ATM 망에 적용 가능한 출력단 버퍼형 Batcher-Banyan 스위치의 성능분석

Performance Analysis of Output Queued Batcher-Banyan Switch for ATM Network

류걸우*, 이규호*

Keol-Woo Yu, Kyou Ho Lee

Abstract

This paper proposes an ATM switch architecture called Output Queued Batcher-Banyan switch (OQBBS). It consists of a Sorting Module, Expanding Module, and Output Queueing Modules. The principles of channel grouping and output queueing are used to increase the maximum throughput of an ATM switch. One distinctive feature of the OQBBS is that multiple cells can be simultaneously delivered to their desired output. The switch architecture is shown to be modular and easily expandable.

The performance of the OQBBS in terms of throughput, cell delays, and cell loss rate under uniform random traffic condition is evaluated by computer simulation. The throughput and the average cell delay are close to the ideal performance behavior of a fully connected output queued crossbar switch. It is also shown that the OQBBS meets the cell loss probability requirement of 10^{-6}.

Key Words : ATM Switch Architecture, Output Queued Batcher-Banyan switch, performance, average cell delays, throughput, cell loss probability, cell re-circulation

* 한국전자통신연구원(ETRI) 라우터기술연구부

1. Introduction

Broadband Integrated Services Digital Network (BISDN) provides integrated services for diverse traffic types[1]. These traffic types range from conventional telephone to wideband services involving images and video. The Asynchronous Transfer Mode (ATM) was selected by the ITU [2,3] as the universal transport mechanism of BISDN services.

ATM is a cell-based switching and multiplexing technique in which data is segmented into 53-octets of fixed-size cells and cells from active sources are statistically multiplexed. According to ITU-T standards and ATM Forum specifications, interface of 155.52 and 622.62 Mbits/sec (SONET OC-3 and OC-12) have been recommended. At these rates, cells are transported at the rate of 350,000 and 1,460,000 cells/sec. In addition, a switch must meet the Quality of Service (QoS) requirements which are specified in terms of the cell loss probability (CLP), cell delay and delay jitter, and throughput. Furthermore, an ATM switch must accommodate a large number of input/output links.

The various architectures for fast packet switches are surveyed in [4,5,6]. However space-division switches based on Banyan networks[7] suffer from two fundamental problems which can limit switch performance. These are the *internal blocking* and the *output contention*. The internal blocking occurs when two input cells have different destination addresses and they are to be routed to the same output port of a switching element(SE). In such a case, one of the cells is either lost or routed to an incorrect destination. The output contention, on the other hand, occurs when two or more cells have the same destination address and are to be routed to the same output port of an SE. Since each destination port can accept only one cell at a time, the other cells will be blocked in the switch.

The internal blocking problem can be avoided by using the Batcher sorting network[8] followed by Banyan switch. Throughput of such a switch is limited to 58%, because of the output contention problem. Considering the statistical nature of traffic presented to the switch[11], queueing methods offer a viable approach to the output contention problem. There are several queueing methods used in ATM switching systems[9]. Examples are the input queued Batcher-Banyan switch with a three-phase algorithm[13] and the Sunshine switch[14]. In the Sunshine switch, the output contention problem is handled by a trap network and K planes of Banyan networks with cell re-circulation. In this approach, up to K cells for each output can be delivered to their destination at the same time, and the remaining cells are fed back to the input side of the sort network, i.e., input queue, for re-entry during the next cell-cycle. However, when cells are recalculated, they may arrive out of sequence at their output. This situation requires a cell re-sequencing mechanism to handle out-of-sequence cells.

As the performance of input and output queueing methods are analyzed in [10,11,12], the output queueing method is shown to have better performance than input queueing method. The Knockout switch[15,16] is an example of a switch which has adopted the output queueing method. One major drawback of this switch is its hardware complexity. The cross-point complexity grows as the square of the switch size. This hardware complexity makes difficult to implement it in large-scale switching systems.

In this paper, we propose a space-division switch called Output Queued Batcher-Banyan Switch (OQBBS)[20]. As described in[19], the design objectives of ATM switch architecture are the enhancement of the throughput and the reduction of the hardware complexity while retaining modularity. In the OQBBS, the design

objectives are met by partitioning the switch fabric in three parts: Sorting Module(SM), Expanding Module(EM), and Output Queueing Modules (OQMs). It is based on the grouping of all N outputs into N/C output groups, where N is number of inputs/outputs and C is the size of output group. Cells from the SM are sorted in the ascending order of cells output group. The EM with expansion ratio K provides K alternative self-routing paths to each output, and cells through the OQM are queued in their destined output queue. If the EM is able to deliver K cells to the same output with the group size of C, then each output buffer can be connected to $K \times C$ outputs of the EM. Since multiple concurrent paths exist to the same output and are shared by all outputs in the same output group, output contention can be minimized. Therefore, the throughput and average cell delay of the OQBBS approaches the ideal performance of fully connected output queued system while reducing the cross-point complexity.

