Shared Buffer Memory Switch for an ATM Exchange

Noboru Endo, *Member, IEEE,* Takahiko Kozaki, *Member, IEEE,*
Toshiya Ohuchi, Hiroshi Kuwahara, and Shinobu Gohara

***Abstract*—This paper proposes an asynchronous transfer mode (ATM) switch architecture called a shared buffer memory switch, whose output cell buffers are shared among all the output ports of the switch. Buffer sharing can reduce the amount of hardware compared with that of a separated buffer memory switch. Moreover, modifying the memory control circuits of the switch makes the memory switch flexible enough to perform functions such as priority control and multicast.**

Experimental measurements and a discussion about the traffic characteristics of switch architecture are carried out to determine how much buffer memory will be reduced through buffer sharing under various traffic conditions and to roughly estimate how many buffers are needed for the switch to meet certain requirements. The resultant estimate shows that buffer sharing reduces the necessary buffer memory to less than 1/5 of what would otherwise be required, and the required buffer size is about 128 cells/output for a 32 × 32 switch when considering bursty traffic conditions.

LSI implementation is also discussed to show that a 32 × 32 switch can be composed of about 12 chips mounted on one printed board.

I. Introduction

ASYCHRONOUS transfer mode (ATM) is a promising technique for transferring and switching various kinds of information, such as telephone speech, data, and motion video, in broadband ISDN, according to the studies of CCITT [1] and others. It can improve the utilization efficiency of transmission and switching equipment by statistically multiplexing fixed length packets (cells) of transferred information on a broad band transmission line (150 or 600 Mbps). The major areas of ATM study are switch architecture, protocol structure and traffic control technique.

This paper proposes a new ATM switching architecture named the shared buffer memory switch. In this switch, the cell buffer memory for the output-queue is shared among all the switch output ports and can be allotted to any output port. As a result, the switch has the potential to reduce the amount of memory required and to be easily implemented in LSI's.

A traffic performance analysis and experimental measurement show the memory reduction effect and give a rough estimation of the buffer size needed for various traffic conditions. The memory used for a cell buffer must be read or written so fast that it has to be implemented in an on-chip memory, which has a severely limited capacity. Therefore, reducing buffer memory or improving buffer memory utilization is very important for ATM switch implementation in an LSI. In addition, such functions as priority control and multicast, which must be given to an ATM switch in some cases, tend to increase the required memory size. The shared buffer memory switch will easily meet such requirements.

Switches implemented in LSI's can also be easily modified in various ways without inherent changes in the basic switch architecture.

Paper approved by the Editor for Communications Switching of the IEEE Communications Society. Manuscript received March 23, 1990; revised July 13, 1990. This paper was presented in part at the IEEE GLOBECOM '90, San Diego, CA, December 1990.

N. Endo, T. Kozaki, and T. Ohuchi are with Central Research Laboratory, Hitachi Ltd., Tokyo 185, Japan.

H. Kuwahara is with Kokusai Electric Co., Ltd., Tokyo 190-11, Japan.

S. Gohara is with Telecommunications Division, Hitachi Ltd., Kanagawa 244, Japan.

IEEE Log Number 9201636.

TABLE I
ATM Exchange System Requirements

Item	Requirement
Application	Public/Private Network
Line Bit Rate	150/600 Mbps
Line Capacity	$>1,000$ ports
Throughput	>150 Gbps
Unit Switch Cap.	>32 ports
Line Utilization	$>70\%$
Cell Loss Prob.	$<10^{-9}$
Service Class	Optional
Multipoint Connection	Optional

II. Requirements for an ATM Switch

The switch architecture requirements for an ATM exchange in the early 1990's are shown in Table I. In the first stage of broad-band ISDN, the line bit rate of 150 Mbps will primarily be used. This bit rate makes the transmission of almost all broad-band information, including HDTV motion-pictures, possible. In a B-ISDN public network, an ATM exchange may be required to accommodate at least 1000 ports for subscriber line or trunk connection. This means the total throughput must be more than 150 Gbps.

To achieve such a large capacity, a multistage switching network must be adopted [2]. A 3-stage-switching network accommodates about 1000 ports. In this case, the unit switch must have more than 32 input/32 output port capacity.

A priority control function is necessary for meeting the different cell loss ratio requirements and switching delay requirements of different service classes. In addition, a multicast function is necessary for broadcasting and multimedia teleconferencing.

III. ATM Switch Architecture

There are three types of ATM switches, an input buffer switch, an output buffer switch and a memory switch. Gener-

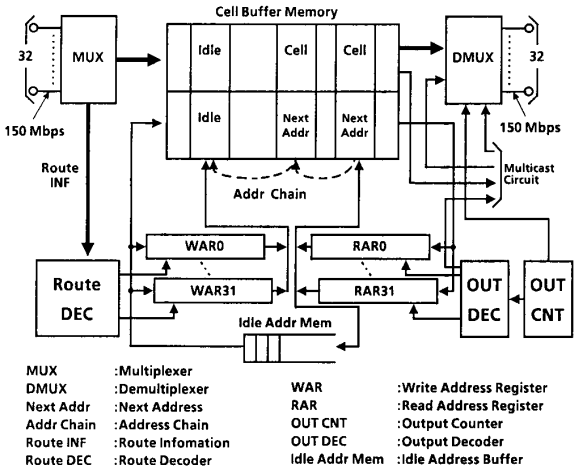


Fig. 1. Shared buffer memory switch.

