# Design and Analysis of a Growable Multicast ATM Switch

Kuochen Wang, *Member, IEEE*, and Ming-Howe Cheng

***Abstract*—This work designs and analyzes a cost-effective growable multicast asynchronous transfer mode (ATM) switch that has a new grouping network structure. The proposed switch can easily be enlarged by using more stages, since both *cell routing* and *contention resolution* are designed to distribute over switch elements. Experimental results indicate that, by allowing valid cells to enter grouping networks from two directions (the west and north sides), the modular ATM switch proposed herein not only meets the ATM performance requirements for both unicasting and multicasting but also uses fewer switch elements and has a shorter cell delay than the ATM switch.**

***Index Terms*—Asynchronous transfer mode, design methodology, large-scale systems, switches.**

## I. INTRODUCTION

TWO approaches can achieve an asynchronous transfer mode (ATM) switch design capable of modular growth [2]: 1) generalization from a single output to group outputs using the Knockout principle [3] and 2) the concept of multistage interconnection networks using the interconnection fabric [4]. For a unicast switch [5], self-routing and output port contention resolution should be performed in a distributed manner so that centralized processing does not become an obstacle to constructing a large switch. For a multicast switch, cell replication and multicast addressing must also be considered. Multicast routing can be performed by broadcasting incoming cells to all output ports and filtering at each output port [6], or by replicating incoming cells into several copies and sending each copy to the corresponding output port [7]. This study proposes a growable multicast ATM switch. The switch is based on the generalized knockout principle. By redesigning *grouping networks* [1], the switch proposed herein uses fewer switch elements than [1] and has a shorter cell delay than [1]. In the following, we design grouping networks and summarize the performance and cost-effectiveness analyses that distinguish the proposed switch from the switch in [1].

## II. DESIGN OF THE PROPOSED ATM SWITCH

### A. Overall Architecture

Fig. 1 illustrates an $N \times N$ two-stage architecture of our proposed growable multicast ATM switch [1]. The switch can be easily extended to more than two stages. It consists of three parts: *input port controllers* $(\mathrm{IPC}_1, \mathrm{IPC}_2)$, *grouping networks* $(\mathrm{GN}_1, \mathrm{GN}_2)$, and *output port controllers* (OPC). Each IPC ($\mathrm{IPC}_1$ or $\mathrm{IPC}_2$) accepts arrival cells, uses the arrival routing information (VCI) to look up the routing table, and attaches new routing information to the front of the cells to route them in the grouping networks. A GN ($\mathrm{GN}_1$ or $\mathrm{GN}_2$) sends multicast cells to broadcast buses, and the cells are delivered to all output groups. The output groups of the GN decide whether to accept these cells or not by checking the multicast patterns. An output port controller stores arrival cells in an output buffer, makes multiple copies with a cell duplicator, looks up the multicasting table to attach new VCI to the front of outgoing cells, and finally sends the cells to output ports.

According to Fig. 1, a set of $M$ output ports of $\mathrm{GN}_1$ forms a group at the first stage [1]. $L_1$ links are dispatched to each output port, and so there are a total of $L_1 \times M$ routing links in each output group. Totally, there are $K$ output groups $(K = N/M)$. Up to $L_2$ cells may be time-division multiplexing. These cells are stored in the output buffer and are then read out sequentially. Adjusting $L_1$ or $L_2$ (called a group expansion ratio) can reach the required cell loss probability. To avoid cells out of sequence, the number of routing links in a group between two grouping network stages in Fig. 1 cannot exceed the number of bits in a cell. To expand the switch, a $\mathrm{GN}_2$ and its associated $M$ output port controllers can be combined into a module, and the module can be replicated $K$ times and integrated with the grouping network $(\mathrm{GN}_1)$ at the first stage as a larger module. In this way, a large modular switch can be constructed by continuing to replicate these modules recursively. Since the four major functions of the switch—cell addressing, cell routing, cell contention resolution, and cell replication—are all handled in a distributed manner, the switch is scalable [1].

