육센터(IDEC) 지원의 MPW 칩으로 제작되었다.

VII. 결 론

본 논문에서는 MPLS 망에서 DiffServ를 고려하여

실시간 트래픽의 QoS를 보장할 수 있는 레이블 통합방안을 제안하였다. 또한 VC-merge 기법을 수행하기 위해서는 새로운 하드웨어의 추가가 필요하기 때문에 VC-merge 기법을 수행할 수 있는 스위치를 VHDL로 기술하여 하드웨어로 구현하였다. 제안한 우선순위 제어 기법을 통하여 실시간 트래픽의 성능이 향상되었고 VC-merge 기법을 이용한 추가 버퍼링은 셀 손실율을 줄였으며, 또한 EPD 알고리즘을 사용하여 전체적인 전송효율이 개선되었다. 버퍼링으로 인한 추가 지연시간은 대부분의 어플리케이션에 영향을 미치지 않음을 볼 수 있었고, 출구 LSR에서의 셀을 재결합해야 하는 시간을 줄일 수 있다.

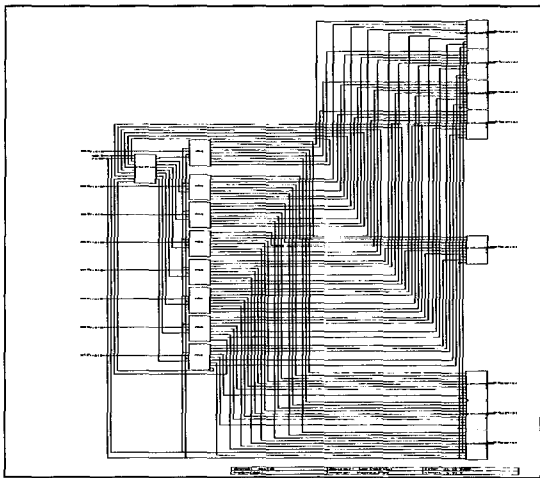


그림 20. 제안한 스위치의 합성 결과
Fig. 20. Synthesis result of the proposed switch.

DiffServ를 지원하는 MPLS 망에서 고속 스위칭이 가능한 VC-merge 모듈 구현은 반드시 고려되어야 하며, 제안한 구조는 향후 인터넷의 급속한 성장에 따른 망의 확장성측면에서 효율적이고 우수한 역할이 기대된다.

본 논문에서는 스위치 Fabric 쪽의 공유 버퍼보다는 VC merge가 가능한 OM의 구현을 목적으로 하였기 때문에 공유버퍼형 스위치 패브릭을 가정하였으며, 주소 번역 모듈을 통해 Reassembly Buffer와 OM으로의 셀 전달 체계가 이미 갖추어져 있다고 가정하였다. 따라서 앞으로 스위치 Fabric의 메모리 구조와 그 주소 전달 체계에 대한 연구를 본 연구와 병합하여 진행할 예정이다.

참 고 문 헌

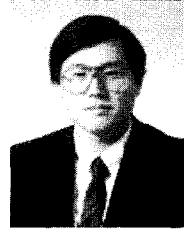
- [1] R. Callon, P. Doolan, N. Feldman, "A Framework for Multiprotocol Label Switching", Internet Draft <draft-ietf-mpls-framework-05.txt>, September, 1999.
- [2] Loa Andersson, Paul Doolan, Nancy Feldman, Andre Fredette, Bob Thoma, "LDP Specification", <draft-ietf-mpls-ldp-06.txt>, October, 1999.
- [3] I.Widjaja, M. Neuts, and J.M.Li, "Conditional overflow probability and profile curve for congestion detection in ATM networks", in Proc. INFOCOM'96, San Francisco, CA.
- [4] WAN Packet Size Distribution. [Online]. Available WWW : <http://www.nlanr.net/NA/Leam>.
- [5] Nick McKeown, "iSLIP: A Scheduling Algorithm for Input-Queued Switches" IEEE Transactions on Networking, Vol 7, No.2, April, 1999.
- [6] HungKei Keith Chow, A. Leon-Garcia, "VC Merge Capable Scheduler Design" IEEE ATM'99, Japan, May, 1999.
- [7] Pankaj Gupta and Nick McKeown, "Design and Implementation of a Fast Crossbar Scheduler," Hot Interconnects VI, August, 1998.
- [8] K.C. Chang, "Digital Systems Design with VHDL and Synthesis", IEEE Computer Society Press, 1999.
- [9] 한국 전자 통신 연구원, "ATM상의 인터넷 서비스 기술개론", 진한도서, 1999년 7월
- [10] 손병찬, "ATM 망에서 동적 임계치를 갖는 Early Packet Discard 방식의 성능 분석", 경북대학교 대학원 전자공학과 석사학위 논문, 1996년 2월
- [11] 신진서, "MPLS를 이용한 고속 IP Network 구축 방안", 미디어 링크, 2000년 6월

저 자 소 개



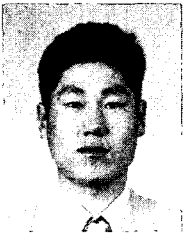
金永哲(正會員)

1981년 : 한양대학교 전자공학과 공학사. 1987년 : University of Detroit, EE, 공학석사. 1993년 : Michigan State University, EE, 공학박사. 1993~현재 : 전남대학교 전자공학과 부교수. 2000~현재 : 전남대학교 반도체설계교육센터 소장 <주관심분야> 초고속통신망, 인터넷 응용, 회로설계, 보안 칩 개발



李貴相(正會員)

1980년 : 서울대 공대 전기공학과 졸업(학사). 1982년 : 서울대 대학원 전자계산기공학과 석사. 1982년 : 금성통신 연구소 근무. 1991년 : Pennsylvania 주립대학 전산학과 박사. 1984년~현재 : 전남대 전산학과·정보통신 연구소 교수. <주관심분야> 멀티미디어통신, 영상처리 및 복원, 테스트, 논리합성, VLSI/CAD



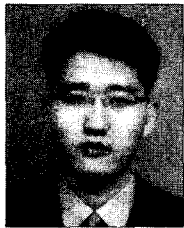
李泰元(學生會員)

1993년 : 전남대학교 전자공학과 공학사. 1999년 : 전남대학교 전자공학과 공학석사. 1999년~현재 : 전남대학교 전자공학과 박사과정. <주관심분야> 초고속통신망, 회로설계



崔德載(正會員)

1982년 : 서울대학교 컴퓨터공학과 공학사. 1984년 : KAIST 전산학과 석사. 1995년 : Univ. of Missouri, 공학박사. 1996년~2001년 : 전남대학교 전산학과 부교수. 2001년~현재 : 전남대학교 전산학과 교수. <주관심분야> MPLS, 서비스관리, 차세대 인터넷, Active Network, VoIP



李東遠(學生會員)

1999년 : 전남대학교 전자공학과 공학사. 2001년 : 전남대학교 전자공학과 공학석사. 2001년~현재 : 삼성반도체 근무중. <주관심분야> MPLS, 회로설계