This paper is organized as follows. The architectural description of the OQBBS is presented in section 2. In section 3, the performance results obtained from simulation are discussed with consideration of the crossbar and Sunshine switches. Concluding remarks are given in section 4.

2. The Structure of the Output Queued Batcher-Banyan Switch(OQBBS)

The Output Queued Batcher-Banyan Switch (OQBBS) is based on the Batcher sorting network[7] and has a partial output queueing topology. For simplicity, we assume N ($= 2^n$) is the number of inputs/outputs of a switch.

C: Output Group Size
K: Expanding Ratio of EM
N: Number of Inputs/Outputs of OQBBS
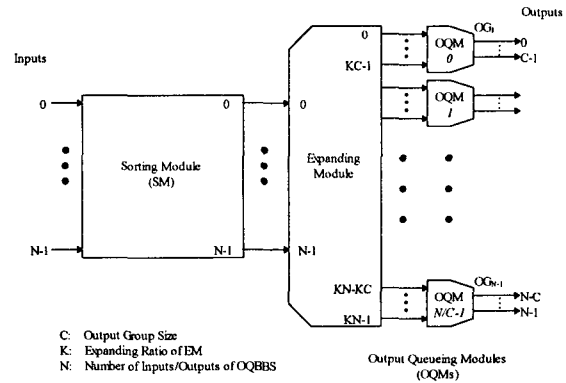
(Figure 1) The $N \times N$ OQBBS Architecture

Figure 1 shows the structure of the OQBBS, which consists of: a Sorting Module (SM), and an Expanding Module (EM), and Output Queueing Modules(OQMs). All N outputs in OQBBS are equally partitioned into small subset of output group(OG) each having $C(= 2^c)$ outputs as shown in the figure. For example, 8 outputs, which are indexed from 0 to 7, can be partitioned into 4 OGs, with 2 outputs for each OG such as {0,1}, {2,3}, ..., {6,7}. Cells are first sorted through the $N \times N$ SM in the ascending order of their output group number. Therefore, cells destined to the same OG are adjacent to each other which creates cell group. Sorted cells are then all moved to the $N \times KN$ EM, where K is the expansion ratio. This is one distinguishing feature over the Sunshine switch, in which cells from the sorting network are examined if there are more than K cells to the same output (K is the number of banyan planes in Sunshine switch) and are re-circulated if necessary. Through the EM, cells are routed to the OQM which corresponds to their OG. The expansion ratio, K, provides K alternative paths to the same output. Therefore total paths to an output group is KC. Output queueing is performed through the OQMs, in which multiple cells to the same output are concentrated at their designated output buffer.

Together with the EM and OQM, each output can accept up to the KC cells at the same time depending on the cells to the other outputs of the same output group. Therefore the OQBBS employ partial output queueing method that can simplify the output queueing complexity.

Figure 2 shows an example of EM which is shown as dashed box for $K = 2$, $N = 8$, and $C = 2$. The internal structure of 8×16 EM forms banyan network configuration with expanded outputs. There are $\log_2 N - \log_2 C = 2$ switching stages with $NK/2 = 8$ switching elements (SEs). Each SE in stage i routes cells to the upper(lower) output port of an SE if their ith most significant address bit is 0(1). Cells from the SM are explicitly partitioned into N/C output groups through the EM. The OQM then collects cells appearing at the right output group and stores them at their destined output buffer until they are transmitted. In addition to, the segmentation and reassembly(SAR) sublayer function can be performed
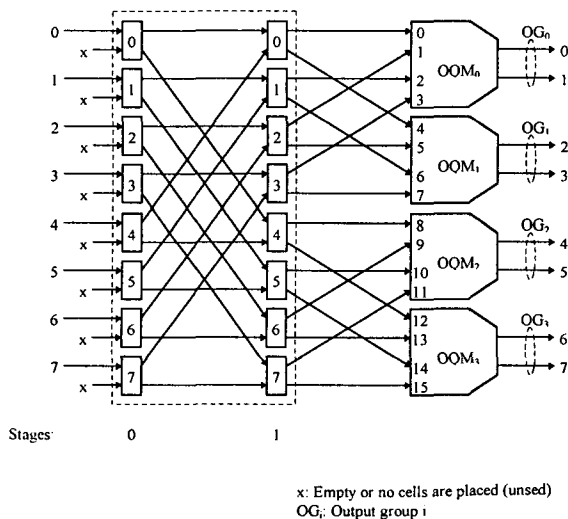
for signaling message and user traffic. The internal structure of OQM has fully interconnected output buffers, so that it can accommodate multiple cells destined to the same output.

