Joint Conference of 1996 International Computer Symposium
December 19~21, Kaohsiung, Taiwan, R.O.C.

USING TRIANGULAR STYLE MEMORY ARCHITECTURE FOR HIGH SPEED SWITCH

Liang-Teh Lee, Po-Hsian Huang, and Der-Fu Tao

Department of Computer Science and Engineering
Tatung Institute of Technology
Taipei 10451, Taiwan
Email: ltlee@cse.ttit.edu.tw

Abstract

This paper presents a new architecture of a high-speed switch which can be used for ATM. The switch, called TSM (Triangular Style Memory), uses Dual-Port memories to construct a triangular style memory for eliminating the bus contention. With the same required throughput and cell loss probability, the performance analysis of the switch shows that the number of buffers and average cell delay can be significantly reduced in the proposed switch while comparing to the ATM switches with central shared memory.

1. Introduction

The ITU (International Telecommunications Union) has standardized the ATM (Asynchronous Transfer Mode) as the multiplexing and switching principle for the B-ISDN (Broadband Integrated Services Digital Network). ATM is a packet and connection-oriented transfer mode based on statistical time division multiplexing techniques [11,12].

In the STDM (Synchronous Time Division Multiplexing) or STM (Synchronous Transfer Mode), the network bandwidth is divided into fixed time slots [11]. Every time slot presents the bandwidth usage of a channel. From the Fig. 1-1a, the bandwidth of channel 1 is still occupied and idle if there is no data to be transferred. If there is many data to be transferred but the bandwidth is not enough in channel 2, these data must be transmitted in next dedicated time slot. In general, the idle phenomenon will waste the network bandwidth, thus the utilization will perform poorly.

ATM adopts another complete different processing style [11]. As showing in Fig. 1-1b, it adds the header which contains the virtual path and virtual channel. The network bandwidth can be used effectively since the bandwidth are shared to any user. If there is no cell to transmit, an idle or an empty cell is sent. For non-consecutive data, ATM is more flexible and effective than STM in bandwidth utilization.

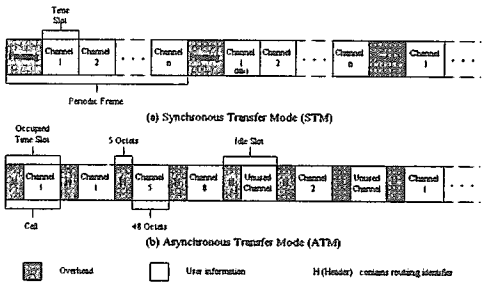

Fig. 1-1. STM and ATM.

ATM is constructed by a fixed length of 53 octets, which is called cell. The cell is divided into 5 octets header and 48 octets payload [11]. The transfer rate of the fixed length cell in ATM network is fast, thus the transmission is fixed and short. This property will make the ATM more suitable to transfer voice and video. In Fig. 1-2, data, voice, and video traffic sources will be segmented into 48 octets stringswhen they go through the ATM interface card. After the 5 octets headers are added in every 48 octets, cells are transmitted into the ATM switching technique in ATM network to transmit cells in high speed, it is necessary to include the translation of VPI/VCI table. The VPI/VCI will be changed into the next VPI/VCI when they go through the ATM switch. After passing some ATM switches, cells will arrive at the destination. By removing the header, the original data, voice, or video will be reassembled to complete the cell transmit. Besides of integrating the data, voice, and video, the ATM can transmit synchronously and also support many kinds of B-ISDN services.

