

640Gb/s ATM 스위칭 시스템의 Throughput 향상을 위한 Self-Load Balancing 메커니즘 연구

장 석 기*, 김 태 영*, 김 훈*, 박 광 채**
조선대학교 일반대학원 전자공학과*, 조선대학교 전자정보통신공학부**
광주광역시 동구 서석동 375번지

A Study on Self-Load Balancing Mechanism for Improvement of Throughput in 640Gb/s ATM Switching system

Suk-Gi Jang*, Tae-Young Kim*, Hoon-Kim*, Kwang-Chae Park**

* Dept. of Electronics Eng., Graduate School of Chosun Univ

** School of Electronics, Information and Communications Eng. Chosun Univ.

375 Seoseck-dong, Dong-gu, Kwangju 501-759, Korea

E-mail : kcpark@mail.chosun.ac.kr

Abstract : Software controlled traffic management becomes more difficult with increasing switch size, because a huge number of connections are multiplexed in a high speed switching system and each ATM connections does not have a fixed bandwidth

It is based on high-speed WDM(wavelength division multiplexed) links that can multiplex a huge number of connections on a statistical basis, and a self-load balancing technique. the Proposed switching system is scalable, and achieves the non-blocking capability without the use of complicated software control, using instead a hardware self-load balancing mechanism

I. 서 론

인터넷 트래픽을 수용하는 일반적인 추세의 급속한 성장은 네트워크 공급자들의 경쟁을 부추기고 있다. 데이터, 음성 및 멀티미디어 통신을 위하여 융통성 있고, 성능이 우수한 플랫폼에 대한 수요가 급속히 증가하고 있다. 비동기 전송 방식(ATM : Asynchronous Transfor Mode)은 플랫폼을 수용하는 핵심기술이다. 멀티서비스 플랫폼은 고속 ATM 백본 네트워크를 이용함으로써 구성된다. 액세스 네트워크는 백본 네트워크와 서로 연결되고, 통신 프로토콜은 백본 네트워크 내부에서 실행된다. 이러한 효과적인 백본 구성은 대

용량의 고품질 데이터 전송을 제공할 수 있는 Tbps급 처리율을 갖는 고속 ATM 스위칭 시스템을 필요로 한다^[1]. 그러나, 스위칭 수용능력은 고속 광전송과 조밀 파장 분할 다중화(DWDM Dense Wavelength Division Multiplexer) 영역과 같은 오늘날의 전송 기술에서의 급속한 성장에 보조를 맞추지 못하고 있다.

다단 스위치 구성은 Tb/s급 ATM 스위칭 시스템을 구축하는데 이점을 가지고 있다. 이러한 구성은 대규모의 스위칭 시스템을 위해 필수적인 확장성을 요구한다^[2-3]. 또 다른 중요한 특징은 과부하로 인한 블로킹을 방지하는 기능이다.

소프트웨어 제어 트래픽 관리가 다수의 커넥션이 고속 스위칭 시스템에서 다중화되고, 각각의 ATM 커넥션은 고정된 대역폭을 갖고 있지 않기 때문에 증가하는 스위치의 크기로 인해 보다 어렵게 된다. 이러한 문제를 해결하기 위하여 quasi-non-blocking Tb/s급 ATM 스위치로 구성하고, 통계적 기초에 의한 다수의 커넥션과 Self-Load Balancing 기술을 복합할 수 있는 고속 WDM 링크를 기초로 한다. 스위칭 시스템은 복잡한 소프트웨어 제어의 사용없이도 하드웨어 Self-Load Balancing 메커니즘을 사용하여 비블로킹할 수 있다. 또한 Self-Load Balancing 메커니즘은 내부의 부하를 자동으로 체크하고, 안정시킴으로써 스위치에서의 과부하를 방지한다.

본 논문에서는 Tb/s급 ATM 스위칭 시스템 구성을 바탕으로 Self-Load Balancing 메커니즘을 제안하고, 제안한 메커니즘을 기반으로 640Gb/s 스위칭 시스템의 개요를 제시한다. 그리고, 640Gbps 스위칭 시스템 실험결과와 하드웨어의 실행을 제시한다.

