

고성능 셀/패킷 스위치를 위한 고속 VOQ 관리기 설계

정갑중, *이범철

경주대학교 컴퓨터전자공학부, Email: gijeong@kyongju.ac.kr

*한국전자통신연구원, Email: bclee@etri.re.kr

Design of High-Speed VOQ Management Scheme for High Performance Cell/Packet Switch

Gab Joong Jeong, *Bhum Cheol Lee

School of Computer and Electronic Engineering, Kyongju University

*Electronics and Telecommunications Research Institute

Abstract

This paper presents the design of high-speed virtual output queue(VOQ) management scheme for high performance cell/packet switch, which has a serial cross bar structure. The proposed VOQ management scheme has been designed for wire-speed routing with a pipelined buffer management. It provides the tolerance of requests and grants data transmission latency between the VOQ manager and central arbiter using a new request control method that is based on a high-speed shifter. The designed VOQ manager has been implemented in a field programmable gate array chip with a 77MHz operating frequency, a 900-pin fine ball grid array package, and 16x16 switch size.

I. 서론

입력버퍼 스위치 구조는 스위치의 각 입력 포트에 단일 버퍼나 다중 버퍼를 가지게 하는 방식으로 나눌 수 있다. 입력 버퍼 스위치이면서 입력 포트 당 단일 버퍼를 가지는 구조는 지연 셀 또는 데이터 패킷을 단일 FIFO 큐 또는 다중 FIFO 큐에 저장한다. 단일 FIFO 큐를 가지는 입력 버퍼는 모든 셀을 목적 출력 포트에 관계없이 단일 큐에 저장하며 단일 큐의 head 셀만이 현재 전송 시간 슬롯에 전송 가능하며 매 전송 슬롯 당 하나의 셀만이 전송된다. 따라서 단일 입력 큐의 head 셀이 전송되고자 하는 목적 출력 포트가 다른 포트의 셀을 출력하기 위해 이미 할당 되어 있을 때 해당 출력 셀은 전송되지 못하고 차단되거나 같은 큐에 저장된 차단된 셀의 뒤에서 대기하는 셀은 head-of-line(HOL) 차단을 경험한다. 본 방식은 스위치의 최대 성능이 58.6%로 한정된다 [1].

입력 버퍼 스위치의 성능을 향상시키기 위해 하나의 입력 버퍼 내에 다중 입력 큐를 관리하는 방식이 개발되었다 [2], [3]. 모든 입력 버퍼 내에 각 출력 포트에 대응하는 큐를 두는 방식의 다중

입력 큐를 관리 함으로써 모든 출력 큐에 대기 셀이 있을 때 모든 입력 버퍼들 내의 각 출력 큐의 head 셀은 동일한 출력 포트로의 전송 가능 셀이 된다. 이러한 방식이 가상 출력 큐(virtual output queue: VOQ) 방식이다. 가상 출력 큐 방식에서 각 큐의 head 셀은 해당 입력 버퍼 내에서 전송 가능하나 단지 전체 셀 중에서 하나의 전송 시간 슬롯 당 하나의 셀만 해당 버퍼에서 전송 가능하다.

많은 입력 버퍼가 각각 현재 전송 시간 슬롯에서 다중으로 하나의 출력 포트로 셀을 전송하고자 하는 경우에 하나의 출력 포트는 하나의 입력 포트로부터 셀을 전송할 수 밖에 없으므로 모든 입력 버퍼들에게 공정한 전송 기회를 할당하기 위한 스케줄링 알고리즘이 필요하다. 셀을 전송하고자 하는 입력 포트들에게 매 전송 시간 슬롯마다 충돌이 없도록 출력 포트를 할당하는 중앙 중재기는 매 슬롯마다 계산량이 매우 많고 복잡할 수 밖에 없으며 스위치의 성능에 많은 영향을 미친다. 따라서 HOL 차단 현상을 가지는 입력 버퍼 스위치에서 스위치의 성능을 높이기 위해 더욱 향상된 스케줄링 알고리즘들이 연구되어 왔다 [1]-[5]. 그러나 기존의 연구들에서 셀 스케줄링 시 입력 버퍼와 중앙 중재기 사이의 요구 및 허가 정보의 전달 시에 발생하는 전송 지연에 대한 연구는 고려되지 않았다. 중재 정보의 전송지연은 ATM과 같이 고속의 짧은 셀 전송을 필요로 하는 응용 시스템에서는 매우 중요한 문제가 된다.