### B. A New Grouping Network

Fig. 2 depicts a novel modular structure of the grouping network at the first stage. The grouping network consists of $N$ *skew buffers* (SB) and $K$ *switch modules*. A switch module at the first stage comprises $N$ *valid bit controllers* (VBC) and an $(L_1 \times M) \times (N - L_1 \times M)$ *switch array*, composed of switch elements. When the west's priority is higher than the north's, the switch element is in a toggle state [1]. Otherwise, it is in a cross state. Since the structure and operation of the grouping network are applicable to both stages 1 and 2, this study only discusses the grouping network at stage 1. In Fig. 2, each switch module accepts a maximum of $N$ cells from $N$ broadcast buses and sends out a maximum of $L_1 \times M$ cells through the $L_1 \times M$ links. The routing links are shared by those cells that are sent to the same output group. The routing information attached to an arrival cell may be a multicast pattern. The multicast pattern is a bit map for all output groups in the grouping network. If the $n$th bit of the multicast pattern for some cell is set to 1, the cell is sent to the $n$th output group. If more than one bit of the

---

Paper approved by N. McKeown, the Editor for Switching and Routing of the IEEE Communications Society. Manuscript received June 5, 1997; revised Nobember 9, 1999. This work was supported by the National Science Council, Taiwan, R.O.C., under Contract NSC85-2213-E-009-120. This paper was presented in part at the IEEE INFOCOM'97, Kobe, Japan, April 1997.

The authors are with the Department of Computer and Information Science, National Chiao Tung University, Hsinchu, Taiwan 30050, R.O.C. (e-mail: kwang@cis.nctu.edu.tw).

Publisher Item Identifier S 0090-6778(00)06149-3.