Notice when $C = 1$. The resulting switch architecture forms as Starlite ($K = 1$) and Sunshine ($K>1$) switches when cells from the SM are all moved to the EM. Since the value C for these switches is fixed to 1, the upper limit of the number of cells accepted by an output in one cell-time is limited by K. Therefore, for small K, these switches require cell re-circulation mechanism. In the OQBBS, since all C outputs are sharing all paths to the same OQM, it can accept up to KC cells concurrently depending on the number of cells to the other outputs in the same output group. As long as the number of cells to each of all output groups does not exceed KC, no cell will be lost in the EM. From this fact, the cell loss probability of the OQBBS can be decreased not only by increasing K, but also by increasing C. Therefore, in the OQBBS, cell re-circulation is not needed when reasonable values of K and C are chosen, which will be given later in Section 3.

3. Performance Assessment of the OQBBS

In this section we will examine the performance of the OQBBS. The cell arrivals at the input of a switch will be modeled by a Bernoulli process. The probability of a cell arriving during one cell-time is ρ, and the destination addresses are uniformly distributed over N outputs. This traffic pattern will allow comparison to be made with other switch architectures. The time required to service one cell from the output buffer is taken to be one cell-time. In addition, all input cells and output destination addresses are considered to be governed by an identical independent random process.

x: Empty or no cells are placed (unsed)
OG$_i$: Output group i

(Figure 2) An Example of 8 ×16 Expanding Module and 4 ×2 Output Queueing Modules for $N = 8$, $K = 2$, and $C = 2$.
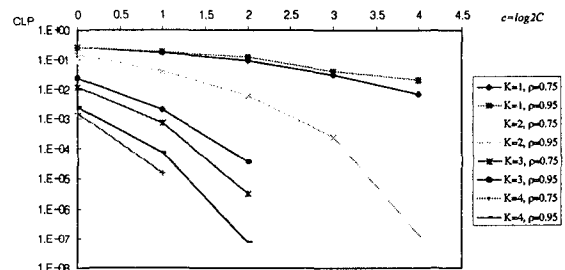
3.1 Cell Loss Probability

The cell loss probability(CLP), throughput, and average cell delay will be examined in this section. In order to maintain a CLP of 10^{-6}, one must determine the influence of varying the expansion ratio K and the channel group size C. The throughput and average cell delay can be evaluated with knowledge of the appropriate parameters range. The ideal performance can be achieved by a fully connected crossbar switch with infinite output buffer. Therefore, the results of the OQBBS can be compared to that of the crossbar and Sunshine switches. Numerical results presented in this section for the performance parameters are obtained by the computer simulation for various values of the switch parameters. The simulation is run for 10^6 cell-times. Over the entire simulation duration, $10^6 \times N$ cells are presented and transmitted by the switch.

In the OQBBS, cell loss is possible through the EM due to the output group contention and output buffer overflow. We focus CLP through the EM, since it serves to limit the performance of OQBBS. The performance analysis for OQM is not included in this paper. However, similar performance analysis for the OQM can be found in [18], which presents an analytical model for the performance analysis and dimensioning of a cell reassembly function block in an ATM switching system. To investigate the upper bound of switch's CLP, we assume buffer sizes are unlimited. The CLP requirement can be met by the various configurations of the OQBBS depending on the combinations of K and C.

Figure 3 shows the cell loss probability versus both an output group size C and the expansion ratio K of EM. The CLP is given for the switch offered loads of 0.75 and 0.95. It is observed that, for a fixed value of K, the CLP exponentially

decreases as C increases. Similar result occur when C is fixed. As we expected, the CLP is shown to increase as the offered load is varied from 0.75 to 0.95. Considering the case $C = 1$ and $K = 1$, the CLP in the figure corresponds to that of the Batcher-Banyan switch.

The CLP for $K \geq 1$ also can be viewed as the CLP of the Starlite($K = 1$) and Sunshine($K > 1$) switch if all cells from sorting network are entered to the Banyan part(s) without going through the trap network. The CLP for $C = 1$ drops slowly with increasing K. This result demonstrates why the Sunshine switch requires a re-circulation method to meet the CLP requirement for values of K less than 8. Furthermore, when C increases, the CLP is also exponentially decreased. Therefore, the advantage of the routing based on the N/C output groups, expanded routing, and shared output queueing topology is evident.

