

---

# 고속 입력 큐 스위치를 위한 고성능 라우팅엔진

정갑중<sup>\*</sup> · 이범철<sup>\*\*</sup>

<sup>\*</sup>경주대학교, <sup>\*\*</sup>한국전자통신연구원

High Performance Routing Engine for an Advanced Input-Queued Switch Fabric

Gab Joong Jeong<sup>\*</sup> · Bhum-Cheol Lee<sup>\*\*</sup>

<sup>\*</sup>Kyongju University, <sup>\*\*</sup>Electronics and Telecommunications Research Institute

E-mail : gjeong@kyongju.ac.kr

## 요 약

본 논문에서는 고속 입력 큐 스위치에서 발생하는 중재정보전달지연 현상을 수용하기 위한 고성능 라우팅엔진의 구조를 제안한다. 제안된 고성능 라우팅엔진은 2.5Gbps의 스위치 입출력 포트 속도에 대해 사용자 셀 데이터의 지연 없이 동작한다. 또한 입력버퍼와 중앙중재기 사이에서 발생하는 요청신호와 허가신호의 전송지연을 수용하는 구조로 설계되었다. 중재정보전송지연 현상의 처리 방법으로는 고속 쉬프터를 사용하여 많은 회로의 추가 없이 구현하였다. 라우팅엔진 내의 세부 블록의 파이프라인 처리를 통하여 저 가격 고성능의 입력 버퍼 설계를 실현하였다.

## ABSTRACT

This paper presents the design of a pipelined virtual output queue routing engine for an advanced input-queued ATM switch, which has a serial cross bar structure. The proposed routing engine 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 routing engine and central arbiter using a new request control method that is based on a high-speed shifter. The designed routing engine has been implemented in a field programmable gate array (FPGA) chip with a 77MHz operating frequency, 16x16 switch size, and 2.5Gbps/port speed.

## Keywords

ATM Switch, Pipelining, Wire-Speed Routing

### 1. Introduction

Input queuing switch architecture can be classified typically by whether a single buffer is allocated in every input port or whether multiple buffers are allocated in every input port. Architecture with input queuing and a single buffer per input port can store blocked packets into a single FIFO queue or into multiple FIFO queues. With a single FIFO queue, all packets are stored in this queue irrespective of their destination. Only one packet at the head of the queue is eligible for transmission. When the destination output port for the packet at the head of the queue is busy, the packets behind it are subjected to the

head-of-line (HOL) blocking. This scheme has been shown to have a maximum throughput of 58.6% [1].

To enhance the throughput of an input-queued switch, schemes with multiple input queues within an input buffer have been proposed [2], [3]. Typically, there is one queue corresponding to every output port in every input buffer. Therefore, every queue has packets, which have arrived from its given input port and which are all destined for the same output port. This scheme is called as virtual output queuing. Under the virtual output queue scheme, packets at the head of the individual queue are eligible for transmission in an input buffer. However, only

one of these packets can be transmitted from the buffer at any given time slot.

In this case, plural input buffers may need to transmit multiple packets to a single output port in one time slot. Since no more than one output port can read from a given input buffer at a time slot, this scheme requires a fair scheduling among the plural input buffers. The centralized arbitration to determine a conflict-free assignment of output ports to input buffers that have packets is quite complex and affects the performance of an input-queued switch. Hence, many research groups' efforts for an input-queued switch having the HOL blocking problem are focused on performance improvement, which requires an enhanced scheduling algorithm for better throughput [1]-[5]. However, these previous works have not considered arbitration latency, which includes transmission latency of request and grant data between plural input buffers and a central arbiter, in cell scheduling and routing. The arbitration latency problem appears seriously if an input-queued ATM switch uses in-band transmission of request and grant data based on high-speed serial links.

In this paper, we introduce the design of a virtual output queue routing engine for an advanced input-queued ATM switch, which has transmission latency of arbitration information and a pipelined approach to improve the throughput [6]. The proposed routing engine as an input buffer of an input-queued ATM switch provides wire-speed pipelined queue management that has dynamically allocated virtual output queues. The dynamic allocation of queues minimizes the amount of packet memory in an input buffer. The routing engine adopts a novel method to manage the transmission latency between input buffers and a central arbiter, which is based on request shifting.