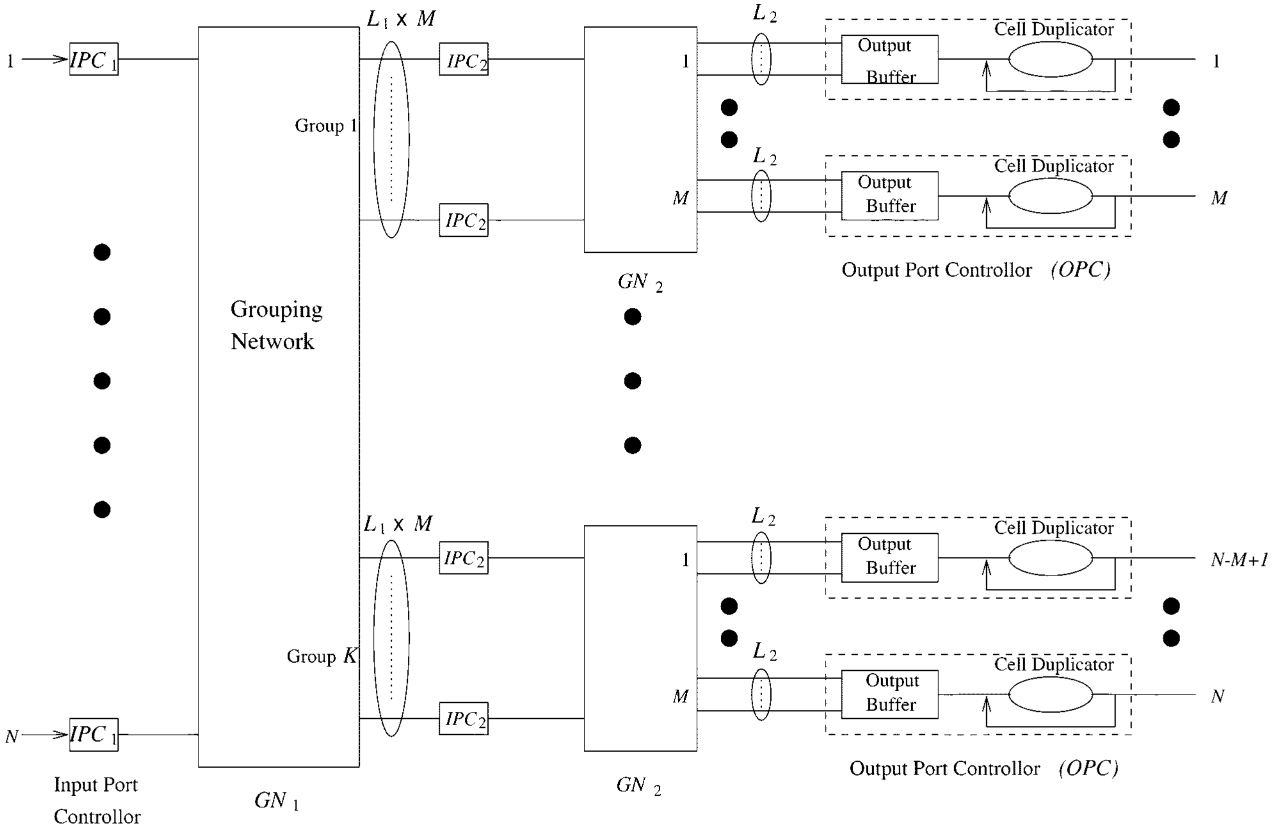

Fig. 1. Overall architecture of our growable multicast ATM switch.

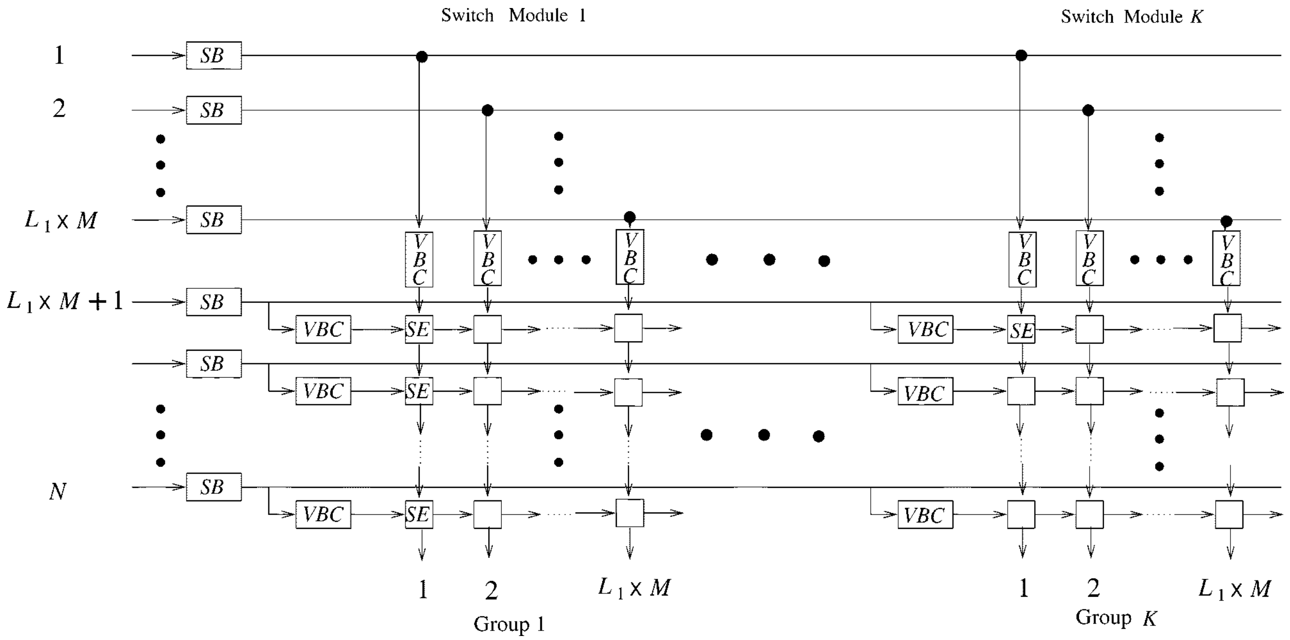

Fig. 2. New grouping network.