2. Various ATM Switch Architectures

Since ATM will provide transport and switching environment for emerging B-ISDN network, a great deal of research is involved, especially in ATM switch

hardware design which has been understood with considerable level. Despite the maturity of this characterization effect, new switch architectures are still proposed continuously. There are several kinds of ATM switch architectures were proposed recently [2-4,7-10,13-17]. In general, both data section and control section of an ATM switch architecture can be classified into three classes. In the data section, Space division [17], Shared medium [4], and Shared memory [2-3,8,10,14,15-16] are classified. Meanwhile, there are three different kinds in the control section -- Linked-list [3,8,10], Hybrid [8,13], and CAM (Content Addressable Memory) based [15]. In this chapter, various ATM switch architectures are described in detail. In next chapter, the TSM switch is presented thoroughly and compares with the switch architectures that are described in this section.

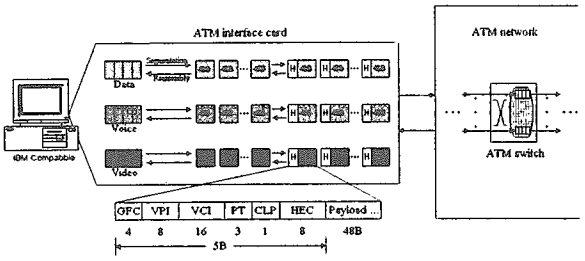

Fig. 1-2. ATM cell and transmission.

2.1 Data Section

The Space division ATM switch is shown in Fig. 2-1. Every input port and output port of the switch are connected by a dedicated communication link. By using the dedicated communication link, cells can be transmitted from input port to output port immediately. If the switch is nonblocking, every output control unit which contains a small dedicated FIFO buffer is employed to avoid switch contention. On the contrary, if the switch is blocking, every input and output control units must contain a small dedicated FIFO buffer to avoid switch contention. Therefore, a large switching element is necessary to exchange cells between input and output ports for this kind of switch architecture. Obviously, the volume and cost of switch is large and high respectively. If the VLSI technique is applied, the technique is difficult to be manufactured in a single-chip.

Figure 2-2 represents a shared medium ATM switch. Both Input controller and Output controller contain a small dedicated FIFO buffer to avoid switch contention. The multiplexer /demultiplexer in input/output port makes cells can keep the same speed in input and output ports within a certain time slot. The speed of the internal operation in switch must be n times faster than that of input and output ports. Since the switch uses the bus to exchange cells, 425-line bus contains a flag line to identify whether there is a cell to be transmitted. If we adopt the VLSI technique to form a single-chip, some defects cannot be avoided, such as the space for the bus is large, the consumption of power is wasted, and the signal noise inflection between lines is huge.

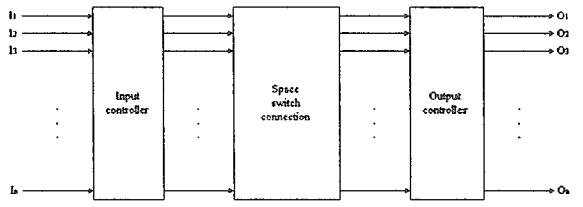

Fig. 2-1. Space division ATM switch.

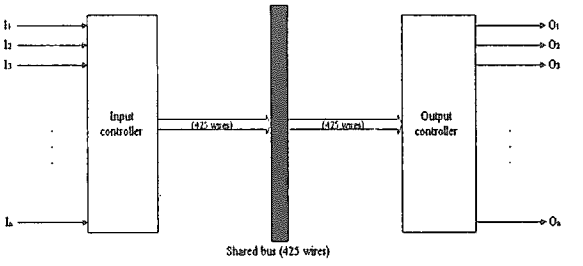

Fig. 2-2. Shared medium ATM switch.

In Fig. 2-3, a new generation of ATM switch architecture called shared memory ATM switch isshown. The input and output ports share a piece of common memory. The input cell and output cell could arrive and transmit at the same time respectively. Using a small shared memory instead of the dedicated FIFO buffer in input and output ports can obtain high ratio of performance/cost [2]. The relative small shared memory is easier implemented in a single-chip ATM switch for VLSI technique.