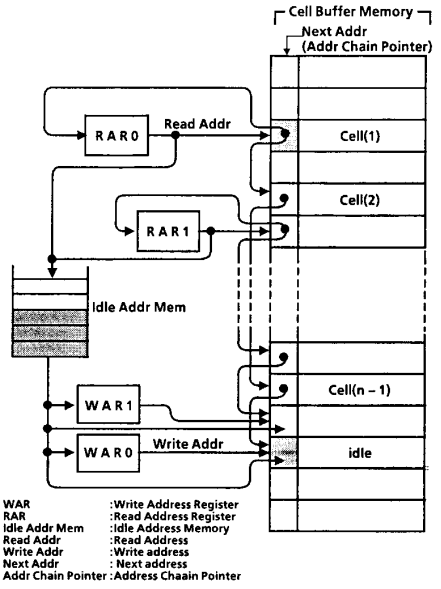


Fig. 2. Output queue chain.

ally speaking, an ATM switch consists of two components: a switching element, through which ATM cells are transferred from an input port to an output port, and a buffer memory, in which the cells are stored for queueing.

Input buffer switches, whose buffers for queueing are in front of the switching element, use less than 60% of their full switching element bandwidth because of Head of Line (HOL) blocking [3]. Output buffer switches, whose buffers are at the rear of the switching element, and memory switches, in which the sequence change between memory-write and memory-read carries out both switching and buffering, can achieve higher utilization efficiency because they do not have degradation factors, such as HOL blocking.

Furthermore, a memory switch [4] can be composed of less hardware than the output buffer switch [5]. This is because the multiplexed cell handled in the memory switch architecture uses the switching element circuit part and control part efficiently. Moreover, modifying the memory read/write control circuit makes the memory switch flexible enough to perform functions such as priority control and multicast. Buffer sharing among the switch output ports can further reduce the total amount of cell buffer required for a memory switch.

Therefore, a shared buffer memory switch architecture is the most practical ATM switch architecture [6]. The high-speed circuit operation requirement imposed on the switch hardware in return for multiplexing can be overcome by submicron CMOS technology and bit-parallel circuit logic.

A. A Shared Buffer Memory Switch Architecture

The shared buffer memory switch proposed in this paper is depicted in Fig. 1. The address chain corresponding to the output is composed of the memory addresses of the buffered cells to a particular output. This chain operates as the output queue. The switch operation is explained as follows.

Before an ATM cell is input into the switch, its data is converted from a serial format to a parallel format, its header is replaced and the routing information, which shows the destination output, is attached to the cell. Subsequently, a cell from an input is multiplexed in MUX with other cells from all the other inputs and is written one-by-one into a cell buffer memory. The writing address is obtained from the WAR. The WAR is designated by the routing information decoded by a route DEC and corresponds to the destination output port of the cell. The route DEC receives the routing information synchronously with the cell data output from the MUX to the cell buffer memory.

An idle address of a cell buffer memory is simultaneously read from an idle address buffer memory (Idle Addr Mem), written into the next address (Next Addr) part of the cell buffer memory in the same address the cell is written in, and stored in the WAR overwriting the former WAR content. This address indicates the writing address of the next cell, which will be input into the output queue corresponding to the WAR. This operation constructs an address chain corresponding to an output and this chain acts as a cell queue for the output. Two address chains of the output queue are shown in Fig. 2. WAR0 and RAR0 correspond to an output port, indicating the end and beginning of the chain, respectively, and WAR1 and RAR1 correspond to another output port. The end of each chain indicates the address into which the next input cell to the output queue will be written.

When a cell is output from the switch, OUT CNT cyclically indicates all the output ports for demultiplexing of the cells read out from the cell buffer memory. The output cell is read out from the cell buffer memory, which has the read address obtained from the RAR indicated by the OUT CNT through the OUT DEC, and is sent out through DMUX. The content of the RAR is also stored in the Idle Addr Mem because this address becomes idle and is replaced by the next chain address read from the Next Addr part of the cell buffer memory. The Idle Addr Mem has a FIFO (first-in–first-out) function and stores idle addresses of the cell buffer memory.

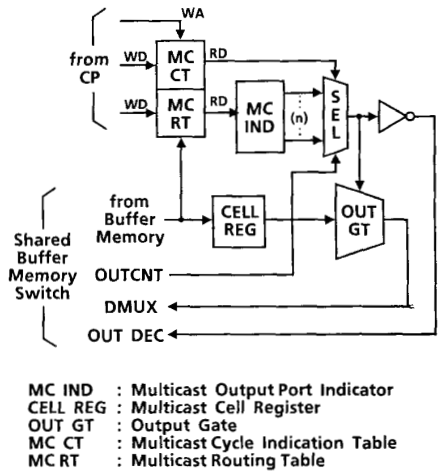


Fig. 3. Multicast circuit.

This mechanism allows temporarily allotment of each cell buffer memory address to any output port as the occasion demands. The next chain addresses for output queues are stored in the same cell buffer memory as the ATM cells, so that the memory for the output queues can be shared among the switch outputs. Because of this dynamic allotment, we call this switch architecture the shared buffer memory switch. This architecture can reduce the required buffer memory quantity through buffer sharing among the output ports, as discussed in Section IV.

B. Priority Control for Various Service Classes

The various types of information handled by an ATM switch have different requirements for switching delay and cell loss ratio. To meet these requirements, an ATM switch has to provide a priority control function that enables it to handle a cell according to the service class to which it belongs. The shared buffer memory switch can easily be modified to handle multiple service classes by adding a WAR and an RAR to correspond to each combination of service class and output port, respectively.