II. Tb/s급 ATM 스위치 구성

640Gb/s급 스위칭 시스템은 0.25 μ m CMOS (Complementary Metal-Oxide Semiconductor) 장치, 고밀도 MCM(Multi-Chip Module) 기술, 광과장 분할 다중화(Optical WDM : Wavelength Division Multiplexing) 인터커넥션 기술을 이용하여 구성된다. scalable 640Gbps 스위칭 시스템의 핵심 부품인 WDM 인터커넥션 시스템은 과장 분할 다중화 라우팅 기술을 이용하여 80Gbps 스위칭 모듈을 서로 연결한다.

고속 인터커넥션 기술은 640Gbps 스위칭 시스템을 구성하기 위한 기술이다. 8개의 80Gb/s 스위칭 모듈의 인터커넥션은 64개의 교차점을 단마다 다루는 것이 필요하다. 일반적인 인터커넥션 기술은 각각의 교차점에서 요구되는 초고속 회선속도를 다룰 수 없기 때문에 비실용적이다. 그러므로 넓은 대역폭을 가지는 광 기술은 전기 스위칭 시스템의 처리를 한계를 극복하는데 있어 핵심이 된다^[5-6]. 이는 각각의 단에서 8개의 80 Gb/s급 스위치들을 다른 단의 스위치와 서로 연결하는 WDM 라우팅 시스템 구축이 목적이다. 구성은 광 라우터를 통하여 8개의 기본 스위치를 서로 연결하기 위해 8개의 다른 과장과 8 \times 8 광 라우터를 사용한다. 광 라우터는 앞단에 있는 모든 기본 스위치들이 여러 과장을 통해 다음 단에 있는 모든 스위치들로 어드레스 되는 것을 가능하게 한다. WDM 인터커넥션 시스템은 8개의 송신기(TX) 블록, 1개의 광 라우터, 8개의 수신기(RX) 블록으로 구성되어 있다. 각각의 TX 블록은 8개의 10Gbps 광 TX 모듈, 8개의 서로 다른 과장, 1개의 광 다중화기로 구성되어 있고, 다중화된 과장을 갖는 8 \times 10Gbps WDM 신호를 전송한다. 각각의 RX 블록은 1개의 역다중화기, 8개의 10Gbps 광 수신기 모듈로 구성되어 다중화된 과장 8 \times 10Gbps WDM 신호를 받아들인다. 송신기와 수신기 모듈의 중심에 있는 광 라우터는 광 라우터로 동작하는 AWGF로 이루어져 있다. 광 라우터가 입력 WDM 신호들을 과장에 따라 여러 출력포트에 교환한다는 것을 의미한다. 출력 채널 수와 과장의 라우팅 관계는 WDM 광 신호가 유입되는 입력채널 수에 좌우되는데, 순환적인 방법으로 실행된다. 따라서 앞단에 있는 1개의 스위치는 하나의 셀을 특별한 과장 신호를 이용하여 다음 단에 있는 스위치에 교환할 수 있다. 이러한 방법으로 8 \times 8 인터커넥션은 80Gb/s 처리율을 갖는 8개의 WDM을 수용할 수 있다.

quasi-non-blocking 다단 ATM 스위치는 다수의 동일한 스위칭 블록을 포함하여 확장이 용이하고, 스위치 non-blocking을 이루기 위해 Self-Load Balancing 메커니즘을 사용한다. 그림 1은 시스템의 구성도이다. 기본적인 스위칭 블록은 이상적인 출력 버퍼형 ATM 스위치와 유사하게 동작한다. 즉, 스위칭 블록에서 하나는 고 우선순위 셀과 다른 하나는 저 우선순위 셀을 위해서 두 가지 유형의 버퍼를 갖는다. 버퍼 스케줄러는 이들 버퍼

의 헤드 부분에 배치되어 먼저 고 우선순위 트래픽을 보내고, 다음에 저 우선순위 트래픽을 보낸다. 분배된 경로들이 주어지면, 두 가지 유형의 고 우선순위와 저 우선순위 트래픽들이 시스템에서 유지될 수 있다.