In every time slot the multiplexer and demultiplexer have to service the input cells and output cells for input and output ports respectively. Thus, the speed of cell transmitted is limited by the number of input and output ports as well as the internal operation speed of the switch. If the internal operation speed is fixed in switch, the more input and output ports, the lower cells transmission speed. Furthermore, if the traffic load is heavy in one of the input ports, the total shared memory will be occupied by this input port. This phenomenon makes other input ports cannot transmit cells and results in the occurrence of the cell loss. Recently, the designation of this kind of switch restricts an input port could occupy at most fifty percent of the shared memory. If the ratio exceeds fifty percent, the excess cells will be lost [1].

In our proposed switch, TSM, the architecture of multiprocessor and multi-bus can overcome the shortcoming of the shared memory switch. Moreover, the TSM uses a small amount of Dual-Port memory to replace the traditional large multi-port shared memory. The switch is relatively easier to be implemented in a single-

chip

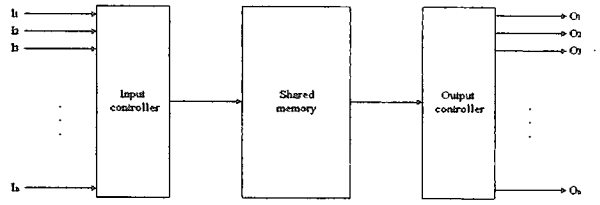

Fig. 2-3. Shared memory ATM switch.

2.2 Control Section

Typically, shared memory ATM switches use Linked-list to control cells. As shown in Fig. 2-4, the cells which are destined to the same output port will be linked together. The sequence of cells in the constructed link corresponds to the order of cells arriving. Finally, these cells will be stored in the shared memory. A shared memory address A[o][i] stores the cell C[o][i] and links to the location which stores the next cell with the same destination as cell C[o][i]. Every output port contains a pair of read and write address registers. The read address register (RA) points to location which stores the cell to be read next , and the write address register (WA) points to a free location for storing next incoming cell.

For writing :
1. Read the shared memory address A[i][j+1] from WA[i].
2. Read a free location A[i][j+2] from the head of the Free Memory address Buffer, FMAB[h].
3. The arrived cell and A[i][j+2], Next Cell Address (NCA), will be written into the memory address A[i][j+1]].
4. In order to write next incoming cell, the A[i][j+2] will be written into WA[i].

For reading :
1.Read the shared memory address A[i][j] from RA[i].
2.The cell and A[i][j+1], NCA, will be read from the shared memory address A[i][j]
3.The shared memory address of A[i][j] will be released and appended to the FMAB[t], where t is the tail of the FMAB.
4.In order to read next cell, the shared memory address A[i][j+1] will be written into RA[i].

From the above algorithm, the whole processing has close relationship with memory address. Unfortunately, the main purpose of the switch is to transfer cell from input port to output port with short delay. The Linked-list is easy to understand and can be implemented widely, but it will waste enormous time in address operation and waste memory to store address. Thus, the entire performance of switch will decrease.

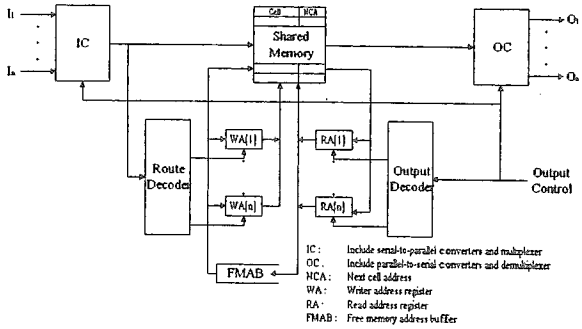

Fig. 2-4. Linked-list based shared memory ATM switch.

Figure 2-5 shows a Hybrid shared and dedicated output buffer ATM switch architecture. Every output port has a dedicated output FIFO buffer to store cell address in the shared memory.