## II. Switch Architecture

The switch system considered in this paper uses a type of an advanced input-queued switch in which a separate queue called virtual output queue (VOQ) is maintained at each input buffer for each output port. Fig. 1 illustrates the advanced switch architecture and the organization of input queue and a request FIFO. We insert the secondary FIFO buffer for each VOQ, called request FIFO (RF), in each

input buffer. The RF in an input buffer stores multiple request bits, which is requested for arbitration of a packet at the head of each output queue. We then add a request FIFO at the central arbiter, which stores the request data that has not been selected to transmit a cell. The residues of request data in the RF of the central arbiter are consecutively considered in the next arbitration. The switch system allows for the transmission of consecutive request data for the new head cell of each VOQ in an input buffer, and it allows for the transmission of the consecutive grant data for each input buffer, despite the transmission latency.

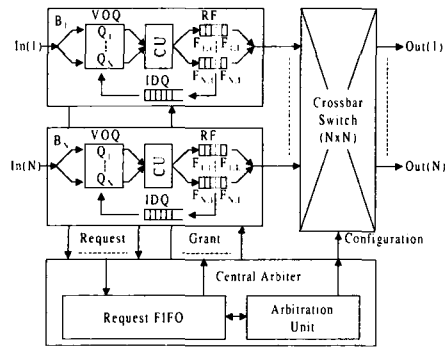


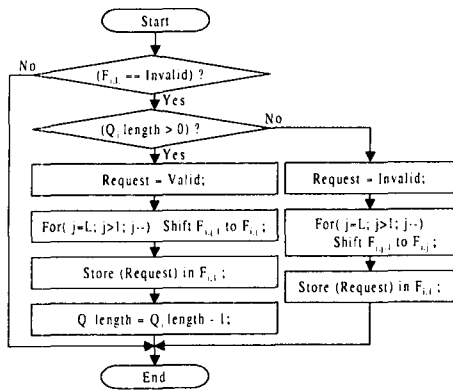
Fig. 1. Advanced switch architecture.

## III. High Performance Routing Engine

The routing engine in an advanced input-queued ATM switch manages VOQs at each input port of the switch. The operations are divided into incoming cell writing, outgoing cell reading, policing, and request control for arbitration latency. Each operation has to be connected in a pipelined manner for wire-speed routing with low cost. Fig. 2 illustrates the architecture of the pipelined VOQ routing engine which consists of a VOQ and idle queue (IDQ) modules, write pointer manager (WPM), read pointer manager (RPM), and request first-in-first-out controller (RFC).

Fig. 2. The architecture of the pipelined routing engine (CF: cell framer, GDI: gigabit data interface, PI: processor interface).

Fig. 3. Pending request management of the request FIFO controller in an input buffer using



high-speed shifters (F: request FIFO element, i: number of output queue, j: number of request bit).

The RFC has request first-in-first-out (FIFO) registers for each VOQ and communicates with a central arbiter. It controls request generation and deletion according to the VOQ and request FIFO status. The RFC generates a valid request signal for each VOQ that has one or more queued cells when the first element of the request FIFO register, of the VOQ, is not occupied with a valid request. It then stores the request signal in the request FIFO register after shifting existing data in the request FIFO, and decreases the VOQ length by 1. The VOQ length in the RFC is smaller than the VOQ length in the policing module (PM) by the number of valid request signals in the request FIFO. If the VOQ length in the RFC is zero and the first element of the request FIFO of the VOQ is not occupied by a valid request, the RFC generates an invalid request signal and just shifts the existing data in the request FIFO. Fig. 3 shows the proposed method of a request management that uses a high-speed shifter.

When a granted output port number arrives at the RFC, from the central arbiter, the RFC deletes the oldest request signal in the request FIFO of the granted output port, and it propagates the granted output port number to the outgoing cell reader (OCR). The OCR propagates the granted output port number to the RPM and PM, and reads an outgoing cell from the ingress buffer memory (INBM) using an output cell address that comes from the RPM. The RPM updates the currently selected port queue of the VOQ module with the next address extracted from the next cell pointer information in the ingress pointer memory

(INPM).