The buffer memory capacity increase caused by added service classes can be reduced due to buffer sharing among the output ports and service classes. If the length of the output queue for a certain service class must be restricted, an up/down counter for counting the cells in the queue can be added. When the number indicated by the counter exceeds a predetermined threshold, a new input cell will be inhibited from putting into the output queue.

C. Multicast Function

A multicast function, used for various distributions or conference services, can be implemented by adding a multicast circuit to a shared buffer memory switch, as shown in Fig. 3. The function of the circuit is as follows.

Multicast cells form a multicast queue in the cell buffer memory irrespective of their destination output. This queue has the same function as an output queue and has a WAR and an RAR. When a multicast cell is sent out from the switch, it is read from the multicast queue chain and stored in a CELL REG.

The cell header is sent to an MC RT to be converted into cell routing information, i.e., the output port numbers to which the cell has to be multicast. The MC RT is a table that stores the multicast routing information for each virtual channel identifier (VCI) and is read using the VCI as the address. The read-out-datum, which is a bit-pattern of each bit corresponding to each switch output, is stored in an MC IND. The routing information corresponding to a VCI is written in the MC RT by a call processor at the time of call set up. An OUT CNT designates each bit in the MC IND by a cyclic one-by-one round trip of all MC IND bits.

When the multicast cell is output, and inverted output of an SEL inhibits the cell buffer memory read out corresponding to the output number indicated by the OUT CNT, and when the multicast cell does not need to be sent out, the SEL output enables the cell buffer memory to be read. When the OUT CNT has passed through all the output ports, multicasting of the cell in the CELL REG is completed and the next multicast cell is read out from the cell buffer memory.

MC CT designates whether or not multicasting must be executed in each cycle of the OUT CNT round trip. The frequency of the multicast execution cycle is determined by the call set up program at the time of bandwidth allocation to a VCI.

Because a multicast cell is copied right when the cell is sent out, the copied cells do not occupy the cell buffer memory or reduce its utilization efficiency. The throughput of the multicast connection is limited to $2/N$ (where N designates the number of switch output ports) in the worst case because only one cell can be multicast during each cycle of the OUT CNT round trip. If an application requires an improvement in the multicast connection throughput, the multicast circuit can be modified.

IV. Traffic Performance of the Shared Buffer Switch

This section estimates the traffic performance of the ATM switch from parameters such as cell loss ratio and switching delay [7]. Of these parameters, the cell loss ratio is more important because high speed transmission in ATM switching makes the switching delay very small and negligible. For example, a 99% cell delay is less than 20 μs under a random traffic condition and with a utilization of 0.8. For the estimation, the cells arrived at the switch randomly or in bursts, and were then either uniformly distributed to all of the switch outputs (balanced model) or concentrated to a few switch outputs (imbalanced).

A. Performance Evaluation Methods for an ATM Switching System

There are three methods for evaluating the cell loss ratio of an ATM switch: theoretical analysis, computer simulation, and hardware experiments. However, estimating small cell loss ratio values from 10^{-7} to 10^{-10} is a problem. If we can make a well-formed analytical model and solve it, then

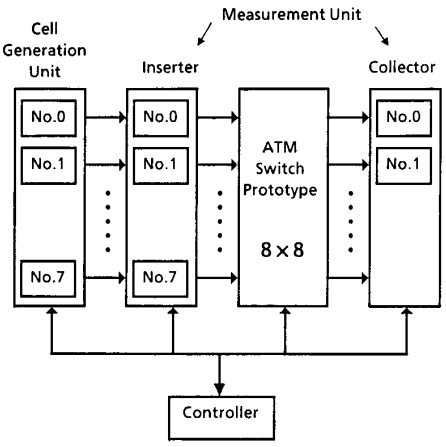


Fig. 4. Experimental system configuration.

the analytical method, whose underlimit is restricted by the computing precision, can calculate cell loss ratios less than 10^{-9}. However, we can only use restricted models under bursty traffic conditions. In addition, shared buffer memory switches require a large amount of calculation because the convolution of each output queue length has to be calculated to evaluate the cell loss ratio of the switch. Furthermore, it is difficult to analyze the cell loss ratio under imbalanced traffic conditions.

On the other hand, in the computer simulation, we can easily change the model, and it is relatively easy to evaluate the performance under various conditions (e.g., bursty or imbalanced traffic conditions). However, the computing power is insufficient to evaluate cell loss ratios less than 10^{-6} [8].

In the hardware experiment, although there are also a few problems with switch size and variety of traffic conditions, evaluation of the cell loss ratios less than 10^{-9} can be achieved under several kinds of bursty traffic conditions and under imbalanced traffic conditions. Therefore, we primarily used the hardware experimental method to evaluate the cell loss ratio.

The configuration of the experimental ATM switching system is shown in Fig. 4. The system is composed of four parts: a shared buffer memory switch, a cell generation unit, a measurement unit, and a controller.

1) Shared Buffer Memory Switch: The main items of this prototype switch are shown in Table II. The switch has three logical queues which correspond to service classes. It also has up/down counters that indicate the number of cells queued in the shared buffer memory belonging to each combination of output ports and service classes, and can prohibit cells from exceeding a queue length threshold for each combination. It can offer any cell loss ratio to any service class by setting a threshold for each service class. It can also emulate a separated buffer memory switch by setting a threshold for each output port.

