

FA 9.4: 350MHz Time-Multiplexed 8-port SRAM and Word-Size Variable Multiplier for Multimedia DSP

Toshinari Takayanagi, Kazutaka Nogami, Fumitoshi Hatori, Naoyuki Hatanaka, Makoto Takahashi, Makoto Ichida, Shinji Kitabayashi¹, Tatsuya Higashi¹, Mike Klein², John Thomson², Roger Carpenter², Ravi Donthi², Denny Renfrow², Jason Zheng², Liane Tinkey², Brandi Maness², Jim Battle², Steve Purcell², Takayasu Sakurai

¹Toshiba Corp., Kawasaki, Japan

²Toshiba Microelectronics Corp., Kawasaki, Japan

³Chromatic Research, Mountain View, CA

A multimedia DSP optimized for digital audio/video applications provides simple flexible cost-effective solution capable of GUI acceleration, MPEG2 decoding, real-time MPEG1 encoding, personal video conferencing, 28.8kbps fax/modem, and audio/sound functions [1]. The main frequency of the chip is 62.5MHz and the supply voltage is 3.3V. The chip is fabricated in 0.5µm triple-metal CMOS, occupies 12.8x14.0mm² and is mounted in a 240 QFP package with a heat-spreader. The chip integrates high-performance custom macro blocks: an interface for Rambus DRAMs (RAC), a 37kb time-multiplexed 8-ported SRAM, 72b scalable datapath and single oxide 3V/5V I/O. The focus here is the SRAM and word-size-variable multiplier.

The 72b x 512-word 8-port SRAM with 36Gb/s bandwidth acts as an instruction/data cache, register file, and buffer queues. Although memory bandwidth is crucial for this type of a highly-concurrent multimedia processor, the area penalty of conventional multi-port SRAM is high. Therefore a time-multiplexed multi-port scheme that performs four serial writes and two serial reads in a cycle using 3-ported memory cells (two read ports and one write port) is used. Write bit line and read bit line are single-railed to further reduce area. The memory cell is 25.8x9.2µm². Decoders and read/write circuits are reduced by time-multiplexed architecture, so the area advantage is high.

The write circuitry is shown in Figure 1 and the timing diagram is described in Figure 2. Write addresses and data of port0,1,2,3 are time-multiplexed by four-phase pulses: PULSE0,1,2,3, that are generated by a digital DLL. Since address decoding naturally takes more time than driving a bit line, a write data input is triggered by a one-phase-later pulse than a write address input. The margin narrows because of wordline delay variation, bit line delay and time-multiplexing pulse width. Phase inversion of addresses (AN[0] and A[0] in Figure 1 for example) is at the input of the 4:1 multiplexers instead of at the output to equalize the number of gate stages of the paths. To maximize margin of the time-multiplexed SRAM, a cycle is divided into four equal phases. In the DLL an input clock goes through 4 stages of variable delay line (VDL) and the fourth output is compared to the initial input clock by a phase lock detector (PLD). PLD provides UP/DOWN signals to a delay control counter (DCC) whose outputs control VDL delay. A feedback loop of VDL, PLD and DCC delays the fourth VDL output exactly by one cycle from the initial clock input. Thus four clocks: CLK0, 1, 2, 3 whose phase is shifted by one fourth of the cycle are generated. One advantage of the digital DLL approach over analog is that it freezes the VDL delay by holding registers of the DCC by a HOLD signal. The chip clock frequency is changed without affecting the four SRAM internal clocks, useful in chip testing and debugging.

The read circuit is shown in Figure 3. Read is twice in one cycle. The address is multiplexed by CLK1. For fast access with single-railed bit lines, an EPROM-like dummy memory cell is used with

a current-mode latch sense amplifier [2]. When accessing a bank, a dummy memory cell of the opposite bank on the other side of a sense amplifier is also accessed. The bit line of the opposite bank acts as a dummy bit line that provides a reference level. The dummy memory cell sinks half the current of the real memory cell. A SRAM macro test chip with latch-to-latch delay measurement circuits, is fabricated separately from the multimedia DSP. Measurements show that the SRAM operates at 87.7MHz at 3.0V and room temperature, indicating that internal write is at 350MHz.

The multiplier is key in the multimedia DSP because filtering, discrete cosine transform (DCT) and other signal processing functions frequently use multiply-and-accumulate (MAC). A word-size-variable multiplier provides high efficiency in the utilization of hardware resources. The multimedia DSP incorporates two 24x24 multipliers and two 18x18 multipliers, each of which can also perform two concurrent 9x9 multiplications. This multiplier array finishes carry-save addition while another ternary adder unit performs final carry-propagate addition in the following cycle. The multiplier employs bit-paired Booth decoding and Wallace tree structure coupled with 4:2 compactors.