For writing :
1.Read a free shared memory address A[i][j] from FMAB[h].
2.Write cell into the shared memory address A[i][j], and store A[i][j] into $FIFO_i$ [t].

For reading :
1.Read A[i][j] from $FIFO_i$ [h].
2.Read the cell from the shared memory address A[i][j].
3.Release the shared memory address A[i][j], and append it to the FMAB[t].

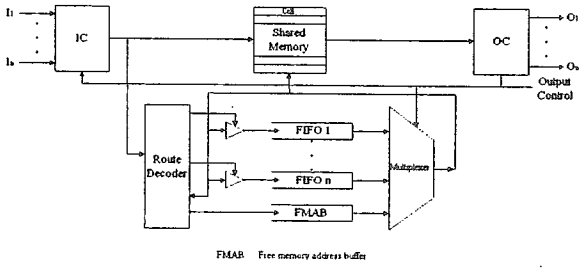

Fig. 2-5. Hybrid shared and dedicated output buffer ATM switch.

The Hybrid method uses hardware to control the access of the shared memory. Therefore, the cell transfer rate is quite high. Using shared memory to store cell in data section can use less memory. The memory utilization will decrease a lot because the method adopts a great number of FIFO buffer to store cell address in control section.

A CAM based shared memory ATM switch, CAM/RAM, is shown in Fig. 2-6. The shared memory is used to store cells and CAM to store tags for referencing cell. A cell can be uniquely identified by the tag, which is composed of the output port number and the sequence number (SN). By searching the desired tag, cells can be accessed efficiently.

For writing :
1.Read the write sequence number, WS[i], from write SN RAM and use [i,WS[i]] (where *i* is the number of output port) as the tag of the cell.
2.Search the first free location (ffl) of the Tag CAM.
3.The C[i][j] and the [i,WS[i]] will be written into the memory address [ffl] and corresponding Tag CAM respectively.
4.Increase the write sequence number, WS[i], by one.
For reading :
1.Read the read sequence number, RS[i], from read SN RAM.
2.Search the corresponding tag in Tag CAM by [i,RS[i]].
3.Read cells from memory address [i,RS[i]].
4.Increase the read sequence number, RS[i], by one.

Using CAM based method to access cell inside the shared memory is independent of cell address. Therefore, the address decoding is not necessary, but it is requires a lot of CAM to store and tag.

To summary, the common phenomenon of the above three control mechanisms is that the huge memory and complicated control circuit needed. By using the shared memory ATM switch, the memory utilization will be degraded significantly. For eliminating the problems of the above three mechanisms, an ATM switch, TSM, is proposed.

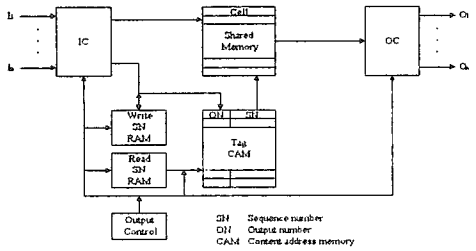

Fig. 2-6. CAM based shared memory ATM switch.

3. The TSM Switch Element

As mentioned in previous section, shared memory ATM switches have the following demerits :
1.Since input and output ports contain the same transfer rate, the performance of dredging cell contention is not well.
2.A central control unit is necessary to control the hole switch, so one powerful CPU is required and the cost is expensive.
3.If the traffic is heavy in the input port, the arriving cells may occupy the whole shared memory. Thus, the cell lose occurs.
4.The control section needs a great deal of memory to store cell address; thus, the memory utilization ratio of storing cells will be low.

