

가변 길이 패킷을 고려한 스위칭 패브릭의 설계

류경숙^o 김무성 최병석

동국대학교 정보통신학과

ksryu@dgu.edu, mskim@dgu.edu, bchoe@dgu.edu

Design of Switching Fabric for variable length packets

Kyoungsook Ryu^o Musung Kim Byeongseog Choe

Department of Information Communication Engineering,

Dongguk University, Korea

최근 인터넷 환경에서의 초고속 스위치의 발전 방향은 기가비트 스위칭이 가능하도록 하고자 스위치 내부의 입/출력 인터페이스 간 패킷 전송에 있어서 ATM 스위치와 같은 고속의 하드웨어 스위칭 기법을 도입하고 있다[1][2]. Abacus와 같은 ATM 스위치는 53 옥텟(Octet)의 고정 셀 스위칭을 하는데 반하여 인터넷 망에서는 가변 길이 패킷을 서비스 해야 하는 차이점이 있다. 일반적으로 인터넷 프레임은 64 바이트에서 1518 바이트의 길이를 가진다. 스위칭 패브릭(switching fabric)에서 이러한 가변 길이 패킷을 효율적으로 스위칭 하기 위해서는 새로운 형태의 스위치 구조가 필요하다.

현재까지 스위칭 패브릭을 적용하여 가변 길이 IP 패킷을 서비스 하기 위한 연구에는 두 가지 접근 방법이 있다. 하나는 패킷들이 입력 단에서 스위칭 패브릭에 진입하기 전에 패킷들을 일정한 셀 단위로 분할하여 전송하고 출력 단에서 재조립하는 방법이며 다른 하나는 가변 길이 패킷을 별도의 가공 없이 서비스 할 수 있도록 스위칭 패브릭 내에 가변 길이 패킷을 저장하기 위한 저장 공간(cross-point buffer)을 확보하는 방법이다 [3][4][5][7]. 전자는 패킷을 분할하고 재조립 하기 위한 시간들이 시스템에 부하로 작용할 수 있고 후자는 단편화(fragmentation) 문제는 해결할 수 있지만 패킷이 스위칭 패브릭을 통과하는 동안에도 지속적으로 가변 길이 패킷들을 서비스해야 하므로 스위칭 패브릭의 속도 면에서 역시 문제가 된다.

제안한 스위치는 크게 두 개의 평면(plane)으로 구성된다. 하나는 데이터 제어 평면(data control plane) 부분이고 다른 하나는 주소 포인터 스위칭 평면(address pointer switching plane)이다.

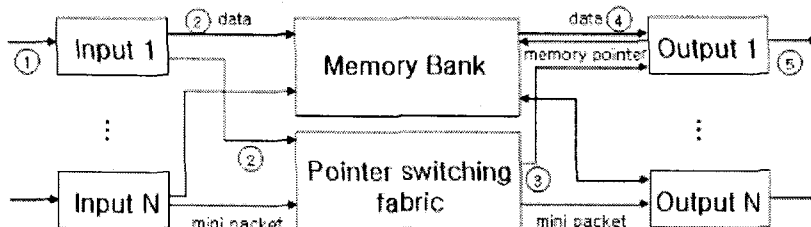


그림 1. 제안한 스위치의 동작

입력 인터페이스에 입력된 패킷을 헤더 정보와 분리하여 데이터 부분은 메모리에 저장하며 동시에 저장된 데이터의 주소 포인터(Address Pointer) 부분만을 스위칭 패브릭을 사용해서 스위칭 하도록 한다. 스

위칭 패브릭은 구현이 간단하면서도 확장성이 우수한 ATM 스위치의 스위칭 패브릭을 사용한다. 데이터는 입력 포트와 출력 포트 번호로 참조되는 독립적인 메모리들의 집합인 메모리 뱅크(Memory Bank)에 저장한다. 각 라인 카드에 라우팅 테이블과 포워딩 엔진을 두어 라우팅 기능을 각 라인 카드에 분산하여 병렬적인 처리가 가능하도록 하고 있다. 성능 향상을 위해 출력 큐는 ASIC 형태의 자동 정렬 큐(ASQ: Auto Sorting Queue)로 구현하여 스케줄링 알고리즘의 복잡성을 회피하고자 한다[6].