2) Cell Generation Unit: The cell generation unit is composed of cell generation circuits and cell multiplexers, and generates cells to evaluate switch performance (see Fig. 5). There are five cell generation circuits in the unit, each of

TABLE II
EXPERIMENTAL SWITCH SPECIFICATIONS

Item	Specification
Architecture	Shared Buffer Memory Switch
Size	8×8
I/O Bit Rate	155.52 Mbps
Priority Class	3(C1, C2, C3)
Priority Control	(1) C1 > C2 > C3 (2) C1 = C2 > C3
Total Buffer Size	2/048 Cells

which generates cells for up to 5 VCI's (virtual channel identifier). The cell multiplexer statistically multiplexes the cells generated by the cell generation circuits, as shown in Fig. 5.

Cell generation is controlled by a statistical generation model which is characterized by the burst length n_i the burst interarrival time t_i and the cell interarrival time C (see Fig. 6). The values of n_i and t_i can be selected from several probability distributions (e.g., uniform distribution or geometric distribution), which are generated using conversion tables from m-sequence random numbers of period 2^{47}. The value of C is chosen at the beginning of each experimental run. Since both n_i and t_i are random variables, it is possible that Cn_i may be greater than t_i, and the tail of the ith burst may overlap with the $(i+1)$st burst. In this case, the $(i+1)$st burst is delayed by $(Cn_i - t_i)$ units of time.

3) Measurement Unit: The measurement unit is composed of inserters and collectors, as shown in Figs. 4 and 5. The cell format is shown in Fig. 7. The inserter sets the sequence number and the time stamp for the arriving cells. The collector calculates the total number of cells discarded in the switch by detecting gaps in the sequence numbers of the arriving cells. The collector also computes the queuing delay in the buffer memory using the difference between the current time and the time stamp of the arriving cells.

4) Controller: The controller is a workstation that sets the routing and input traffic distribution data for the experimental system, collects the data (e.g., cell losses) from the experimental system, manages the sequence of events, and calculates cell loss ratios and queueing delays. The control program is written in the C language.

B. Cell Loss Ratio Evaluation Under Balanced Input Traffic

This section and the next section describe results obtained from the experimental ATM switching system in Fig. 4.

1) Cell Loss Ratio for Random Traffic Conditions: In this experiment, cells were randomly generated, and were uniformly distributed to the output ports of the switch. Fig. 8 shows the relationship between the cell ratio and the buffer size of an ATM switch with a line utilization rate $\rho = 0.8$. The buffer size was normalized by the number of ATM switch output ports. Ten samples were measured to calculate the 99% confidence interval, and 10^9 cells per line were generated in each sample.

On the separated buffer memory switch, the logarithm of the cell loss ratio decreased linearly with the buffer size. This result agrees with the calculation using the $M/D/1/k$ model (broken line in the Fig. 8) [9].

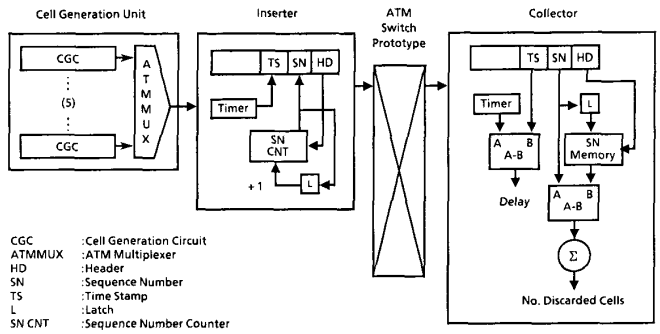


Fig. 5. Measurement unit function.

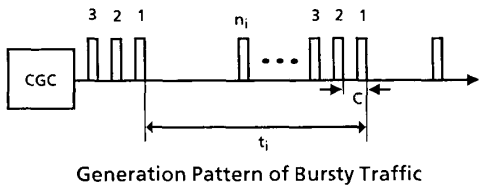


Fig. 6. Cell generation circuit.

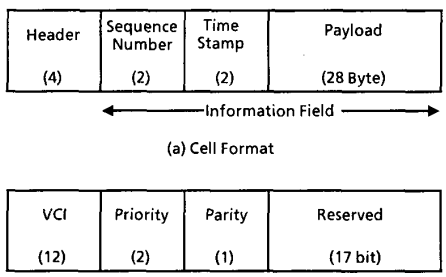


Fig. 7. Experimental system cell format.

The cell loss ratio was more sensitive to the buffer size of the shared buffer memory switch than on the separated buffer memory switch. The logarithm of the cell loss ratio for the shared buffer memory switch decreased more rapidly with buffer size than that for the separated buffer memory switch, as shown in Fig. 8.

On average, 7.1 buffer cells per output port were enough to satisfy a cell loss ratio less than 10^{-7} for the shared buffer memory switch, whereas 35 cells were needed for the separated buffer memory switch. Thus, the buffer size of the shared buffer memory switch was reduced to about one fifth that of the separated buffer memory switch when there were a total of eight input/output ports and the required cell loss ratio was 10^{-7}. This result shows the advantage of sharing the buffer memory in a switch. The cell loss ratio for a shared buffer memory switch can be calculated by mean of a convolution of the queue length distributions of each output port using the $M/D/1$ model [10]. The calculated result is shown by the broken line in Fig. 8.