(Figure 3) Cell Loss Probability versus various K, ρ, and $C = 2^c$ for $N = 64$.

Table 1 shows the CLP of OQBBS with $K = 3$ and $C = 8$ for the input sizes of 256 and 1024. It is shown that the switch can meet CLP requirement of 10^{-6} even for the $N = 1024$. In [14], the Sunshine switch can meet the CLP requirement with $K = 3$ and $M/N = 0.1$, while the same requirement can be met by OQBBS without cell-re-circulation. Throughout this section, we only considers a case for $K = 3$ and $C = 8$.

(Table 1) Cell Loss Probability of the OQBBS with $K = 3$ and $C = 8$.

	$N = 256$	$N = 1024$
$\rho = 0.55$	0.0	0.0
$\rho = 0.75$	0.0	1.0×10^{-8}
$\rho = 0.95$	6.5×10^{-8}	1.0×10^{-7}

The throughput is defined as the cell transmission rate at the switch output over arrival rate ρ to the switch. It can be defined as a function of ρ and is denoted by $T(\rho)$. Normalized throughput is then $T(\rho)/\rho$. As known that the throughput of crossbar switch with infinite buffers remains constant with the offered load, its normalized throughput becomes 1.0 over offered load which is optimal. This can be achieved with the OQBBS. The throughput of OQBBS is observed to be identical to the crossbar switch. The high throughput is due to the low cell loss probability and output queueing method.

3.2 Average Cell Delay

We assume that common buffers are shared for all inputs to an output, and cells are processed First In First Out(FIFO) discipline through the output buffer. Figure 4 shows average cell delays for both the OQBBS and crossbar switch with unlimited output buffers for $N = 64$. The average cell delay up to the offered load of 0.75 is less than two cell-times. When the traffic load increases from 0.75 to 0.95, the average cell delay increases quickly from 2 to 11.5 cell-time units. The average cell delay for OQBBS is ideal by showing identical results to the crossbar switch. The largest cell delay is also shown in the figure and is denoted by squares.
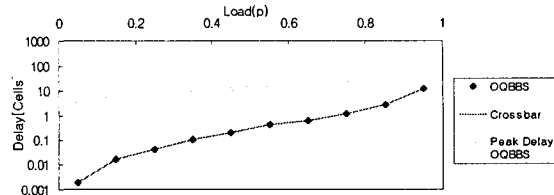
(Figure 4) Average Cell Delay of the OQBBS and Crossbar switch for $N = 64$.

As the cell delay is related to the buffer occupancy, buffer size of 20 is reasonable for a load of 0.75. As shown in Table 2, average cell delay does not change significantly with increasing switch size. This is because the difference in maximum throughput and average cell delay are negligible for switch size beyond 16[10,17].

Therefore, the average cell delay of the OQBBS is not sensitive for input size, but to switch offered loads. Considering both the throughput and cell delay, the OQBBS can provide ideal performance.

(Table 2) Average Cell Delay of the OQBBS with $K = 3$ and $C = 8$.

	$N = 64$	$N = 256$	$N = 1024$
$\rho = 0.55$	0.35082	0.3547	0.35601
$\rho = 0.75$	1.18584	1.19842	1.20198
$\rho = 0.95$	11.37395	11.50720	11.53028

4 Concluding Remarks and Future Work

We have presented Output Queued Batcher -Banyan Switch (OQBBS), which is based on the channel grouping and output queueing method. Under uniform random traffic condition, the OQBBS has been shown to provide very good performance such as the throughput of 1.0 and average cell delay of less than 12 cell-time. The CLP requirement of 10^{-6} can be achieved with $K = 3$ and $C = 8$. The complexity of the OQBBS is

shown to be simple as compared with the Sunshine switch, in which cells are re-circulated.

Further work may include the performance analysis for other traffic patterns such as bursty and hot-spot traffic pattern. The multi/broad-casting functions are also need be studied.

References

[1] Joseph Y. Hui, *Switching and Traffic Theory for Integrated Broadband Networks*, Kluwer Academic Publishers, 1990.

[2] ITU-T Recommendation I.121, "Broadband Aspects of ISDN," Geneva, Switzerland, April, 1991.

[3] ITU-T Recommendation I.150, "BISDN Asyn -chronous Transfer Mode Functional Character -istics," Geneva, Switzerland, December, 1995.

[4] H. Ahmadi, "A Survey of Modern High Performance Switching Techniques," *IEEE Journal on Selected Areas in Communications*, vol.7, no.7, Sep. 1989, pp. 1091-1103

[5] Fouad A. Tobagi, "Fast Packet Switch Architecture for Boadband Integrated Services Digital Networks," *Proceedings of the IEEE*, vol. 78, no.1, Jan. 1990.

