

VLSI Implementation of a Switch Fabric for Mixed ATM and IP Traffic

Chi-Ying Tsui, Louis Chung-Yin Kwan, Chin-Tau Lea
Department of Electrical and Electronic Engineering
Hong Kong University of Science and Technology
Clear Water Bay, Hong Kong

etsui@ee.ust.hk, eelouis@ee.ust.hk, eelea@ee.ust.hk

Abstract- A VLSI implementation of a multistage self-routing ATM switch fabric is presented. The size of the switch is 16x16 and can handle the OC-12 (622 Mbps) link rate. Based on a bit-slice architecture, the entire 16x16 switch is implemented using four identical chips. The switch has multiple paths, created by a randomizer in front of the routing stages, between each input-output pair. The switch uses an input/output-buffering scheme and contains no buffers inside the fabric. A priority structure, which supports four levels, allows the delay sensitive ATM cells to be switched with the shortest latency. It also enables the non-interleaving routing scheme of IP cells. The switch fabric was designed and fabricated using MOSIS 0.8 μ m technology and was tested to run at 93MHz with 3.3V supply voltage.

I. Introduction

Shared-buffer architecture has dominated the field of commercial ATM switches over the years for its simplicity and small size of physical memory. However the speed of the memory has to be scaled up with the number of ports. This limit is also encountered in the shared medium approach. As the demand of capacity soars, multistage architectures have become an inevitable trend in ATM switch design. Of particular interest to us is the Banyan-type network [1]. The control of the switching is distributed over each switching element. So, multiple input/output transmission happens in the same cycles. As a result, the switch is more scalable with respect to the size of the switch. In this paper, we describe a VLSI implementation of a multistage self-routing ATM 16x16-switch fabric, which is designed for mixed ATM and IP traffic [2,3]. It is designed to handle the line rate of OC-12 (622 Mbps). The switch is featured by simple routing strategy, internal non-buffering, and multiple in/out paths. Multiple paths are created by a randomizer in front of the routing stages. The switch uses in/out-buffering scheme, and as a result, the order of cell transmission is always maintained.

II. A 16x16 ATM/IP Switch

The 16x16 ATM/IP switching system is shown in Fig.1. The switch is designed to handle ATM cells and IP packets simultaneously. The I/O processors contain ATM and IP queue multiplexed into the switch fabric. Although an IP packet is fragmented into ATM cells, inside the network no IP packet interleaving occurs, i.e., consecutive IP cells belong to the same IP packet. This greatly simplifies the IP processing but presents a major challenge for the switch design. It is accomplished with a combination of priority, unbuffered switch fabric and correct buffer management in the port processor. It is achieved by assigning the lowest priority to the first cell of an IP packet. It

guarantees that once the first cell of an IP packet reaches a destination port, the remaining cells will be transmitted with a higher priority than other IP packet cells destined for the same output port. The other IP packets can only be received until the current IP packet transmission is over [2,3]. For unbuffered switches, cells which are blocked will re-try in the next time slot. As a result, the sum of the latency of the switch fabric and the time for the acknowledgement signal to transmit back has to be less than the whole cell transmission time. The proposed switch fabric is a multistage interconnection network(MIN) as shown in Fig.2. Four 4x4 switching elements(SE) are used in each stage instead of the usual 2x2 SEs. For the same input/output port number, more stages are needed for a smaller SE design. So, both the latency and internal blocking probability are higher in a switch with 2x2 than the one with 4x4 SEs. In the proposed switch, two extra stages are placed in front of the 2 routing stages. The path setup in these 2 extra stages is randomized. For a particular virtual circuit connection, the path the cells go through inside the switch is different for different time slot. It is shown that randomization is necessary for internal blocking switch to provide immunity to congestion[3].

The operation of the switching system is as follows, the Input Processor processes the input cells. It captures the ATM/IP cell header and then look up information in the translation table. An Internal tag will be attached in the head of cell. Firstly, the cells go through 2 stages of randomizer. The randomizer randomly selects one non-blocking in/out combination and set-up the paths in the crossbar accordingly. The two randomizing stages provide paths from each input port to any one of the 16 entrances of the routing stages. In the routing stage, the control determines the routing from the tag information which contains destination, priority and activity. Since multiple paths are created by the randomizer, for each virtual circuit connection, the path it goes through is different every cell time slot.

III. Switching Element Design

The data transmission is designed to be either 8-bit or 16-bit in parallel. The switch design uses a bit-slice approach. It is formed by multiple 4-bit 16x16 switch chips to minimize the number of I/O on each chip. For a 8-bit system, as shown in Fig.3, two identical slices are used and they operate independently. Consistency in the internal path setup is ensured by duplicating the cell tag to each slice and making the pseudo-random sequences in the randomizer of each slice to be the same. The basic component of the switch is the 4x4 switch element which can be used for both the router and the randomizer. It is a 4-bit 4x4 crossbar switch with an extra one-bit crossbar for the acknowledgement signal. Fig.4 shows the structure of a 4x4 router. There are five one-bit datapath which

contains pipeline registers and a crossbar. Two stages of registers are required as the control unit needs the tag, which is available in two consecutive clock cycles, to setup the paths in the crossbar. The control unit(ROUT) compares the destination fields from the 4 inputs. The results showing any output port conflict between 2 inputs are restored in a register. In the next cycle, the comparator compares the priority fields to determine the relative priority between any 2 inputs. The arbiter will then resolve the destination and priority conflict and the decoder decodes the destination and the masked activity field to generate 16 path control signals for the crossbar setup.