In this section, we propose a switch, called TSM, which applies distributed processing, multiprocessor, and multi-bus techniques to enhance the speed of the switch [6]. The structure of a 4x4 TSM ATM switch element is shown in Fig. 3-1. It is organized by five parts :
1.PE_i (Processing Element *i*; where *i*=1, 2, 3, 4) : Each PE contains a processor unit, a DMAC (Direct Memory Access Controller), and a routing table. The processor unit is used to control the cell from the input port to the dedicated Dual-Port memory location. The DMAC controls the cell transfer among the memory modules. The routing table translates the cell header of VPI/VCI.
2.BUS_i (BUS *i*; where *i*=1, 2, 3, 4) : The multi-bus is employed to accelerate the speed of the cell input, transport, and output.
3.M_{ii} (Memory module *ii*; where *i*=1, 2, 3, 4) : Each memory module contains a Dual-Port memory and a CASU (Check and Send Unit). The Dual-Port memory is used to store valid bit, output port number, and cell. The function of CASU is to transfer cells from Dual-Port memory to O_n (Output port *n*).
4.M_{ij} (Temporary memory *ij*; where *i, j*=1, 2, 3, 4) : M_{ij} is a temporary memory which contains a 8 bits Dual-Port memory or Dual-Port register that is used to transfer cells between M_{ii} and M_{jj}.
5.RA_i/WA_i (Read and Write Address register *i*; where *i*=1, 2, 3, 4) : Every memory module has a pair of RA/WA. RA saves the next cell address which will be sent to the output port, and WA saves the empty memory location to store the next input cell.

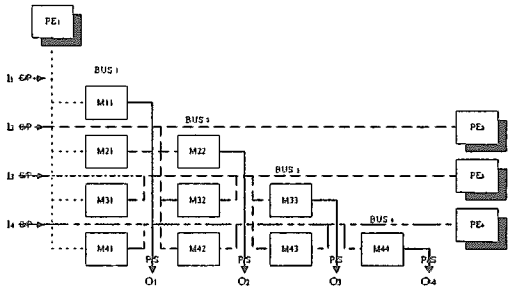

Fig. 3-1. The TSM ATM switch architecture.

There are many S/P (serial-to-parallel converter) and P/S (parallel-to-serial converter) products which can convert efficiently. But, the TSM need not S/P and P/S units inside the TSM switch for the sake of extending the switch size easily. This section will be divided into two subsections, Input part and Output part. We will make a detail description for the proposed switch.

3.1 Cell Storage Mechanism

When the cell arrives at the input port of the TSM switch, it goes through the S/P converter for internal processing. After the PE_i translates the VPI/VCI of the cell header, the cell and the 2-bits DN, the outgoing port number that the cell desired, will be written to the desired

Dual-Port memory. Before writing information to the Dual-Port memory, it is necessary to verify the available capacity of the Dual-Port memory. If WA_i is not equal to RA_i-1, which means the dedicated Dual-Port memory is not full, the cell is stored to the dedicated Dual-Port memory directly. Otherwise, the DMA method is employed to transfer the cell to another Dual-Port memory that is used less than fifty percent, or the cell is discarded, which is the same as the other switch mechanisms. Once the cell is stored, the valid bit of the selected memory location will be set to "1".

Inside the TSM switch, every output port has a Dual-Port memory to store cells. When the dedicated Dual-Port memory is full, the new incoming cells must be transferred to other Dual-Port memory whose remaining space is more than fifty percent. From the description discussed above, TSM has a merit of high efficiency transfer rate since it can be thought as a dedicated input/output buffer ATM switch. However, if the traffic load is heavy in some input port, and the dedicated Dual-Port memory is not enough, other Dual-Port memory whose traffic load is light can be shared. TSM has the merit of good memory utilization.

3.2 Cell Output Mechanism

Every output port contains one Dual-Port memory and one CASU as shown in Fig. 3-2. The function of the CASU will check the valid bit. A cell exists in memory if the valid bit is set to "1". The DN will be compared with O_n, which the CASU output port number. If DN is equal to O_n, the cell will be sent to O_n through P/S. On the contrary, the cell will be stored back to its desired outgoing Dual-Port memory. Using the DMAC, the cell will be sent to the original dedicated Dual-Port memory. If the valid bit is set to "0", there is no cell in Dual-Port memory and the CASU will be idle. The above operation can proceed simultaneously on n output ports.