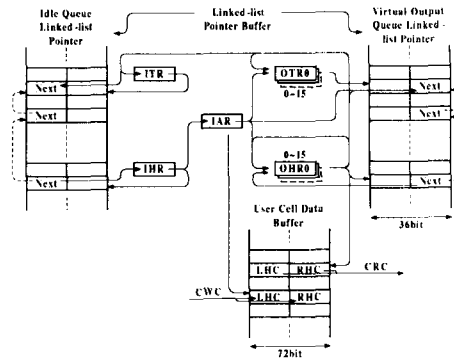
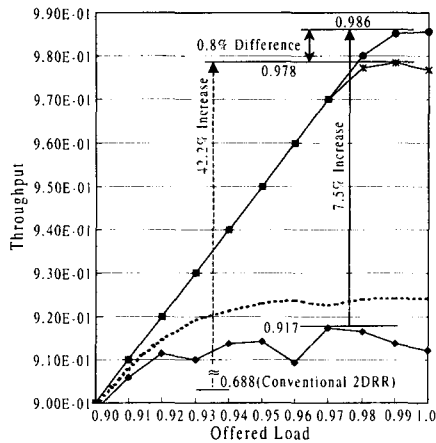


Fig. 4. Entire data path for the pipelined pointer management (OHR: output queue head register, OTR: output queue tail register, IHR: idle queue head register, ITR: idle queue tail register, IAR: idle address register, LHC: left half cell, RHC: right half cell, CWC: current writing cell, CRC: current reading cell).

The RPM stores the current outgoing cell address in the next destination port queue of multicast leaves if the outgoing cell is a multicast cell. A new arriving cell is stored in the INBM by the incoming cell writer (ICW) after receiving a new cell writing address from the WPM. The WPM updates the destination output port queue of the current incoming cell in the VOQ module with the new cell writing address. The RFC aggregates the new cell arriving and multicast cell stitching information from the PM through the backpressure controller (BPC). Fig. 4 shows entire data paths to manage all VOQs and an idle queue that fully share the INPM, a single dual-port synchronous SRAM, including the multicast cell address stitching.

Fig. 5. A comparison of the switch throughput for request shifting and conventional methods in identically distributed Poisson arrivals (switch size: 16x16, input buffer size: 128-cell, request latency: 1 timeslot, grant latency: 1 time slot).

- Pipelined iSLIP (3-iteration and 2-bit request FIFO).
- \*--- Pipelined 2DRR (2-bit request FIFO).
- ◆--- Conventional iSLIP (3-iteration).
- ..... Output queuing.



#### IV. Performance

In conventional arbitration of plural requests, the simulation result shows that the highest saturation throughputs are 0.688 for 2DRR and 0.917 for iSLIP with a dynamically allocated 128-cell buffer. In the proposed arbitration of requests using the request shifting method, a significant performance improvement is achieved. With the proposed method, the highest saturation throughputs are 0.978 for 2DRR and 0.986 for the iSLIP with the same input buffer size and buffer allocation scheme as in the conventional arbitration, when each transmission latency for request and grant is one cell slot and the request FIFO length is 2. The average throughput as a function of the offered load for uniform, independent, and identically distributed Poisson arrivals is shown in Fig. 5.

#### V. Conclusions

In this paper, we proposed a new architecture of the VOQ routing engine for an advanced input-queued ATM switch. The proposed routing engine provides wire-speed routing at low cost using a pipelined buffer management. A novel request shifting method to manage the transmission latency of request and grant data between input buffers and central arbiter is adopted to the design of the proposed VOQ routing engine. The wire-speed pipelined VOQ routing engine has been implemented in a FPGA and has been tested

for all functions in a core ATM switch system which has 16x16 switch size, 2.5Gbit/s port speed, and 40Gbit/s aggregated switching capacity.

#### References

- [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. McKeown, 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.
- [6] G. J. Jeong, J. H. Lee, and B. C. Lee, "Design of pipelined routing engine for input-queued ATM switches," *Electron. Lett.*, vol. 37, no. 2, pp. 137-138, Jan. 2001.