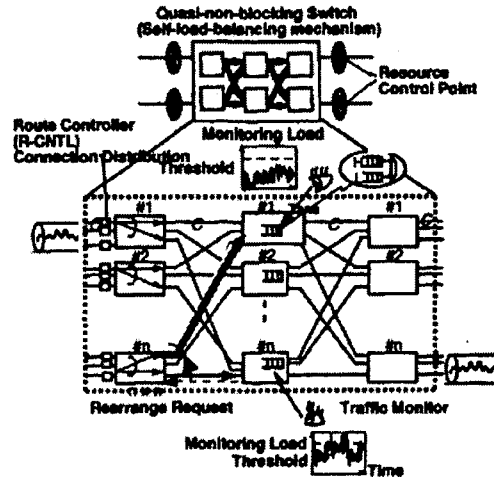


그림 1. Tb/s급 ATM 스위치 구성도

시스템은 커넥션을 바탕으로 한 라우팅을 이용한다. 다단 스위치에서 같은 가상채널(VC : Virtual Channel)에 속한 셀이 동일한 루트를 설정하는 동안 각각의 커넥션은 몇몇의 다경로로부터 두 번째 단에서 사용되는 스위치에서 부합하는 각각의 운송 루트를 선택할 수 있다. Self-Load Balancing 메커니즘은 전송 루트를 선택하고, 피드백 시스템의 종류인 하드웨어를 사용함으로써 실행된다. 하드웨어 실행 트래픽 모니터는 내부 링크의 평균 셀 전송부하인 잔여 대역폭을 계산한다. 이 정보는 내부 부하를 분배하고, 비교하기 위해 인입 가상채널의 스위칭 루트를 설정하여 루트 제어장치(R-CNTL : Router-Controller)에 되돌려 보내진다. 그러므로, 이 메커니즘은 약간의 출력회선에 부하의 집중을 막을 수 있고, 스위치를 non-blocking되게 할 수 있다.

III. Self-Load Balancing 메커니즘

Self-Load Balancing 메커니즘의 기본적인 방법은 예상되는 트래픽 부하에 도달할 때까지 최대 잔여 대역폭을 가지는 지점에 가상채널 루트를 둬으로써 내부 루트들 사이에서 입력 트래픽의 부하를 동등하게 분배하는 것이다. Self-Load Balancing 메커니즘은 루트 제어장치(R-CNTL)와 하드웨어로 실행되는 트래픽 모니터로 이루어져 있다. 이 메커니즘은 스위치가 연결 배치 요구를 받아들이고, 두 번째 단 스위치에서 임의의

640Gb/s ATM 스위칭 시스템의 Throughput 향상을 위한 Self-Load Balancing 메커니즘 연구

출력포트로 부터 과부하를 찾아낼 때 작용한다. 내부루트를 결정하고, 과부하 상태를 찾아내기 위해 트래픽 모니터는 다음의 저역 통과 필터를 이용하여 내부링크의 평균 셀 전송부하인 잔여 대역폭을 계산한다.

$$Ave(t) = (1 - \alpha) \times Ave(t-1) + \alpha \times Mon$$

식에서 Ave는 계산된 평균 부하를 나타내고, Mon은 모니터링 부하를 나타낸다^[4]. 저역 통과 필터는 트래픽 부하에서 파동의 고주파 성분을 제거하고, 평균 분배 부하를 계산할 수 있다. 식의 차단 주파수는 α 에 달려 있다. 따라서, α 는 가상채널의 최대 허용속도를 고려하여 설정한다. 이 정보를 이용하여 R-CNTL은 다음과 같이 각각의 가상채널의 내부 스위칭 루트를 제어한다.

■ 규칙 1 : 연결배치 요구를 받아들일 때, R-CNTL은 잔여 대역폭이 최대인 내부루트를 선택한다. 루트의 부하가 임계치 Th1을 초과한다면, 연결 배치는 거절된다. 이 상태에서 3단에 설치된 동일한 기본 스위치에 연결된 내부루트의 모든 부하는 한계를 초과하고, 만약 연결이 받아들여진다면, 과부하 상태가 발생할 것이다. 따라서, 임계치 Th1은 스위칭 시스템이 제공하는 최대 허용 부하에 의해서 설정된다.

이것은 시스템에서 액세스 트래픽의 도입을 막는다. 이러한 메커니즘은 최소 부하를 가지는 루트에 가상채널을 추가하므로써 두 번째 단계에 있는 스위칭 자원의 균형을 맞춘다. 시스템의 내부 블록킹 확률을 평가하고, 링크속도에 대한 과부하 확률의 상태를 그림 2에서 보여준다. 80Gbps 스위칭 블록 24(8×3)개를 이용하여 구성된 640Gbps 3단 스위칭 시스템에서 내부 블록킹 확률에 대해 살펴보면, 증가된 링크속도는 감소된 블록킹 확률을 수반한다는 것을 알 수 있다. 10Gbps 링크속도는 640Gbps ATM 스위칭 시스템에서 quasi-non-blocking (블록킹 확률 <math> < 10^{-10}</math>) 능력을 보장하기에 충분하다.

