

FA 7.3: A 5Gb/s 8 x 8 ATM Switch Element CMOS LSI Supporting Five Quality-of-Service Classes with 200MHz LVDS Interface

Yasuo Unekawa, Keiko Seki-Fukuda¹, Kenji Sakaue¹, Takehiko Nakao, Shin'ichi Yoshioka, Tetsu Nagamatsu, Hideaki Nakakita, Yasuyuki Kaneko, Masahiko Motoyama, Yoshihiro Ohba, Koutarou Ise, Masayoshi Ono, Kuniyuki Fujiwara, Yuichi Miyazawa, Tadahiro Kuroda, Yukio Kamatani, Takayasu Sakurai, Akira Kanuma

Toshiba Corp./Toshiba Microelectronics Corp., Kanagawa, Japan

The switch element (SE) is a 622Mb/s, 8 x 8 shared-buffer ATM switch LSI for backbone LAN and WAN applications. The SE has 5Gbps bandwidth, supporting 5 QoS classes delay priority and link-by-link multicast. Up to a 32x32 switch with 20Gbps bandwidth can be configured using multiple SEs and distributor/arbitrator (DA) LSIs.

Figure 1 shows the logical queue structure realized in SE. Since the best-effort classes will require much larger per-link buffer than that in SE, the SE can dedicate a small-cell buffer to each class and link, and feed the buffer congestion status signals back to switch access (SA) LSIs that use the expandable per link huge buffer memories. As a result of throughput simulations and chip size consideration, the 64-cell buffer per class, which corresponds to the total buffer size of 320 is selected. Link-by-link multicast function is supported using a shift register type address generator.

Figure 2 shows the block diagram of SE. The input cell interfaces establish bit synchronization and cell synchronization, and transform cell data bit-width from 4b to 128b. The routing information is extracted from the cell header. The control logic contains cell counters, threshold registers and comparators. The routing information is transferred to the control logic to decide the cell to be received. The cell data is written to and read from the cell buffer with 128b in parallel sequentially from link0 to link7. Read/write addresses of the cell buffer are controlled by a shift register type address generator using the information of cell counters and flow-control signals received from DA or SA. The control logic generates flow-control signals in time-multiplexed fashion to upstream LSIs using flow-control signals received from the downstream LSIs, and the status of the SE cell counters to avoid the cell buffer overflow. The output cell interfaces transform the bit-width from 128b to 4b. Finally, LVDS output buffers convert the level of output cells from CMOS to LVDS.

Each link to or from the SE consists of 5 pairs of LVDS buffers - 4b data and 1b clock. LVDS interfaces have the following advantages: 1) low-power consumption at high-speed operation, 2) high CMRR and low EMI, 3) low-cost, 4) small-signal distortion, and 5) few level conversion ICs through backplanes and PCBs. Figure 3 shows the schematic diagram of LVDS driver and receiver for the SE.

The shift-register type address generator shown in Figure 4a has a dedicated structure for read/write address management of the cell buffer. The shift register is more advantageous in implementing the multicast function than the linked-list because unicast cells and multicast cells can be handled in the same way. Only one clock period is needed to enqueue a multicast cell for any combination of output links. The address generator in the SE has 320 entries, each with 8b for the output link map, 1b for the multicast identifier, 3b for the class of delay priority, 9b for the address pointer to the shared buffer and 1b for the address pointer parity.

Since each data line running through the address register is connected to 320 flip-flops, the data line has high capacitance and long RC delay. Therefore, the data line swing has been reduced and some of the flip-flops were replaced with the sense amplifying flip-flops [2]. The address generator has three operation modes: enqueueing, dequeueing, and shift. Each operation is performed in a clock period. Enqueueing is performed when a cell data is input to the cell buffer. First, the vacant entries in the address generator are searched in parallel by the hierarchical search circuit shown in Figure 4b. Then the output link map, the multicast identifier, and the class of delay priority are written to the vacant entry which is the lowest in the column, and the address pointer held by that entry is used to write a cell data into the cell buffer. Dequeueing is performed when a cell data is outputted from the cell buffer. In this operation, the entries which have the specified output link and class of priority are searched. If the search is successful, the address pointer in the lowest entry to be found is transferred to the cell buffer to read out the cell. If the search fails, an idle cell is output. Shift suppresses the vacant entries. In this operation, vacant entries are searched in the address generator. If a vacant entry is found, the output link maps, the multicast identifiers, and the classes of delay priority and the address pointers above that entry are shifted downward. Each of the above mentioned operations is repeated 8 cycles during one cell period for 8x8 switching.

The SE supports ready type flow control. That is, backpressure signals or non-ready signals are immediately generated when the number of input cells exceeds the flow control threshold. Backpressure signals are serialized by each link and transferred in parallel with the bit clock and the synchronization signal. As for the guaranteed traffic such as CBR and rt-VBR, backpressure signals generated are global since the bandwidth of the traffic is well regulated. As for the best-effort traffic such as nrt-VBR, ABR, and UBR, link-by-link backpressure signals and global backpressure signals are generated for unicast traffic and multicast traffic, respectively, since the bandwidth of the traffic is difficult to predict in advance.