[6] Raed Y. Awdeh, H.T. Mouftah, "Survey of ATM switch architecture," Computer Networks and ISDN Systems 27, 1995, pp. 1567-1613.

[7] L. R. Goke and G. J. Lipovski, "Banyan Networks for Partitioning Multiprocessor Systems," in *Proc. 1st Annu. Int. Symp. Comput. Architecture*, Dec. 1973, pp. 21-28.

[8] K. E. Batcher, "Sorting Networks and Their Applications," in *Proc. Spring Joint Comput. Conf.*, AFIPS, 1968, pp. 307-314.

[9] M. G. Hluchyj and M. Karol, "Queueing in Space-Division Packet Switching," in *Proc. INFOCOM 88*, New Orleans, LA, Mar. 1988, pp. 334-343.

[10] J. J. Karol, M. G. Hluchyj, and S. P. Morgan,

"Input versus Output Queueing on a Space-Division Packet Switch," *IEEE Transactions on Communications*, vol. COM-35, Dec. 1987, pp. 1347-1356.

[11] Ilias Iliadis and Wolfgang E. Denzel, "Analysis of Packet Switches with Input and Output Queueing," *IEEE Transactions on Communications*, vol. 41, no. 5, May 1993, pp. 731-740.

[12] Enrico Del Re and Romano Fantacci, "Performance Evaluation of Input and Output Queueing Techniques in ATM Switching Systems," *IEEE Transactions on Communications*, vol. 41, no. 10, Oct. 1993, pp. 1565-1575.

[13] J. Y. Hui and E. Arthrs, "A Broadband Packet Switch for Integrated Transport," *IEEE Journal on Selected Areas in Communications*, vol. SAC-5, Oct. 1987, pp. 1264-1273.

[14] J. J. Giacopelli, J. J. Hickey, W. S. Marcus, W. D. Sincoskie, and M. Littlewood, "Sunshine: A High-Performance Self-Routing Broadband Packet Switch Architecture," *IEEE Journal on Selected Areas in Communications*, vol. 9, no. 8, Oct. 1991.

[15] Y. S. Yeh, M. G. Hluchyj, and A. S. Acampora, "The Knockout Switch: A Simple, Modular Architecture for High-performance Packet Switching," *IEEE Journal on Selected Areas in Communications*, vol. SAC-5, Oct. 1987, pp. 1274-1283.

[16] K. Y. Eng, M. G. Hluchyj, and Y. S. Yeh, "A Knockout Switch for Variable-length Packets," *IEEE Journal on Selected Areas in Communications*, vol. SAC-5, Dec. 1987, pp. 1426-1435.

[17] H. Ahmadi, W. E. Denzel, C. A. Murphy, and E. Port, "A High-Performance Switch Fabric for Integrated Circuit and Packet Switching," *Int. J. Digital Analog Cabled Syst.*, vol. 2, no. 4, 1989, pp. 277-287.

[18] G. Park, S. Kang, C. Han, "Performance Evaluation of a Cell Reassembly Mechanism with Individual Buffering in an ATM Switching System," *ETRI Journal*, vol. 17, no. 1, pp. 23-36, April 1995.

[19] Y. B. Kim *et al*, "An Architecture of Scaleable ATM Switching and Its Call Processing Capacity Estimation," *ETRI Journal*, vol. 18, no. 3, pp. 107-125, Oct. 1996.

[20] Keol W Yu and T S Chung, "Output Buffered Type Asynchronous Transfer Mode(ATM) Switch," *U.S.Patent* 5612951, Mar. 18, 1997.

● 저자소개 ●

류걸우
1983. 2 조선대학교 기계공학과 공학사
1990. 5 미국 University of Massachusetts 대학 Computer Science 석사
1993. 5 미국 University of Massachusetts 대학 Computer Science 박사
1994. 7~현재 한국전자통신연구원(ETRI) 선임연구원
관심분야 컴퓨터 통신, 네트워크

이규호
1980 경북대 전자공학과 공학사
1982 경북대 대학원 전자공학과 공학석사
1998 The University of Gent, Belgium, 정보/컴퓨터공학과 공학박사
1986~1988 미국 AIT Inc, 연구원
1983~현재 한국전자통신연구원(ETRI) 책임연구원
관심분야 System Modeling and Development Methodology, High-end Switched Router System Architecture, ATM-based Network, High-speed Protocol Implementation Technology, Parallel Processing Architecture