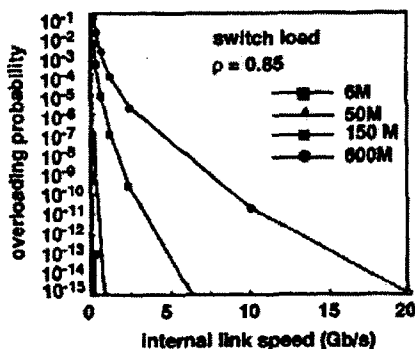


그림 2. 3단 ATM 스위치에서 고속 링크 영향

비록 블록킹이 시스템에서 자주 발생하지 않는다

하더라도 블록킹 방지 메커니즘은 필요하다. 이를 위해, Self-Load Balancing 메커니즘은 연결 재배열 기술을 사용하므로써 역시 블록킹을 방지한다. 블록킹을 방지하기 위한 또 다른 임계치 Th2는 셀 전송품질을 보장하기 위해 설정하는 것이다.

규칙 2 : 만약 트래픽 모니터가 두 번째 단 스위치의 어느 출력포트에서 예정된 간격동안 지속된 과부하 (측정된 평균 부하 > Th2)를 관찰하면, Self-Load Balancing 메커니즘이 실행되고, 첫 번째 단 스위치로부터 연결 재배열을 강요하는 back-pressure 신호가 발생한다. 앞에서 언급한 고 우선순위 연결이 시스템에서 우선한다. 그러므로, 과부하 상태를 줄이기 위해 R-CNTL은 다른 루트를 설정하고, 첫 번째 단 스위치는 과부하 점으로부터 최대 잔여 대역폭을 가지는 점까지 저 우선순위 연결을 재배열하고, 재배열 점에서 평균부하가 오직 임계치 Th1 이하일 때 행해진다. 연결 재배열을 수행하는 첫 번째 단 스위치에서 과부하 상태가 끝날 때까지 랜덤 선택은 수행된다. 만약 재배열 점에서의 평균부하가 임계치 Th1을 초과한다면 R-CNTL은 시스템으로부터 선택된 가상채널을 떨어뜨린다.

Th2 - Th1(Th2 > Th1)의 값이 VC의 최대 허용속도에 일치하도록 설정하기 때문에 재배열 점에서 안전한 셀 전송을 수행하는 것이 가능하다. Self-Load Balancing 메커니즘의 시뮬레이션 결과를 그림 3에서 보여준다.

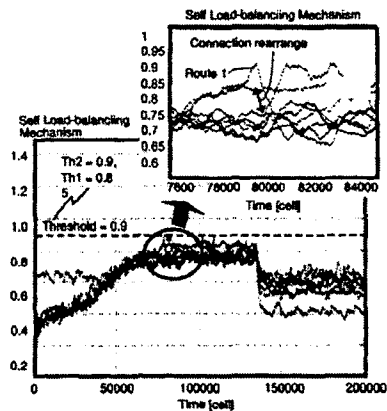


그림 3. Self- Load Balancing 메커니즘

고속 링크 영향 때문에 과부하는 시스템에서 좀처럼 발생하지 않는다. 그러므로, 고 부하에 고속연결을 추가하여 의도적으로 과부하를 줄일 수 있다. 루트 1의 time=79000에서 발생하는 과부하 상태에서 연결 재배열은 각각의 시간에서 수행되고, 블록킹은 종결된다. 연결 재배열은 스스로 발견하는 원리에 의해 수행되지만 고속 인터커넥션 영향과 규칙 2에서 실행된 메커니즘에 의해 적절히 이루어진다. 따라서, 내부 트래픽은 단순한 하드웨어에 의해 자동적으로 제어될 수 있다.

연결 재배열 메커니즘이 OPTIMA 시스템에서 수행될 때, 셀 배열의 보전을 확보하기 위해서 Self-Load Balancing 메커니즘이 수행되는 단순 하드웨어와 고속 연결은 640Gb/s ATM 스위칭 시스템이 quasi-non-blocking 능력을 갖도록 한다.