The SE uses 0.35 μ m double-metal CMOS. Figure 5 shows the SE chip layout. Power consumption at 200/100/50MHz, 3.3/2.5V operation is estimated to be approximately 4.0W. The die is 17.5x17.5mm² and is packaged in a 447-pin ceramic PGA. Figures 6 and 7 show measured waveforms of LVDS signals transferred between PCBs through a 30cm backplane and a chip micrograph of 200MHz LVDS interface.

Acknowledgments:

The encouragement by Y. Unno, O. Ozawa, J. Iwamura and K. Maeguchi is appreciated.

References:

- [1] ATM Forum, "Traffic Management Specification Version 4.0," June, 1995.
- [2] Matsui, M., et al., "200MHz Video Compression Macrocells Using Low-Swing Differential Logic," ISSCC Digest Of Technical Papers, pp. 76-77, Feb., 1994.

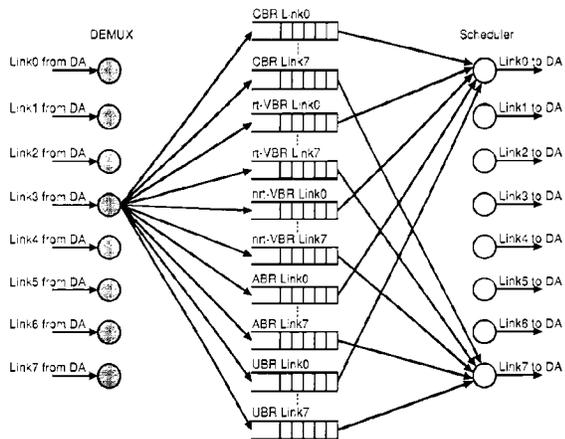
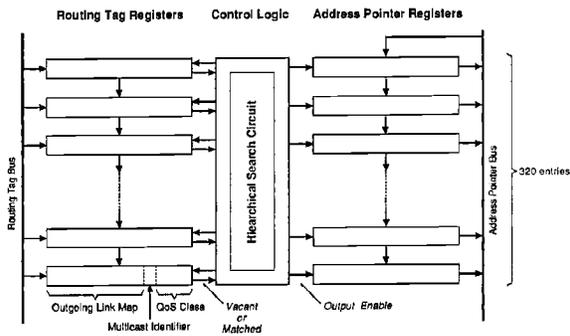


Figure 1: Logical queue structure.



(a)

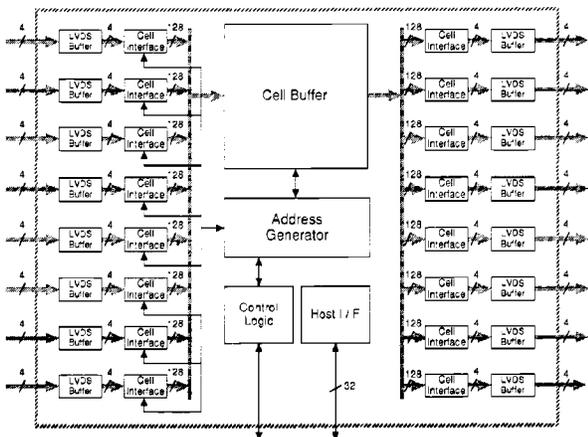
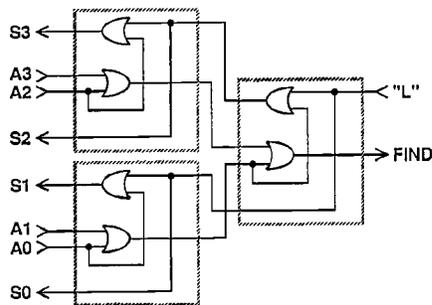


Figure 2: Functional block diagram.



A_n : to be "H" level when the entry n is matched with the searching condition.
 S_n : to be "H" level when any entries lower than the entry n is matched with the searching condition.
 ($n=0, 1, 2, 3, \dots$)

(b)

Figure 4: (a) Shift register type address generator. (b) Hierarchical search circuit.

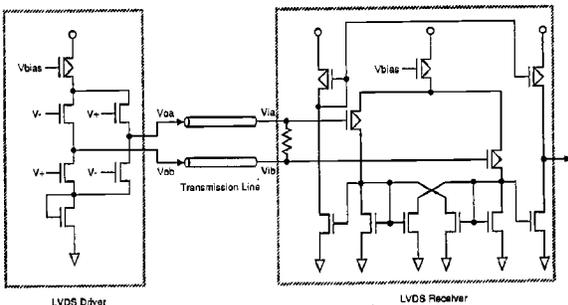


Figure 3: LVDS driver and receiver schematic.