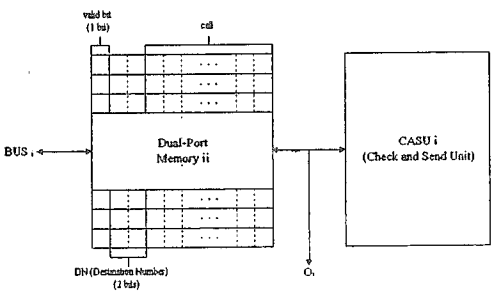

Fig. 3-2. The memory module.

The advantages of the TSM can be summarized as follows :

1. Every input/output port controlled by one processor will decrease the control complexity, and increase the flexibility. Moreover, low-end processor can be used to reduce the cost of the switch.
2. The inter-processing speed is n times faster than the single input and output ports. Due to adopt the multi-bus, at the same time one cell arrive the input port and n cells output. The switch contention will effectively be eliminated, and thus the band-width gain of whole switch can reach n x n.
3. Since the output port speed is n times of the input port speed, the amount of total memory required will greatly be reduced.

4. Performance Analysis

Some analyses of the TSM switch performance which containing the total memory required, MTOB (Maximum Total Output Bandwidth), cell loss probability, and cell delay are detail discussed in this section.

4.1 The Total Memory Required and Maximum Total Output Bandwidth of TSM

Comparison of the total memory required and bandwidth gain for four 4x4 ATM switches, Linked-list, Hybrid, CAM, and TSM, with 256 cell buffer respectively is summarized in Table 1. It shows that the total memory required for the TSM is 109,408 bits which is the best one among the four switches. The bandwidth gain of the TSM is 16 times greater than other switches.

Table 1: Comparison of Linked-list, Hybrid, CAM, and TSM.

	Link list	Hybrid	CAM	TSM
cell storage (decode)	RAM (decode) 256x424=108,544	RAM (decode) 256x424=108,544	CAM/RAM (none) 256x424=108,544	Dual-Port RAM (none) 64x424=108,544
look-up	link : RAM 256x8=2,048	FIFO 128x8x4=4096	tag : CAM 256x(2+7)=2,304	DN : RAM 64x8x2=512
write and read reference (queue length checking)	R/W address registers (additional counters) 2x4x8=64	none (additional counters) none	sequence number registers (compare W and R numbers) 2x8x7=56	R/W address registers (compare R/W registers) (6x8)+(2x8x6)=96
free memory address storage (additional overhead)	FMAB (pointer maintenance, extra memory block) 256x8=2,048	FMAB (extra memory block) 256x8=2,048	CAM valid bit (none) 256x1=256	RAM valid bit (none) 256x1=256
total memory required (unit : bit)	112,704	114,688	111,160	109,408
bandwidth gain (output/input)	1	1	1	16

(sample switch size is 4x4 with 256 cell buffer capacity)

In estimating the throughput, the processing time of the PE and the CASU needs 5ns individually, and the access time of the Dual-Port memory is 20ns. Hereby, one cell arriving the TSM through the PE, the Dual-Port memory, and the CASU to output port is 30ns. Since a cell contains 53 bytes (424 bits), the MIB (Maximum Input Bandwidth) of the switch can be calculated as :

$$MIB = 424 / (n \times T_s)$$

where n is the number of inputs, and T_s is the switching time (30ns).

For an n x n TSM switch, since the switching operation time for processing n inputs is n x T_s, the maximum bandwidth of an output port is n x MIB. Therefore, MTOB (Maximum Total Output Bandwidth) = MIB x n^2.