IV. 성능분석

효과적인 고속 ATM 스위칭 시스템을 개발하는데 있어 하드웨어 기술의 확장성에 있어서 전체 스위칭 용량은 스위치(N)의 기본적인 크기 즉, 각 단(S)에서 스위치 수와 인터커넥션의 회선속도(L)에 좌우된다. 그리고, 스위치 수와 인터커넥션의 회선속도는 하드웨어 기술에 커다란 영향을 준다. 광WDM 기술을 효과적으로 활용하기 위해 시스템은 N=8Gbps, S=8Gbps, L=10Gbps로 설정한다.

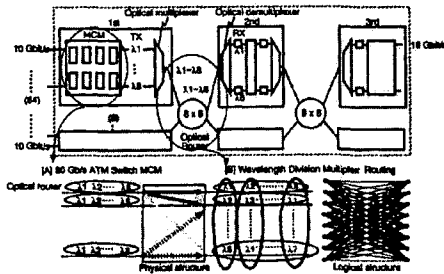


그림 4. 640b/s Throughput을 갖는 시스템 구성도

시스템은 일치하는 3개의 스위칭 단으로 이루어져 있다. 이들 스위칭 단은 8개의 WDM 신호를 경유해서 광 AWG 라우터를 통해 서로 연결된다. 그림 5에서는 스위칭 시스템을 사용하여 기본 스위칭 기능과 WDM 라우팅 기능과 3개의 단들 사이에서 안정된 셀 스위칭을 확인하였다. 또한 기본 80Gb/s 스위칭 모듈의 출력 파형과 일치하는 622Mbps 인터페이스와 WDM 라우팅 인터커넥션의 출력파형과 일치하는 10Gb/s 인터페이스에서의 셀 스위칭 파형 형태를 보여준다.

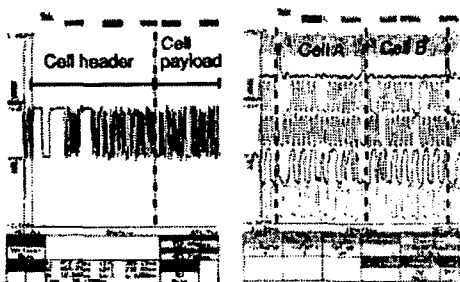


그림 5. 10Gb/s와 622Mb/s 인터페이스의 출력 스위칭 파형 패턴

V. 결론

Tbps급 ATM 스위칭 시스템은 Self-Load Balancing 메커니즘과 고속 인터커넥션 효과의 이점으로 인해 Throughput이 향상되고, quasi-non-blocking 한다. Self-Load Balancing 메커니즘은 640Gb/s 스위칭 시스템을 quasi-non-blocking하도록 자동적으로 체크하고, 내부부하를 균등하게 하므로써 스위치에서 과부하를 방지한다.

640Gb/s ATM 스위칭 시스템을 바탕으로 실험하였고, 하드웨어 실험결과를 제시하였다. 스위칭 시스템은 개선된 0.25 μm CMOS 장치, 고밀도 MCM 기술, 광WDM 인터커넥션 기술을 이용하여 제작되는데, 640Gb/s 스위칭 시스템의 핵심 부분인 WDM 인터커넥션 시스템은 WDM 라우팅 기술을 이용하여 80Gbps 스위칭 모듈들을 서로 연결시킨다. 640Gbps 스위칭 시스템은 고속 멀티미디어 네트워크를 위한 플랫폼의 구축에 있어 응용될 수 있다.

참고 문헌

- [1] H. Ishikawa, "Evolving from narrowband," *IEEE Commun. Mag.*, vol. 30, no. 8, pp. 32-36, 1992.
- [2] R. Melen and J. S. turner, "Non-Blocking Networks for Fast Packet Switching," *Proc. IEEE Infocom'89*, pp. 548-557, 1989.
- [3] K. Y. Eng, M. J. Karol, and Y. S. Yeh, "A Growable Packet (ATM) Switch Architecture: Design and Application," *IEEE Trans. Commun.*, pp. 423-430, 1992.
- [4] K. Shiimoto and N. Yamanaka, "An admission control scheme based on measurements of instantaneous utilization," *IEICE Trans. Commun.*, Dec. 1997.
- [5] E. Munter, J. Parker, and P. Kirkby, "A High-Capacity ATM switch Based on Advanced Electronic and Optical Techniques," *IEEE Commun. Mag.*, pp. 64-71, 1995.
- [6] M. Tsukada, A. Misawa, J. Nishikido, Y. Shimazu, and H. Nakano, "Experiments on photonic cell switching with an optical input buffer," *Electron. Lett.*, 30, 13, pp. 1081-1082, 1994.