본 논문에서는 중재 정보의 전송 지연을 가지는 고속 ATM 스위치 시스템에서 전송 지연을 고려한 VOQ 관리기의 구조 및 설계와 스위치의 성능 향상에 대해 연구하였다. 본 논문에서 제안된 VOQ 관리기는 VOQ의 동적 할당을 지원하고 파이프라인 방식의 큐 관리를 이용한 wire-speed 라우팅을 지원한다. 단일 공유 버퍼를 이용한 다중 큐의 동적 할당은 버퍼 내의

필요 메모리의 크기를 줄이면서 성능을 향상시키고 고속 쉬프팅 방식을 이용한 전송 요청 정보의 관리는 전체 스위치 시스템의 성능 향상을 가져온다.

II. 스위치 구조

본 논문에서는 입력 버퍼 내에 각 출력 포트 당 큐를 가지는 독립된 VOQ를 관리하는 입력 버퍼 스위치 구조를 사용한다. 그림 1에 전체 스위치의 구조와 중재 요청 FIFO의 구조를 나타내었다. 각 VOQ당 본 논문에서 제안된 요청 FIFO(request FIFO)를 할당하였다. 입력 버퍼 내의 요청 FIFO는 각 VOQ의 중재 요청이 이루어진 head 셀에 대한 다중 요청 비트를 저장한다. 그리고 중앙 중재기에서는 중재에서 탈락된 모든 입력 버퍼의 모든 VOQ에 대한 요청 신호를 저장하는 요청 FIFO를 관리한다. 중앙 중재기의 요청 FIFO에 저장된 요청 신호는 다음 중재 시에 연속적인 중재가 이루어진다. 본 논문에서 연구된 스위치 구조는 중재 정보 전송 지연에도 불구하고 입력 버퍼의 각 VOQ의 새로운 head 셀을 위한 연속적인 중재 요청 신호의 전송을 가능하게 하고 각 입력 버퍼의 연속적인 허가 신호 전송을 가능하게 한다.

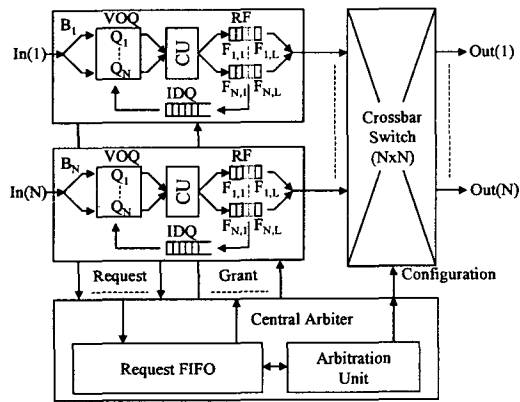


그림 1. 스위치 구조

III. 제안된 고속 VOQ 관리기

제안된 구조의 VOQ 관리기는 기능적으로 입력 셀 쓰기 기능, 출력 셀 읽기 기능, 폴리싱 기능, 요청 FIFO 관리 기능으로 나뉘어진다. 각 기능 블록은 파이프라인 방식으로 연결되고 wire-speed 라우팅을 저가격으로 설계 가능하게 한다. 그림 2에 제안된 VOQ 관리기의 전체 구조를 나타내었다. 제안된 VOQ관리기의 전체

블록은 VOQ관리 모듈(VOQ), 자유 큐 관리 모듈(idle queue: IDQ), 폴리싱 모듈(policing module: PM), 읽기 포인터 관리 모듈(read pointer manager: RPM), 쓰기 포인터 관리 모듈(write pointer manager: WPM), 입력 셀 쓰기 모듈(incoming cell write: ICW), 출력 셀 읽기 모듈(outgoing cell reader: OCR) 그리고 요청 FIFO 관리 모듈(request FIFO controller: RFC)로 크게 나누어진다. 그 외에 시스템 인터페이스를 위한 CSIX(common switch interface) 모듈(cell framer: CF), 기가비트 시리얼 인터페이스 모듈(gigabit data interface: GDI), back-pressure 관리 모듈(back-pressure controller: BPC) 및 프로세서 인터페이스 모듈(processor interface: PI)이 있다. 제안된 VOQ 관리기는 외부의 듀얼 포트 싱크로너스 메모리(dual port synchronous SRAM)를 포인터 메모리(ingress pointer memory: INPM)와 버퍼 메모리(ingress buffer memory: INBM)로 사용한다.