Considering a 4x4 TSM switch :

MIB = 424 / (n x T_s) = 424 / (4x30ns) = 3.5 Gb/s

MTOB = MIB x n^2 =3.5 Gb/s x 4^2 =56 Gb/s

Table 2 presents the MIB and the MTOB of 4x4, 8x8, and 16x16 TSM switches. The MIB is in inverse proportion to the number of input ports is the same as other ATM switches. Yet, the MTOB is in proportion to n^2. Consequently, the more the number of links of each switch, the higher total MTOB will be obtained. The proposed TSM switch not only can support higher cell transfer rate but also can effectively avoid cell contention

Table 2: The MIB and the MTOB of 4x4, 8x8, and 16x16 TSM switch.

Number of input/output port	MIB	MTOB
4	3.5 Gb/s	56 Gb/s
8	1.77 Gb/s	113 Gb/s
16	0.88 Gb/s	225 Gb/s

4.2 Simulation Model and results

Figures 4-1 and 4-2 show the simulation model of TSM and Linked-list ATM switch respectively. The simulation program, smpl (simulation program language), which is adopted to simulate these switching mechanisms. The Poisson distribution is employed to generate traffics. The processing time of PE, Temporary memory, CASU, and Dual-Port memory are 5ns, 5ns, 5ns, and 20ns respectively. The time unit for transferring a cell is 2.74us with 155 Mb/s link transfer rate.

Figure 4-3 presents the relationship between the cell loss probability and the buffer size of the TSM and Linked-list switch architectures. The diagram shows that the cell loss probability of TSM always smaller than that of Linked-list in various buffer size.

Figure 4-4 shows mean delay on the ordinate axis and offered load on the abscissa. If the offered load in a certain value is fixed, it is clear to observe that the value of mean delay in TSM curve is smaller than that in Linked-list.

5. Conclusions

A new architecture of high speed ATM switch called TSM has been proposed in this paper. The distributed processing, multiprocessors and multi-bus, and the triangular style Dual-Port memory are employed inside the TSM to accelerate the switching speed. The circuit to implement the TSM switch is simple and needs small amount of Dual-Port memory units. If a low-end processor such as 33 MHZ clock rate microprocessor is used, the total cost of TSM will be reduced significantly.

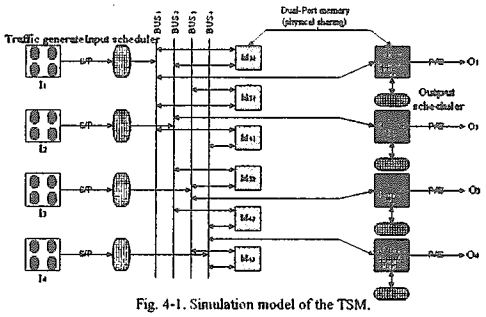

Fig. 4-1. Simulation model of the TSM.

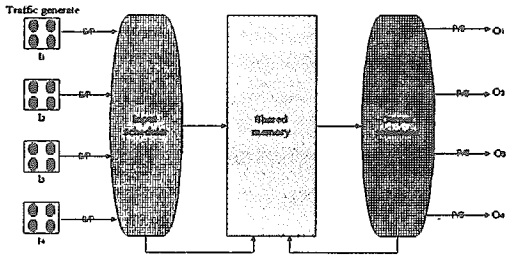

Fig. 4-2. Simulation model of the Linked-list ATM switch.

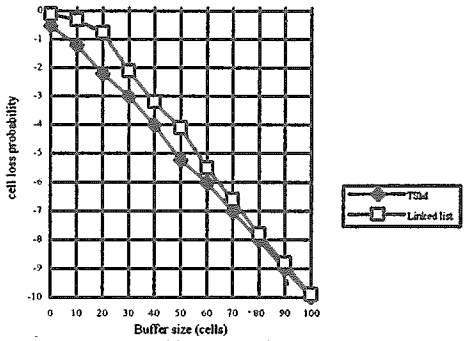

Fig. 4-3. The cell loss probability vs. the different buffer size for a single-stage switch.