Switch size	8x8 ports
Switch size expandability	up to 32x32 ports (with extra LSIs)
Cell buffer size	320 cells
Number of QoS classes	5
Flow control	Yes
Multicast function	Yes (by link)
Cell discarding	Yes (based on CLP)
Others	Dynamic flow control threshold setting Input link rotation

Table 1: ATM switch element LSI functions.

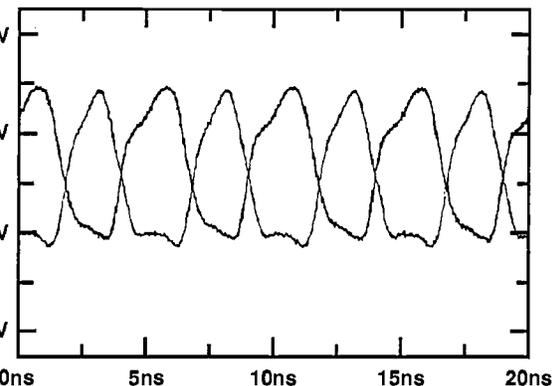


Figure 6: LVDS measured waveform.

Figure 7: See page 428.

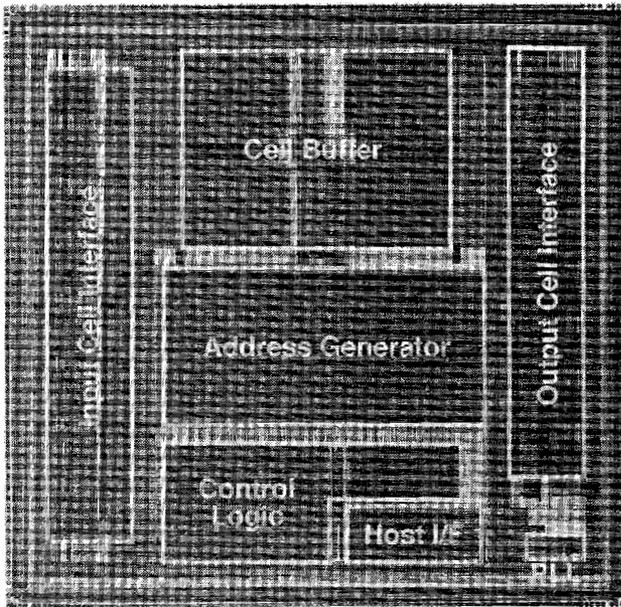


Figure 5: Chip layout.

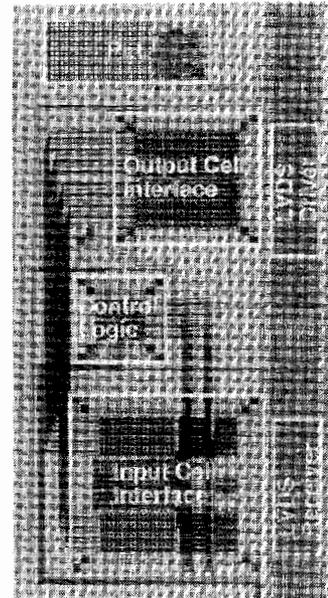


Figure 7: LVDS interface chip micrograph.

FA 7.4: A 2.5GB/s 32:1 / 1:32 Multiplexer / Demultiplexer Chip Set for SONET Communications
(Continued from page 121)

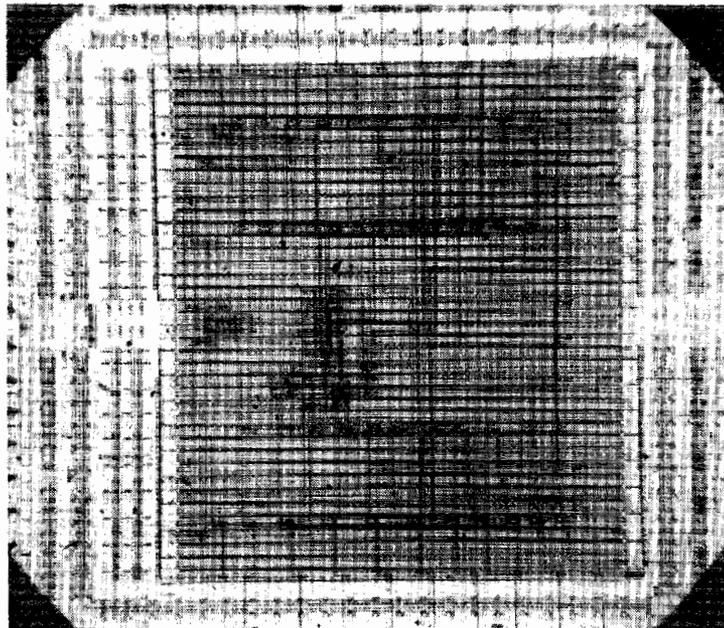


Figure 5: 1:32 demultiplexer microchip plot.