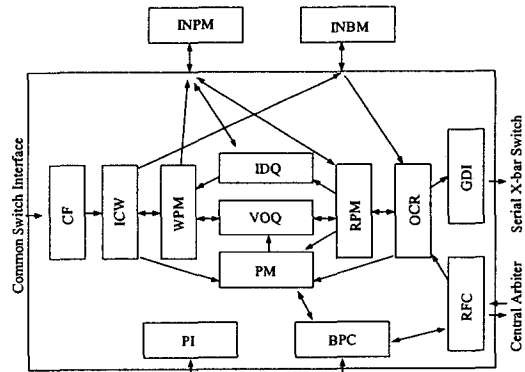


그림 2. 제안된 VOQ 관리기

요청 신호 관리기는 요청 신호 저장을 위해 각 VOQ에 대응하는 FIFO 레지스터로 구성되어 있으며 중앙 중재기와 통신한다. 또한 VOQ와 요청 FIFO의 상태에 따라 요청 신호의 생성 및 소거 기능을 수행한다. 요청 신호의 생성 및 소거 알고리즘은 VOQ에 하나 이상의 대기 셀이 저장되어 있고 해당 VOQ에 대응하는 요청 FIFO 레지스터의 첫 번째 요소가 유효 요청 신호가 아닐 때 하나의 유효 요청 신호를 발생시킨다. 그리고 발생된 요청 신호는 중앙 중재기로 보내짐과 동시에 해당 요청 FIFO 레지스터에 저장된 기 발생된 요청 신호를 쉬프트 시키면서 마지막 요소에 현재 발생된 요청 신호를 저장한다. 요청 신호가 발생된 후 해당 VOQ의 길이는 1 감소한다. 따라서 요청 신호 관리기 내의 각 VOQ의 길이는 폴리싱 모듈 내의 VOQ의 길이 보다 요청 FIFO에 저장된 유효

고성능 셀/패킷 스위치를 위한 고속 VOQ 관리기 설계

요청 신호 수 만큼 작은 값을 저장하고 있다. 요청 신호 관리기 내의 VOQ 길이가 0이고 요청 FIFO의 첫번째 요소가 유효 요청 신호가 아니면 요청 신호 관리기는 무효 요청 신호(*invalid request*)를 생성하며 동시에 요청 신호 FIFO 내에 저장된 기 발생 요청 신호를 쉬프트 시킨다. 그리고 요청 신호 FIFO의 첫번째 요소가 유효 요청 신호이면 각 VOQ에 대기 셀이 있다 하더라도 더 이상의 요청 신호를 발생 시키지 않는다. 그림 3에 본 논문에서 제안된 고속 쉬프트를 사용한 요청 신호 관리 알고리즘을 흐름도로 나타내었다.

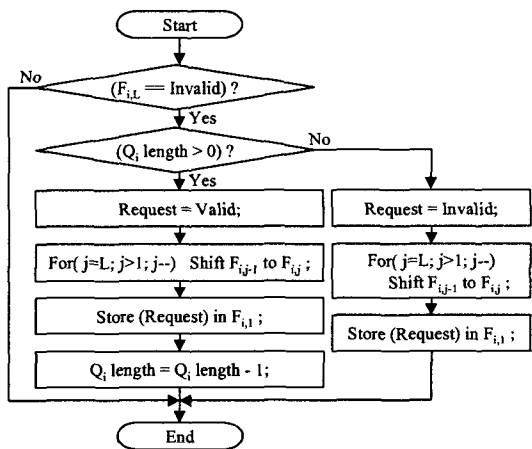


그림 3. 요청 신호 관리 알고리즘 (F: request FIFO element, i: number of output queue, j: number of request bit).

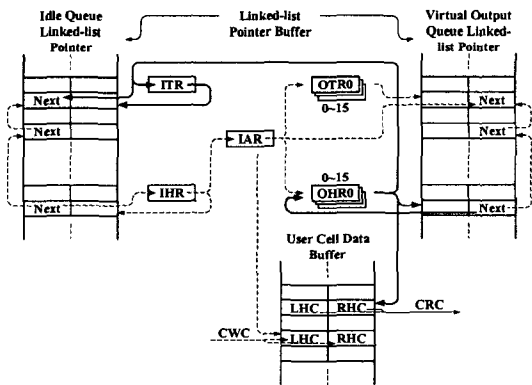


그림 4. 출력 셀 읽기 시의 포인터 데이터 패스 (OHR: output queue head register, OTR: output queue tail register, IHR: idle queue head register, ITR: idle queue tail register, IAR: idle address register, LHC: left half cell, RHC: right half cell, CWC: current writing cell, CRC: current reading cell).