multicast pattern is set to 1, the cell is sent to more than one corresponding output group. A VBC monitors whether there is an arrival cell in a time slot. A valid bit $(v)$ and two priority bits $(p_0p_1)$ are used to check whether a cell is valid or not and resolve contention in an output group, respectively. If there is no arrival cell, the VBC sends an empty cell with the lowest priority $(p_0p_1 = 11)$ and the valid bit indicating an invalid cell $(v = 1)$ to the switch array. If there is an arrival cell, but the arrival cell is not intended for this output group, the VBC also sets the valid bit of the cell to 1 and sends it to the switch array. The first $(L_1 \times M)$th VBC's send cells to the north side of the switch array by vertical links. The other VBCs send cells to the west side of the switch array by horizontal links. To adjust cell timing, the cell at $i$th input port is first skewed $(i-1)$ bits as $i \leq L_1 \times M$,

TABLE I

$L_1$ AND $L_2$ FOR DIFFERENT GROUP SIZES $(M)$ WITH $\lambda = 0.9$ AND $\rho = 0.95$

| loss | ratio | $M = 256$ | $M = 128$ | $M = 64$ | $M = 32$ | $M = 16$ | $M = 8$ | $M = 4$ | $M = 2$ |
|---|---|---|---|---|---|---|---|---|---|
| $10^{-9}$ | $L_1$ | 1.3 | 1.45 | 1.7 | 2.05 | 2.65 | 3.55 | 5.05 | 7.55 |
| $10^{-10}$ | | 1.35 | 1.5 | 1.75 | 2.15 | 2.8 | 3.8 | 5.3 | 8.05 |
| $10^{-9}$ | $L_2$ | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 |
| $10^{-10}$ | | 12 | 12 | 12 | 12 | 12 | 12 | 12 | 12 |