Figure 4 illustrates how the two concurrent 9x9 multiplications (MUL9 mode) work in the 18x18 multiplication array. A parallelogram in the figure represents a partial product selector (PPSEL) array. In the MUL9 mode, the upper right region is used for the lower 9x9 multiplication (A0 x B0) and the lower left region is used for the upper 9x9 multiplication (A1 x B1). Outputs of the PPSEL cells in the shaded regions are forced to "0" in the MUL9 mode. One complication is that the border of two 9x9 multipliers falls in the middle of a bit pair [9:8] of Booth decoding. This is worked around by duplicating Booth decoder cell of that bit pair and dividing the Booth decoder lines in the middle of the PPSEL row. Two Booth decoder cells control the right half of the row and the left half separately in MUL9 mode, while they control the same in 18x18 mode. In addition it is necessary to kill carry propagation across the boundary of the two 9x9 multipliers in carry-save addition stage because the upper multiplication result is broken otherwise. In this way each 9x9 multiplication result automatically flows into appropriate portions of the registers.

Figure 5 shows the 4:2 compactor cell using CMOS pass transistor logic. Dual-rail logic is used in the cell for speed, while the outputs are single-railed because the wiring area penalty is large in block-level routing. This compactor cell is about 50% faster than conventional CMOS combination adder. The carry kill function is realized by replacing output inverters of specific portions with NOR gates. The area of these additional gates are absorbed by adjacent cells. In PPSEL row, area overhead of implementing additional transistors for the MUL9 mode is small since the PPSEL layout is limited by metal. Thus area for the MUL9 mode is about 3%. The SRAM micrograph is shown in Figure 6. The 24x24 multiplier micrograph is shown in Figure 7.

Acknowledgments:

Encouragement by W. Patterson, M. Farmwald, D. Burns, J. Iwamura, K. Maeguchi, K. Kawate, O. Ozawa, and Y. Unno is appreciated. Chip fabrication and support by H. Hayashida, K. Hama, M. Sato, H. Momose, K. Kawate, Y. Nitobe, N. Kiryu, J. Tsunoda, M. Yoshitomi, and S. Ishiwata are acknowledged.

References:

- [1] Purcell, S., "A VLIW and SIMD Vector Processor for PC Multimedia," Microprocessor Forum, Oct., 1995.
- [2] Kobayashi, T., et al., "A Current-mode Latch Sense Amplifier and a Static Power Saving Input Buffer for Low-Power Architecture," Symp. on VLSI Circuits, pp. 28-29, June, 1991.

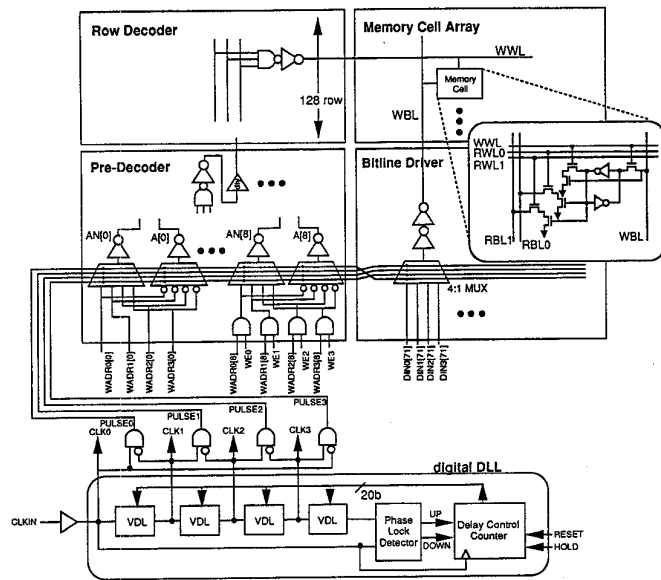


Figure 1: SRAM write circuit.

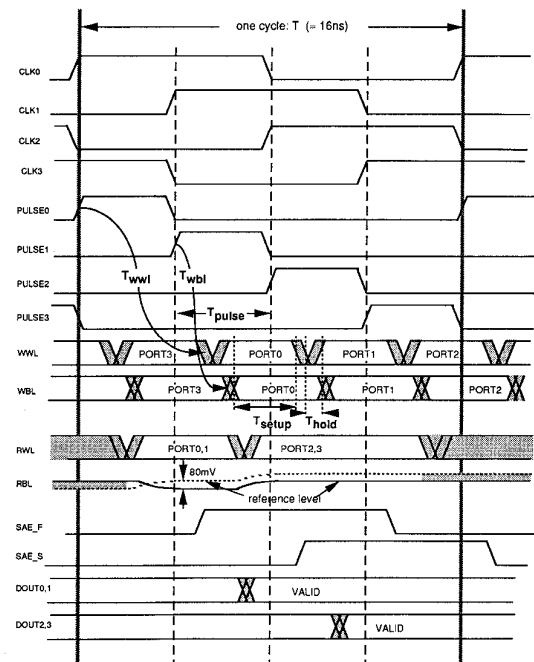


Figure 2: SRAM timing diagram.