For the 4x4 randomizer SE, the structure is basically the same as the router except that only one stage of register is needed as the header is not required for the path setup. A counter and ROM are used to generate the pseudo-random sequence. The ROM table stores the predefined path set-up control signals. For non-blocking output path set-up, the number of combination is equal to $4P_4(24)$. As a result, the ROM table has only 24 pre-determined entries. A counter is used to generate the address for the ROM. It counts from 0 to 23 and counts whenever a new cell transmission cycle starts. The counter receives a start-of-cell signal from the source and a new destination pattern is then loaded. It is desirable to provide a different pseudo-random destination sequence in different randomizers in a slice. In our design, two bits from the tag are arbitrarily chosen as an offset to the counter whenever a new cell transmission cycle starts. The offset is the same for the corresponding randomizers in all bit slices. The modified count value is then applied to the ROM address input. This method has two advantages. The destination pattern is different in every randomizer in a chip and a fixed sequence no longer exists which assure the quality of randomness.

IV. Results

The 16X16 switch fabric has been implemented and fabricated using MOSIS 0.8um CMOS technology. Fig. 5 shows the chip-photo and the layout of the switch fabric (A more detail die photo will be submitted later). The design parameters and measurement results are summarized in Table 1. With the chip running at 93MHz, the maximum throughput is about 1.5 Gbps per channel for a 4 chip, 16 bit switch. Because of blocking, the effective throughput is less, but is still more than adequate to handle the line rate of 622Mbps.

Operating Voltage	3.3V
Transistors Count	44,958
Chip Area	5.6*4.6 mm ²
Maximum Frequency	93MHz

Table 1. Chip Design Summary

REFERENCES

[1] C-L Wu and T-Y Feng, "On a class of multistage interconnection networks," IEEE Trans. On Computers, vol-29, no. 8, pp. 694-702, Aug. 1980.

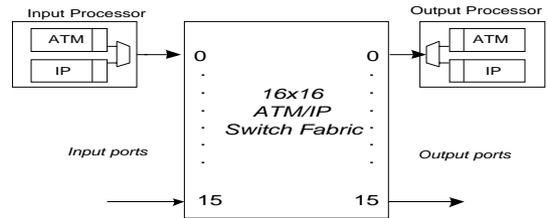


Figure 1. 16x16 ATM/IP Switch

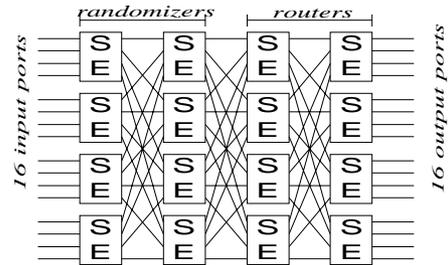


Figure 2. Multistage Self-Routing Switch.

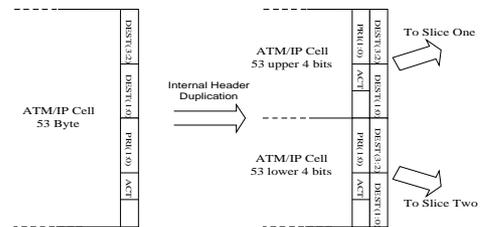


Figure 3. Internal Header Duplication

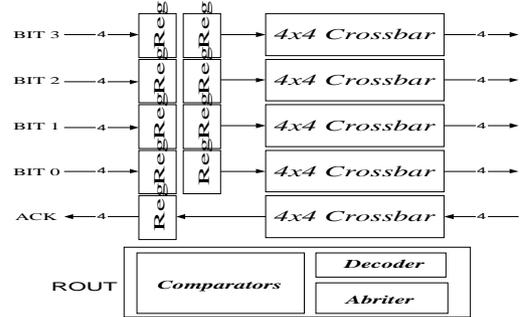


Figure 4. 4x4 Router Structure

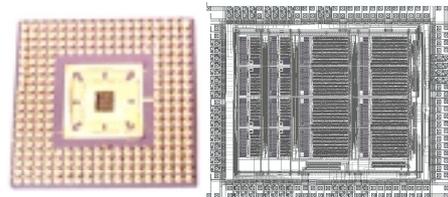


Figure 5 Chip Photo and Layout of the ATM switch fabric

[2] Chin-Tau Lea, Chi-Ying Tsui, Bo Li, Louis Kwan, Stanley Chan, Angus Chan, "A/I Net: A Network That Integrates ATM and IP", in *IEEE Network*, Jan/Feb 1999.

[3] Chin-Tau Lea, "A Multicast Broadband Packet Switch", *IEEE Transactions on Comm.*, pp.621-630, April 1993.