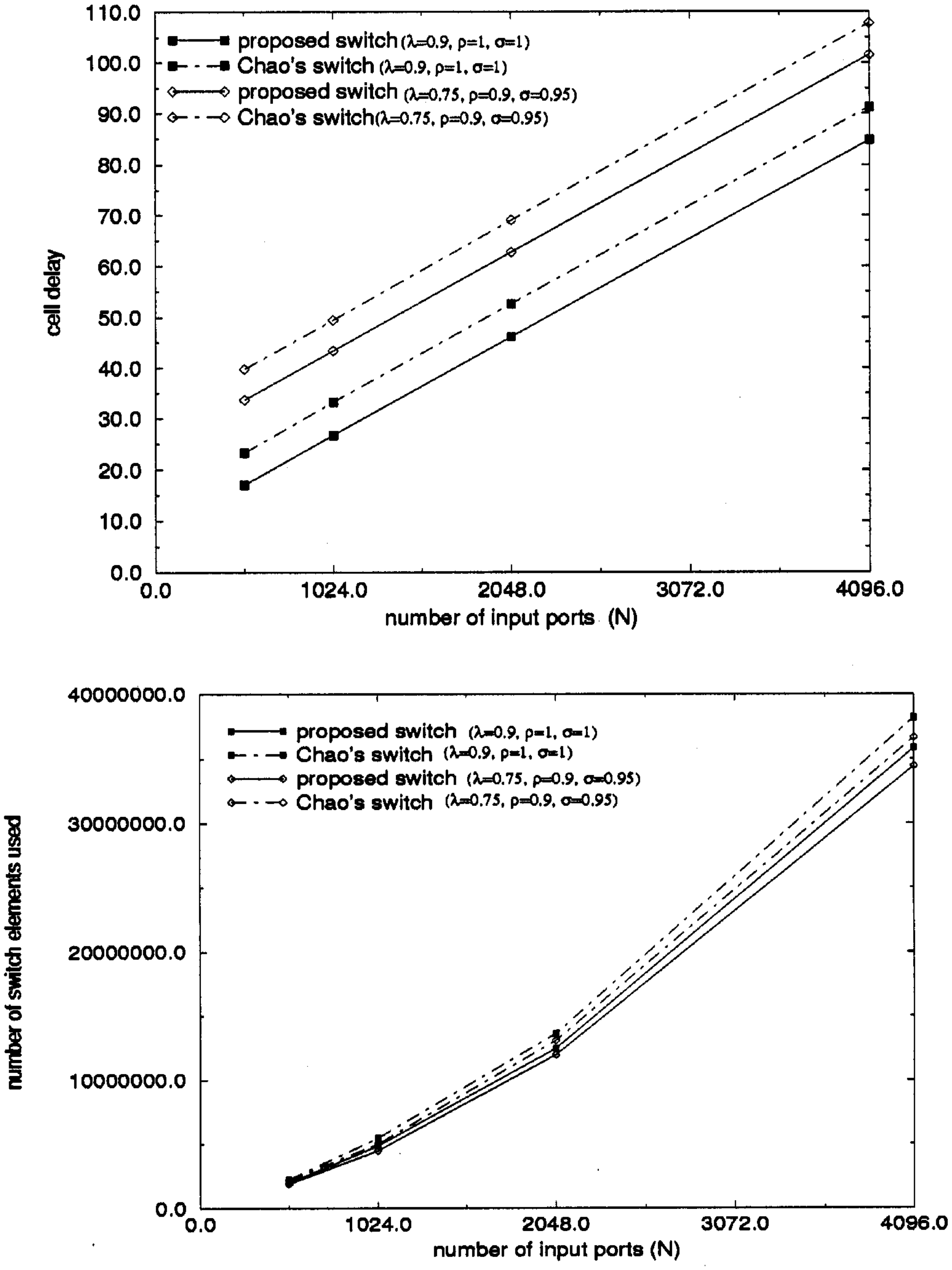

Fig. 3. Comparison with Chao's switch in terms of cell delay and number of switch elements used for $M = 256$.

or $i - (L_1 \times M + 1)$ bits as $i > L_1 \times M$ by using an SB before it is delivered to the corresponding VBC. The grouping network proposed herein and Chao's grouping network [1] differ mainly in that the incoming cells enter our grouping networks from two directions (the west and north sides) instead of one (the west side). The Chao's switch sends empty cells to the north side of the grouping networks. The proposed approach can reduce switch elements and cell delay by a factor of $(L_1 \times M)/N$ compared to the design in [1].

## III. EXPERIMENTAL RESULTS

### A. Practical Combinations of $L_1, L_2$, and $M$

To ensure a low cell loss probability, appropriate combinations of $L_1, L_2$, and $M$ must be chosen. By assuming that the number of arrival cells for an IPC has a Possion distribution with a rate $\lambda$, the number of output groups in the two grouping networks $(\text{GN}_1, \text{GN}_2)$ to which an incoming cell is multicast has a geometric distribution with parameter $\rho$, and an OPC has geometric distribution service time with parameter $\sigma$. A cell may arrive at an input port at any time (in continuous time), accounting for why Possion arrival is used for each input port. It takes a fixed time for an output port to process a cell. A cell that must be multicast to $N$ virtual channels in an output port is duplicated $N$ times for the $N$ virtual channels, accounting for why geometric service (discrete time) is used for each output port. Given a cell loss probability of $10^{-9}$ and $10^{-10}$, respectively, Table I lists the suitable combinations of $L_1$ and $L_2$ for different group sizes $(M)$ with $\lambda = 0.9$ and $\rho = 0.95$. For example, when $M = 256$, $L_1$ should be 1.35 and $L_2$ should be 12 to satisfy the cell loss probability of $10^{-10}$. It is always desirable to choose a small $L_1$ and $L_2$ for a large group size $(M)$ to minimize the total number of switch elements and reduce the number of links. Table I also shows that $L_2$ is always 11 and 12 for the cell loss probability of $10^{-9}$ and $10^{-10}$, respectively. In addition, the larger the $\lambda$ implies a larger $L_1$, the smaller the $\rho$ implies a larger $L_1$. Notably, in the event of a tie of priorities, the switch element is in a cross state. That is, the switch favors upper input ports, which is unfair. Properly adjusting the routing link expansion ratios $L_1$ and $L_2$ can make the cell loss probability of each input port arbitrarily small to satisfy the performance requirement of all broad-band services [1]. In this way, the issue of fairness can be resolved. Another means of resolving unfairness is to add a circular shifter (like a barrel shifter) at the front of a grouping network to rotate the positions of input ports entering the grouping network on a per-cell basis.