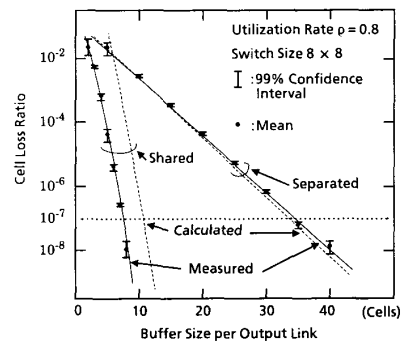


Fig. 8. Cell loss ratio of shared buffer memory switch and separated buffer memory switch (random traffic).

2) Cell Loss Ratio for Bursty Traffic Conditions: In this experiment, cells were generated in a bursty manner. Bursts were randomly generated and the burst length was geometrically distributed. The cells were uniformly distributed to the output ports of the switch.

Fig. 9 shows the relationship between the cell loss ratio and the buffer size of the ATM switch when the line utilization rate $\rho = 0.8$ and the mean burst length was ten cells. The trend of buffer size reduction through buffer sharing under bursty traffic conditions was almost the same as that under random traffic conditions. However, the sample variance under the bursty traffic conditions was greater than that under the random traffic conditions due to the increase in instantaneous bit rate deviation under the bursty traffic conditions.

On the separated buffer memory switch, the relationship between the algorithm of the cell loss ratio and the buffer size was also linear under bursty traffic conditions. The experimental result agreed with the calculated result using the $M^{[x]}/D/1/k$ synchronous server model (broken line in the Fig. 9) [11].

For the shared buffer memory switch, 121.3 buffer cells per output port were enough to satisfy a cell loss ratio less than 10^{-7}. However, 655 cells were needed for the separated buffer memory switch, as shown in Fig. 9. Therefore, under the bursty traffic conditions, the shared buffer memory switch reduced the buffer size to 0.19 that of the separated buffer memory switch when there were a total of eight input/output

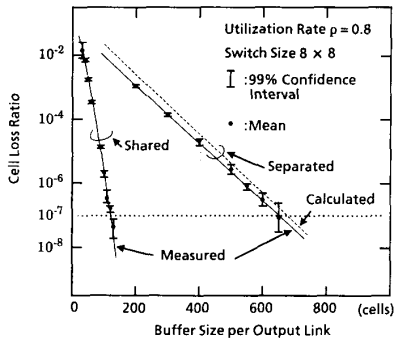


Fig. 9. Cell loss ratio of shared buffer memory switch and separated buffer memory switch (bursty traffic).

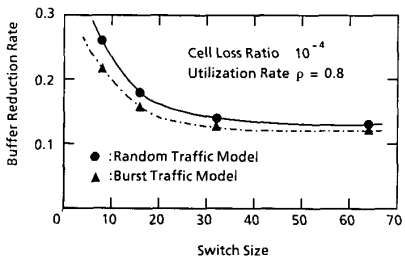


Fig. 10. Buffer reduction through buffer sharing.

ports and the required cell loss ratio was 10^{-7}. This buffer size reduction ratio was almost the same as that under the random traffic conditions.

The buffer size must be further reduced as the switch size becomes larger because the buffer sharing becomes more effective. Fig. 10 shows the buffer reduction ratio through buffer sharing as a function of the switch size estimated by computer simulation. This is the ratio of buffer size of the shared memory switch to that of the separated buffer memory switch required for the same cell loss ratio.

C. Cell Loss Ratio Under Imbalanced Input Traffic

If the input cells are not uniformly distributed to each output port on the shared buffer memory switch, the switch performance becomes worse than that under balanced input traffic conditions. In particular, Kamoun and Kleinrock [12] have shown that control is necessary to prevent the performance degradation of a shared buffer memory switch when one output port of the shared buffer memory switch is overloaded. They use the buffer sharing model of multi $M/M/1$ queues to demonstrate that the performance degrades for all output ports. In our experiment, we considered a more general case when the offered traffic was bursty. The traffic statistics were the same as before, i.e., the burst interarrival time was geometrically distributed and the burst length was ten cells. We illustrate how the same simple control strategy can prevent performance degradation of all other output ports of the shared buffer memory switch when one particular output is overloaded.

Imbalanced traffic conditions presented one problem. When the traffic was heavy, the confidence interval of the measured

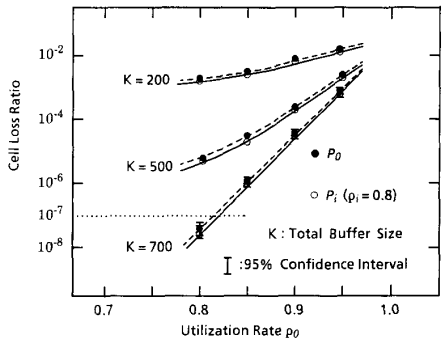


Fig. 11. Cell loss ratio of shared buffer memory switch (imbalanced).

cell loss ratio was large. Therefore, a large number of samples were needed to obtain a reasonable confidence interval. This may be due to coinciding bursts from many cell generation circuits multiplexed together, resulting in large bursts. To minimize the number of samples needed, the cell interarrival time was changed from one unit of time (C in Fig. 6) to eight.

Fig. 11 shows the relationship between the cell loss ratio and the utilization rate ρ_0 of a particular output port of a shared buffer memory switch. P_0 and P_i are the cell loss ratios of the overloaded output port 0 and the nonoverloaded output port i ($i = 1, 2, \cdots, 7$), respectively. The utilization rates of other output ports, defined by ρ_i, are fixed at 0.8. It is evident from the plots of Fig. 11 that as ρ_0 increases, the cell loss ratio of all output ports (not only the overloaded output port 0) becomes worse. In particular, the cell loss ratio degrades more than ten times as ρ_0 changes from 0.8 to 0.85. Furthermore, we confirmed by measurements that P_i also degrades as ρ_0 is increased even for a low value of ρ_i.