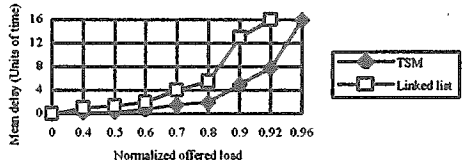

Fig. 4-4. The delay vs. the different offered load for a single-stage switch.

According to the performance analysis, the TSM uses less Dual-port memory units and obtains less cell loss probability .

In designing specific circuit , the switch bandwidth gain will reach n^2, where n presents the number of input

and output ports. Considering a 4x4 TSM switch, the maximum throughput of the switch can reach 56 Gb/s. The TSM not only can provide the cell transfer rate but also can effectively avoid cell contention.

References

[1] Causey, Jeffrey W. and Kim, Hyong S., "Comparison of Buffer Allocation Schemes in ATM Switches: Complete Sharing, Partial Sharing, and Dedicated Allocation," *IEEE ICC*, pp. 1164-1168, 1994.

[2] Denzel, W. E., Engbersen, A. P. J. and Iliadis, I., "A flexible shared-buffer switch for ATM at Gb/s rates," *Computer Networks and ISDN Systems*, pp. 611-624, 27(1995).

[3] Endo, N. *et al.*, "Shared Buffer Memory Switch for an ATM Exchange," *IEEE Trans. Commun.*, vol. 41, pp. 237-245, Jan. 1993.

[4] Fayet, C., Jacques, A. and Pujolle, G., "High speed switching for ATM: the BSS," *Computer Networks and ISDN Systems*, pp. 1225-1234, 26(1994).

[5] Handel, R. and Huber, M. N., Integrated Broadband Networks, *Addison-Wesley Publishing Company*, June 1991.

[6] Hwang, Kai, Advanced Computer Architecture, *McGraw-Hill, Inc.*, 1993.

[7] Jiang, Xi and Meditch, James S., "A high-speed integrated services ATM/STM switch," *Computer Networks and ISDN Systems*, pp. 459-477, 26(1993).

[8] Joan, G. H. and Andrzej, J., "ATM shared-memory switching architecture," *IEEE Network*, pp. 18-26, August 1994.

[9] Kim, Hyong S., "Design of a fault-tolerant multichannel ATM switch for BISDN," *Computer Networks and ISDN Systems*, pp. 29-43, 27(1994).

[10] Kuwahara, H. *et al.*, "A shared Buffer Memory Switch for an ATM Exchange," *IEEE ICC*, pp. 118-122, June 1989.

[11] McDysan, D. E. and Spohn, D. L., ATM theory and application, *McGraw-Hill, Inc.*, 1995.

[12] Prycker, D. P., Bell, A., Antwerp and Belgium, "Asynchronous Transfer Mode," *Ellis Horwood Limited*, pp. 131-133, 1991.

[13] Ronald, P. Bianchini and Hyong, S. Kim, "The Tera Project: A hybrid queueing ATM switch architecture for LAN," *IEEE J. Select. Areas Commun.*, vol. 13, no. 4, pp. 673-684, May 1995.

[14] Sanjeev, K. and Dharma, P. A., "A share-buffer direct-access(SBDA) switch architecture for ATM-based networks," *IEEE ICC*, pp. 101-105, 1994.

[15] Schultz, K. J. and Gulak, P. G., "CAM-Based Single-Chip Shared Buffer ATM Switch," *IEEE ICC*, pp. 1190-1195, 1994.

[16] Takahiko Kozaki, Morihito Miyagi and Ichiro Kohashi, "PVC Reservation on Shared Buffer Type ATM Switch for Data Communication," *IEEE ICC*, pp. 391-396, 1994.

[17] Tobagi, F. A. and Kwok, C., "The tandem Banyan switching fabric: A simple high-performance fast packet switch," *Proc. IEEE, Infocom'91*, pp. 1245-1253.