### B. Comparison with Chao's Switch

This study used a two-stage switch architecture to compare its switch with Chao's switch [1] in terms of cell delay and number of switch elements used. Because the cell delay in a VBC resembles that in a multicast pattern masker (MPM) of Chao's switch [1], this study only compares the main cell delay resulting from other components. In addition, only the number of switch elements used is compared since the complexity of an address broadcaster along with the MPM's in Chao's switch compares to that of the VBC's in this study's switch. The comparison uses two traffic conditions: 1) unicasting for $\lambda = 0.9$, $\rho = 1$, and $\sigma = 1$ and 2) multicasting for $\lambda = 0.75$, $\rho = 0.9$, and $\sigma = 0.95$. For $M = 256$, as shown in Fig. 3, the proposed switch always reduces 5–7 cell delays more than Chao's switch for both unicasting and multicasting. For example, when $N = 1024$, this study's switch reduces 6.50 cell delays for unicasting and 6.02 cell delays for multicasting. The switch proposed herein always performs better than Chao's switch in terms of cell delay under different group sizes and different traffic conditions. For $M = 256$, according to Fig. 3, the proposed switch reduces the number of switch elements used by 12% more than Chao's switch when $N = 1024$ and by 14% when $N = 512$, both for unicasting and multicasting. Notably, the proposed switch can reduce more switch elements than Chao's switch as $N$ becomes closer to $M$.

## IV. CONCLUSION

This study has presented a recursive modular architecture for a growable multicast ATM switch that has a novel and efficient structure of grouping networks. Experimental results demonstrate that the proposed ATM switch meets the ATM performance requirements either for unicasting or multicasting, is also is very cost effective due to using fewer switch elements (implying reductions in cost, size, weight, and power consumption), and finally has a shorter cell delay than Chao's switch under various group sizes and traffic conditions (unicasting and multicasting).

## ACKNOWLEDGMENT

The authors would like to thank the anonymous reviewers for their valuable comments that improved the quality of this letter.

## REFERENCES

[1] H. J. Chao and B. S. Choe, "Design and analysis of a large-scale multicast output buffered ATM switch," *IEEE/ACM Trans. Networking*, vol. 3, pp. 126–138, Apr. 1995.

[2] K. Y. Eng, M. J. Karol, and Y.-S. Yeh, "A growable packet (ATM) switch architecture: Design principles and applications," *IEEE Trans. Commun.*, vol. 40, pp. 423–430, Feb. 1992.

[3] M. J. Karol and C.-L. I, "Performance analysis of a growable architecture for broadband packet (ATM) switching," *IEEE Trans. Commun.*, vol. 40, pp. 431–439, Feb. 1992.

[4] J. Hui and T. H. Lee, "A large scale ATM switching network with sort-banyan switch modules," presented at the GLOBECOM'92 Conference Record, vol. 1, Dec. 1992.

[5] Y. S. Yeh, M. G. Hluchyj, and A. S. Acampora, "The Knockout switch: A simple modular architecture for high performance packet switching," *IEEE J. Selected Areas Commun.*, vol. SAC-5, pp. 1274–1283, Oct. 1987.

[6] M. J. Karol and M. G. Hluchyj, "The Knockout switch: Principles and performance," in *Proc. 12th Conf. Local Comput. Networks*, Oct. 1987, pp. 16–22.

[7] X. Liu and H. T. Mouftah, "Design of a high performance nonblocking copy network for multicast ATM switching," *Proc. IEE Commun.*, vol. 141, no. 5, pp. 317–324, Oct. 1994.