In an exchange system, the overload of a particular output port should not affect the performance of other output ports. Therefore, some overload control is necessary in the shared buffer memory switch. Several control schemes have been proposed [12] [13] to prevent performance degradation of the shared buffer memory switch under imbalanced traffic conditions. Of those schemes, we adopted the sharing with a maximum queue length scheme (SMXQ) [12]. In this scheme, a queue length threshold is chosen for the logical queue pertaining to each output port, and if the queue length exceeds the threshold, the arriving cells are discarded. Of course, the sum of these maxima must be greater than the total buffer size in order to take advantage of buffer sharing.

Let K be the total buffer size of the shared buffer memory, and $M = \alpha K$ be the threshold of the output queue length. When $\alpha \geq 1$, the switch works as a completely shared buffer memory switch (CS), and when $\alpha \leq 1/N$ (where N is the number of output ports), the switch works as a completely separated buffer memory switch.

Fig. 12 plots P_0, P_i as a function of α. The input traffic is the same as that used for Fig. 11, and ρ_0 is 0.95. As the value of α changes from 1 to 0.5, P_i decreases because the cells destined to the overloaded output port are discarded. However, as the value of α approaches $1/N$, the effect of buffer sharing decreases, and the cell loss ratio P_i increases again. Therefore,

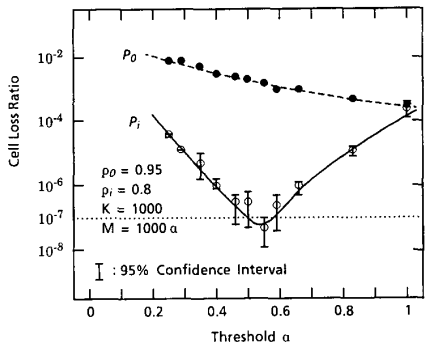


Fig. 12. Cell loss ratio of CS and SMXQ(1).

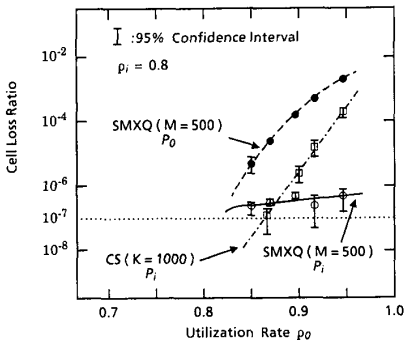


Fig. 13. Cell loss ratio of CS and SMXQ (2).

there is an optimal value of α, which is close to 0.5. The cell loss ratio P_i with SMXQ control strategy is 100 times lower than that with CS at a threshold value $\alpha = 0.5$.

Fig. 13 compares the performance of the SMXQ and CS policies as a function of ρ_0. Because P_0 is almost the same as P_i for the CS policy, we show only P_i. The cell loss ratio P_i at $\rho_0 = 0.85$ is 10 times higher for SMXQ than for CS. However, as ρ_0 increases, the cell loss ratio P_i for SMXQ is almost constant, while that for CS increases rapidly. Consequently, under imbalanced traffic conditions the performance of SMXQ is better than that of CS. Figs. 14 and 15 compare the performance of SMXQ and CS as a function of the buffer size K under balanced and imbalanced loads, respectively. Here, too, only P_i is considered. For balanced traffic conditions, the total buffer size required to satisfy a cell loss ratio of the 10^{-7} with the SMXQ policy is about 1000 cells, which is roughly 1.5 times that needed (700 cells) with the CS policy, as shown in Fig. 14. However, for imbalanced traffic conditions the total buffer size required to satisfy the same cell loss ratio with SMXQ policy is about 1150 cells, as shown in Fig. 15, and is roughly the same as that under the balanced traffic conditions. On the other hand, as seen from the CS cell loss gradient in Fig. 15, it is obvious that CS requires a larger buffer capacity for the same performance on the imbalanced load. These results show that SMXQ control effectively prevents shared buffer memory switch performance degradation under imbalanced traffic conditions.

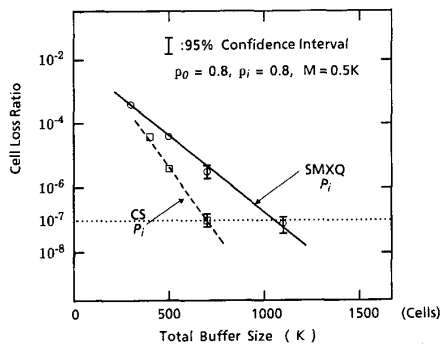


Fig. 14. Required buffer size on CS and SMXQ (balanced).

V. LSI Implementation

The shared buffer memory switch is partitioned into two kinds of custom-made LSI's and one kind of commercial LSI, as depicted in Fig. 16. The BFM LSI and the CTL LSI are custom-made, and the Idle Addr Mem is a commercially available LSI. They are all mounted on a printed board. The cell buffer memory, MUX, and DMUX in Fig. 1 are contained on the BFM LSI, and the Route DEC, WAR, RAR, OUT DEC, and OUT CNT are contained on the CTL LSI.

The requirements for this LSI implementation study are as follows:

1) Switch size: 32 input/32 output ports (150 Mbps/port)
2) Buffer quantity: 4 k cells (128 cells/port)
3) Service class: 3 classes
4) Cell size: 53 bytes
5) LSI Technology: submicron CMOS process

The buffer quantity is calculated as follows.