제안한 스위치의 동작을 간략하게 살펴보면 다음과 같다.

- ① 스위치의 입력 인터페이스에서 패킷이 들어오면 라우팅 헤더 부분을 분리하고 라우팅 테이블을 룩업(lookup)한다.
- ② 헤더와 룩업 정보를 바탕으로 제안한 스위치 구조에서 QoS와 스위칭의 목적으로 사용하게 될 미니 패킷을 작성하여 주소 스위칭 패브릭으로 보냄과 동시에 메모리에 데이터를 저장한다.
- ③ 출력 인터페이스에 도착한 미니 패킷은 ASQ에 의해서 입력 포트에서 지정한 우선 순위에 따라 오른 쪽부터 전송 순서대로 정렬(sorting) 된다.
- ④ 해당 미니 패킷이 전송될 순서(HOL: Head Of Line)가 되면 메모리 포인터 주소를 참조해서 메모리 뱅크에서 데이터 부분을 읽어 오고 참조한 메모리 인덱스 부분은 해제(free) 된다.
- ⑤ 메모리 뱅크에서 읽은 데이터와 라우팅 헤더를 결합하여 다음 홉(hop)으로 전송한다.

스위치에 진입한 모든 패킷은 입력 단에서 라우팅 테이블 룩업을 위해서 헤더 정보를 읽히게 되는데 이 과정에서 필요한 계산들을 미리 하게 되면 효과적이기 때문에 제안한 스위치 구조에서는 QoS 처리를 위한 스케줄링과 출력 큐에서의 정렬을 위한 정보는 입력 포트에서 처리하도록 한다. 출력 포트에서 조절할 경우 보다 정확하고 안정적일 수 있지만 헤더 정보를 한 번 더 읽어야 하고 고속 스위치에서 QoS를 복잡하게 계산하는 것은 지연을 초래할 수 있다.

성능 평가 결과 제안한 구조는 초고속 인터넷 망의 고속 패킷 스위칭과 사용자 요구를 동시에 만족할 수 있는 구조로 적합하다는 것을 확인하였다.

참고문헌

- [1] H. Chao and B. Choe, "Design and Implementation of Abacus Switch: A Scalable Multicast ATM Switch," IEEE JSAC. vol. 15, no. 5, Jun., 1997.
- [2] H. Chao and B. Choe, "Design and Analysis of a Large-Scale Multicast Output Buffered ATM Switch," IEEE/ACM Trans. Networking, vol. 3, no. 2, Apr., 1995.
- [3] E. Oki and N. Yamanaka, "Scalable Crosspoint Buffering ATM Switch Architecture Using Distributed Arbitration Scheme," in Proc. IEEE ATM '97 Workshop, 1997.
- [4] C. Minkenberg and T. Engbersen, "A Combined Input and Output Queued Packet-Switched System Based on PRIZMA Switch-on-a-Chip Technology," IEEE Commun. Mag., pp. 70-77, 2000.
- [5] K. Yoshigoe and K. Christensen, "An Evolution to Crossbar Switches with Virtual Output Queuing and Buffered Cross Points," IEEE Network, pp. 48-56, Sep./Oct., 2003.
- [6] H. Jonathan Chao, "A VLSI Sequencer Chip for ATM Traffic Shaper and Queue Manager," IEEE JSAC, vol. 27, no.11, Nov, 1992.
- [7] "Cisco 12000 Series Gigabit Switch Routers," http://www.cisco.com/warp/public/cc/pd/rt/12000/prodlit/gsr_ov.pdf