중앙 중재기로부터 요청 신호 관리 모듈로 하나의 허가 신호와 출력 포트 번호가 입력 버퍼로 도달하면 입력 버퍼 내의 요청 신호 관리 모듈은 허가된 출력 포트에 해당하는 요청 신호 FIFO에 저장되어있는 요청 신호 중 가장 오래 대기한 요청 신호를 삭제하며 전송 허가된 출력 포트 번호를 출력 셀 읽기 모듈로 전달한다. 출력 셀 읽기 모듈은 허가된 출력 포트 번호를 읽기 포인터 관리 모듈과 폴링 모듈로 전송하고 읽기 포인터 관리 모듈로부터 입력 받은 출력 셀 주소를 이용해 데이터 버퍼 메모리로부터 출력 셀을 읽는다. 읽기 포인터 관리 모듈은 현재 VOQ 모듈의 선택된 포트의 큐를 포인터 버퍼 메모리로부터 추출된 다음 셀 주소를 이용해 갱신한다. 이와 같은 데이터 버퍼 메모리와 포인터 버퍼 메모리로부터 하나의 셀을 읽기 위한 파이프라인 방식의 포인터 주소 데이터 흐름을 그림 4에 나타내었다.

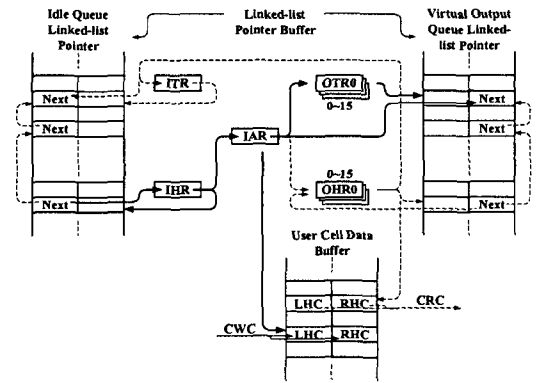


그림 5. 입력 셀 쓰기 시의 포인터 데이터 패스

읽기 포인터 관리 모듈은 현재 출력 셀이 멀티캐스트 셀인 경우 출력 셀의 주소를 다음 목적 출력 포트의 큐에 다시 저장한다. 새로운 도착 셀은 입력 셀 쓰기 모듈이 쓰기 포인터 관리 모듈로부터 새로운 쓰기 셀 주소를 받아 셀 버퍼 메모리에 저장하며 동시에 쓰기 포인터 관리 모듈은 현재 도착 셀의 목적 출력 포트에 해당하는 VOQ의 포인터를 현재 도착 셀이 저장될 새로운 쓰기 주소로 갱신한다. 그림 5에 새로운 도착 셀의 처리를 위한 파이프라인 방식 포인터 데이터의 흐름을 나타내었다. 요청 신호 관리 모듈은 *back-pressure* 관리 모듈을 통해 새로운 입력 셀의 도착 정보와 멀티캐스트 셀의 다음 목적 출력 포트 정보를 입력 받아 전체 VOQ의 상태를 관리한다. 그림 6에 셀 읽기, 쓰기 및 멀티캐스트 셀 처리를 포함하여 VOQ와

자유 큐가 하나의 이중 포트 싱크로너스 SRAM인 포인터 버퍼 및 데이터 버퍼 메모리를 완전 공유하는 전체 포인터 데이터 흐름을 나타내었다.