The total buffer size required at a cell loss ratio of 10^{-9} for an 8×8 switch with SMXQ under balanced traffic conditions is about 1500 cells, as shown in Fig. 14. This gives about 188 cells buffer/output port for an 8×8 switch. The buffer size for a 32×32 switch can be estimated as 107 cells (rounded to 128 cells/output port) by using the buffer reduction rate shown in Fig. 10. Only the restriction of the maximum queue length (SMXQ) for imbalanced traffic conditions is considered at present.

The BFM LSI is arranged in the form of a bit-slice, i.e., each LSI corresponds to each bit of an 8 bit parallel format byte of an ATM cell. Cell data input or output from each BFM LSI are further converted into or from the 53 bit parallel form, which has a one cell bit-length, and these parallel formed data are written into or read from the buffer memory.

Therefore, the cycle time of the buffer memory can be roughly calculated as follows:

$$1/150(\text{Mbps}) \times 8(\text{bit-slice}) \times 53(\text{bit-cell-length})/ 32(\text{port-multiplexed})/2(\text{write and read}) = 44 \text{ ns}.$$

Considering various margin factors, the memory cycle time is determined as 25 ns . The capacity of the buffer memory is required to be more than 212 kbits = 4(kcells) $\times$ 53(bytes) $\times$8(bits)/8(bit-slice) for each BFM LSI. Because numerous

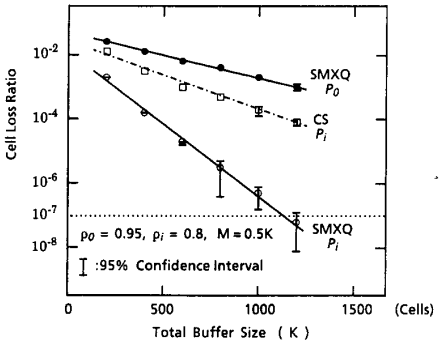


Fig. 15. Required buffer size on CS and SMXQ (imbalanced).

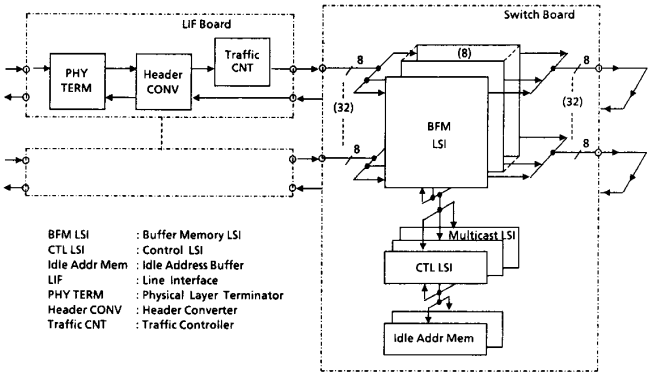


Fig. 16. LSI partitioning and board mounting.

interconnections are required between the buffer memory array and the peripheral circuits, such as the address decoder or read/write register, the buffer memory has to be implemented in an on-chip memory. The cycle time and capacity mentioned above are nearly the upper limits of present on-chip memory technology.

The CTL LSI provides functions such as buffer memory read or write control and service class control. These functional specifications can be changed in detail by modifying only the CTL LSI without redesigning the BFM LSI. One CTL LSI can control up to two service classes, one of which is prior to the other. For three service classes, two CTL LSI's are required and the classes on one LSI are prior to the classes on the other LSI. The LSI specifications for a 32 × 32 switch are summarized in Table III.

The multicast function requires one more CTL LSI, as shown in Fig. 16.

To produce a 600 Mbps-ATM switch, the input cells are distributed to four input ports and the cells from the four output ports are multiplexed on one output line. To preserve the cell sequence, the cells toward the four output ports are chained in one output queue and the cell sequence of the input-distribution is made to coincide with that of the output-multiplexing. The 600 Mbps-switch can be achieved by modifying the CTL LSI and by adding distributors and multiplexers to the switch.

TABLE III
LSI SPECIFICATIONS FOR A 32 × 32 SWITCH

LSI		Gate		Memory		Signal Pins	Chip Count	LSI Process
		Count	Delay	Count	Cycle Time			
BFM	Custom Made	46 KG	<0.5 ns	258 Kbit	<25 ns	186	8	Submicron CMOS
CTL	Gate Array	50 KG		—	—	319	2	Submicron CMOS
Idle Addr Mem	Commercially Available FIFO	—	—	4KW × 9bits FIFO	<35 ns	26	2	—
					Total Chip Count		12	

A 32×32 ATM switch can be mounted on one printed board (300 mm × 300 mm), as shown in Fig. 16.

VI. CONCLUSION

We proposed a shared buffer memory switch, as the ATM switch with the least amount of hardware composed of various switch architectures. This architecture decreases the buffer memory capacity by sharing the cell buffer memory among all the switch output ports.

A shared buffer memory switch also has the flexibility to meet the detailed specification variations of functions such as

priority control, queue length restriction, multicast connection and bandwidth expansion of the switch input or output ports from 150 Mbps to 600 Mbps, by redesigning the buffer memory read or write control circuit.

Buffer sharing was evaluated by a traffic analysis and experimental measurements. These evaluations showed that the required buffer capacity for a shared buffer memory switch is less than $1/5$ that required for a separated buffer memory switch, and needs about 128 cells/output-port for a 32×32 switch when the input-traffic model has bursty conditions and when the destination addresses are uniformly distributed. The hardware reduction allows a 32×32 switch (150 Mbps/port) to implement the switch in about 12 CMOS LSI chips and to be mounted on one printed board.