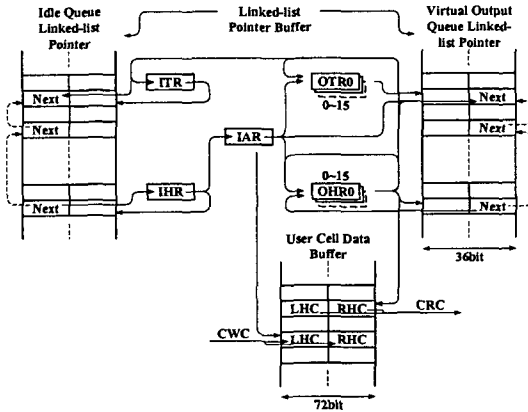


그림 6. 제안된 VOQ 관리기의 전체 포인터 데이터 패스

IV. 설계된 CHIP의 특성

본 논문에서 제안된 VOQ 관리기는 FPGA를 이용하여 설계 및 구현하였다. 표 1에 설계된 VOQ관리기의 특성을 나타내었다. 설계된 VOQ관리기는 OC-48c의 포트 속도를 지원하며 작렬 크로스바 스위치를 사용하는 초고속 백본 네트워크의 코어 ATM 스위치 시스템에서 기능 및 동작을 검증하였다. 설계된 VOQ 관리기는 개발된 스위치 시스템의 입력 포트 버퍼 관리기 뿐만 아니라 출력 포트의 버퍼 관리기에도 동시에 사용하여 스위치 시스템 개발에 적용하였다. 또한 VOQ관리기의 입력 데이터 형식은 CSIX(common switch interface) 데이터 패킷을 수용하도록 설계하였으며 CSIX를 이용하여 스위치 패브릭 외부의 상용 입력 및 출력 포트 프로세서와 연동하였다. 컴퓨터 시뮬레이션을 통해 128 셀 입력 공유 버퍼 크기를 가지는 VOQ 관리기는 random uniform traffic에서 2DRR 중재 알고리즘을 사용할 경우 42.2%의 포화 스위치 성능 향상을 나타내었고 iSLIP 중재 알고리즘을 사용할 경우 7.5%의 성능 향상을 나타내었으며 최대 스위치 성능은 98.6%로 나타났다.

표 1. 설계된 VOQ 관리기의 CHIP 특성

FPGA device	XCV1000E-6	INBM data bus width	72bit
package	900pin FGA	buffered cell size	640bit
I/O	LVTTL	INPM data bus width	36bit
used I/O	629pin	managing pointer packet size	64bit
estimated power dissipation	13.2W	experimental operating freq.	62.5MHz
equivalent gate count	380K	experimental port speed	OC-48c (2.5G)
CLB usage	40%	experimental switch size	16x16
max. operating freq.	77MHz	experimental one cell time	160ns
max. bit rate per port	5Gbit/s	aggregated switch throughput	40Gbit/s

V. 결론

본 논문에서는 입력 버퍼 구조의 고성능 셀/패킷 스위치를 위한 새로운 고속 VOQ 관리기의 구조를 제안하고 설계 및 기능 검증을 하였다. 설계된 VOQ 관리기는 파이프라인 방식을 이용한 wire-speed 라우팅을 지원하고 저가격의 고속 VOQ 관리기를 구현할 수 있게 하였다. 또한 제안된 VOQ 관리기는 대용량 ATM 스위치에서와 같은 고속 스위치 패브릭에서 나타나는 입력 버퍼와 중앙 중재기 사이에 존재하는 중재 정보 전송 지연을 새로운 요청 신호 쉬프트 방식으로 지원하였으며 파이프라인 방식의 라우팅에 의한 입력 스위치의 성능 향상을 시뮬레이션을 통해 검증하였다. 설계된 VOQ 관리기는 FPGA를 이용하여 Chip으로 구현되었으며 16x16 스위치 크기, OC-48c 포트 속도, 40Gbps의 용량을 가지는 백본용 코어 ATM 스위치 시스템에 실장 되어 모든 기능 검증을 완료하였다.

참고문헌

- [1] H. Obara, S. Okamoto, and Y. Hamazumi, "Input and output queueing ATM switch architecture with spatial and temporal slot reservation control," *Electron. Lett.*, vol. 28, no. 1, pp. 22-24, Jan. 1992.
- [2] N. McKewon, P. Varaiya, and J. Walrand, "Scheduling cells in an input queued switch," *Electron. Lett.*, vol. 29, no. 25, pp. 2174-2175, 1993.
- [3] R. O. LaMaire and D. N. Serpanos, "Two-dimensional round-robin schedulers for packet switches with multiple input queues," *IEEE/ACM Trans. Networking*, vol. 2, no. 5, pp. 471-482, 1994
- [4] N. McKeown, "The iSLIP scheduling algorithm for input-queued switches," *IEEE/ACM Trans. Networking*, vol. 7, no. 2, pp. 188-201, 1999.
- [5] P. Gupta and N. McKeown, "Designing and implementing a fast crossbar scheduler," *IEEE Micro*, vol. 19, no. 1, pp. 20-28, 1999.