ACKNOWLEDGMENT

The authors wish to express their gratitude to Dr. H. Shirasu of Kohgakuin University, Mr. E. Amada of the Central Research Laboratory, Hitachi Ltd., and to Mr. K. Ohtsuki and Mr. T. Kato of Telecommunications division, Hitachi Ltd. for their valuable discussion and guidance regarding ATM switch architecture. We would also like to thank Mr. K. Asano and Mr. M. Ogino of the Central Research Laboratory, Hitachi Ltd. for their logic design and estimation of hardware quantity.

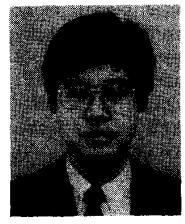
REFERENCES

[1] "Recommendation I.121," CCITT Document AP IX-143–E, July 1988.
[2] Y. Sakurai, N. Ido, S. Gohara, and N. Endo, "Large scale ATM multi-stage switching network with shared buffer memory switches," in *Proc. ISS'90*, vol. 4, 1990, pp. 121–126.
[3] J. Hui, "A broadband packet switch for multi-rate service," in *Proc. ICC '87*, 1987, pp. 782–788.
[4] A. Thomas, J. P. Coudreuse, and M. Servel, "Asynchronous time-division techniques: An experimental packet network integrating video-communication," in *Proc. ISS '84, 32.C.2*, 1984.
[5] K. Y. Eng, M. G. Hluchyj, and Y. S. Yeh, "A knockout switch for variable length packets," in *Proc. ICC '87*, 1987, pp. 794–799.
[6] H. Kuwahara, N. Endo, M. Ogino, and T. Kozaki, "A shared buffer memory switch for an ATM exchange," in *Proc. ICC '89*, 1989, pp. 118–122.
[7] N. Endo, T. Ohuchi, T. Kozaki, H. Kuwahara, and M. Mori, "Traffic characteristics evaluation of a shared buffer ATM switch," in *Proc. GLOBECOM'90*, vol. 905.1, 1990.
[8] T. E. Eliazov, V. Ramaswami, W. Willinger, and G. Latouche, "Performance of an ATM switch: Simulation study," in *Proc. INFOCOM.*, June 1990, pp. 644–659.
[9] L. Kleinrock, *Queueing Systems, Vol. 1: Theory*. New York: Wiley.
[10] A. E. Eckberg and T.-C. Hou, "Effects of output buffer sharing on buffer requirement in an ATDM packet switch," in *Proc. INFOCOM.*, Mar. 1988, pp. 459–466.
[11] W. W. Chu, "Buffer behavior for batch Poisson arrivals and single constant outputs," *IEEE Trans. Commun.*, vol. COM-18, pp. 613–618, Oct. 1970.
[12] F. Kamoun and L. Kleinrock, "Analysis of shared finite storage in a computer node environment under general traffic conditions," *IEEE Trans. Commun.*, vol. COM-28, pp. 992–1003, July 1980.
[13] M. I. Irland, "Buffer management in a packet switch," *IEEE Trans. Commun.*, vol. COM-26, pp. 328–337, Mar. 1978.

Noboru Endo (S'81–M'86) received the B.E., M.E., and Ph.D. degrees from Tohoku University, Sendai, Japan, in 1981, 1983, and 1986, respectively.

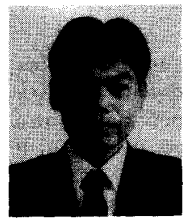
In 1986, he joined Central Research Laboratory, Hitachi, Ltd. Since then, he has engaged in the research of fast packet switching systems and B-ISDN switching systems. From 1990–1991, he was with Bellcore, Red Bank, NJ.

Dr. Endo is a member of the Institute of Electrics, Information, and Communication Engineers (IEICE) of Japan, and the Information Processing Society of Japan.



Takahiko Kozaki (M'88) received the B.E. and M.E. degrees from the University of Tokyo, Japan, in 1981 and 1983, respectively. In 1983 he joined the Central Research Laboratory, Hitachi, Ltd., Tokyo, Japan. He worked for the design of digital codec LSI for telecommunications from 1983 to 1985. Since 1986, he has been engaged in the research of fast packet-switching systems and BISDN switching systems.

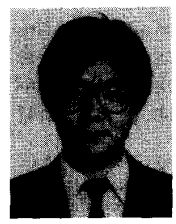
Mr. Kozaki is a member of the Institute of Electrics, Information, and Communication Engineers (IEICE) of Japan.



Toshiya Ohuchi received the B.E. and M.E. degrees from Tohuku University, Sendai, Japan, in 1985 and 1988, respectively. Since 1988, he has been with Central Research Laboratory, Hitachi, Ltd., engaged in the research of fast packet-switching systems and B-ISDN switching systems.

Hiroshi Kuwahara received the B.S. degree in physics from Tohoku University, in 1963 and joined Hitachi, Ltd.

He had been engaged in developments of an electronic switching system, a voice Codec and an ATM switching system at Totsuka Works and at Central Research Laboratory of Hitachi, Ltd. He is currently developing a digital cordless telephone system for personal communications at Kokusai Electric Co., Ltd.



Shinobu Gohara received the B.S. degree from the Tokyo Institute of Technology in 1977. He joined Hitachi, Ltd., in 1977.

He was concerned with the development of the ISDN system. He is currently engaged in the R&D of ATM switching systems.