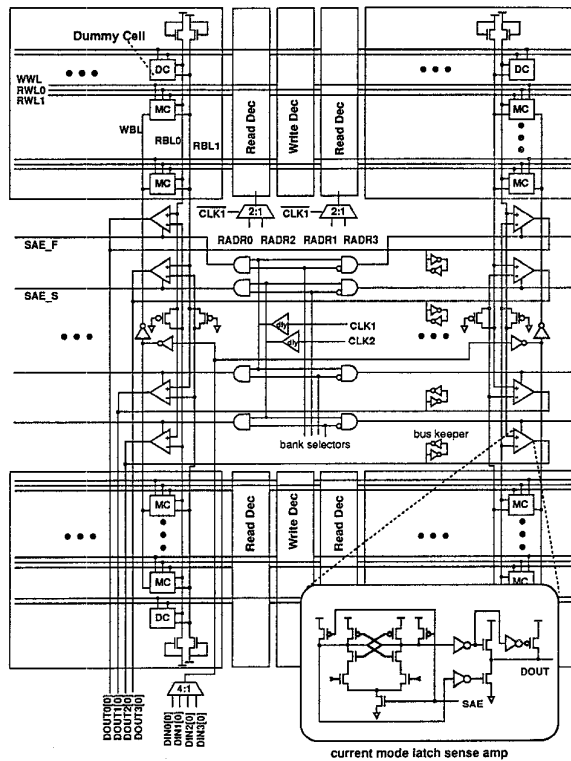


Figure 3: SRAM read circuit.

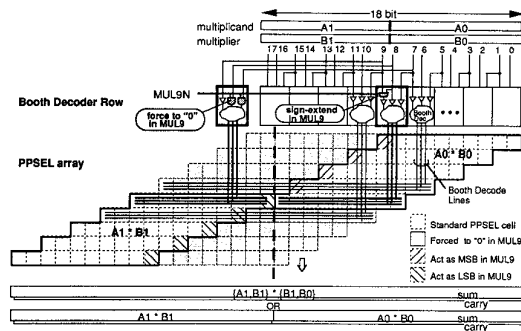


Figure 4: 18x18 multiplier partial product selector array.

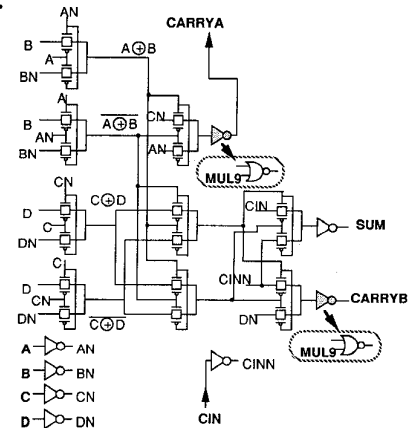


Figure 5: 4:2 compactor circuit.

Figures 6 and 7: See page 451.

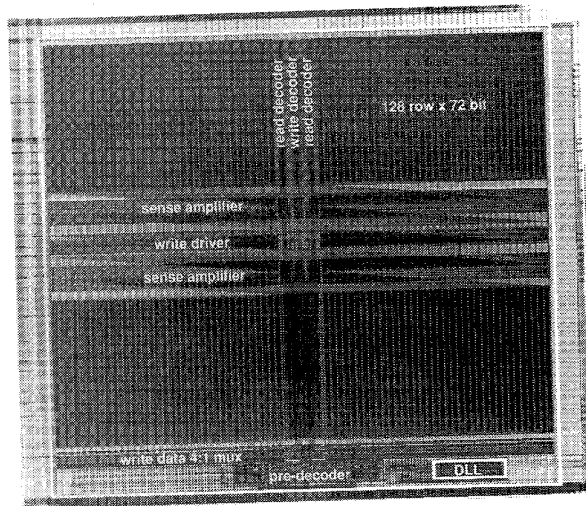


Figure 6: SRAM macro micrograph.

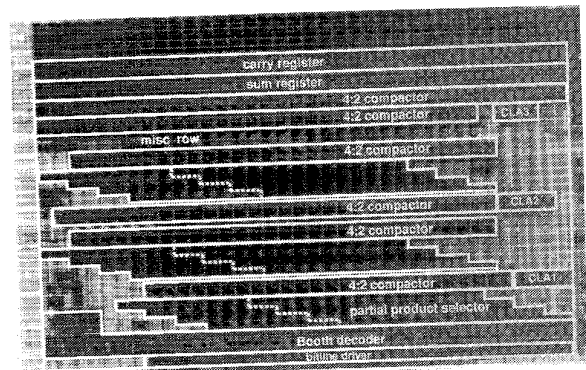


Figure 7: 24x24 multiplier macro micrograph.

FA 9.5: A 1V 100MHz 10mW Cache using Separated Bit-Line Memory Hierarchy and Domino Tag Comparators
(Continued from page 153)

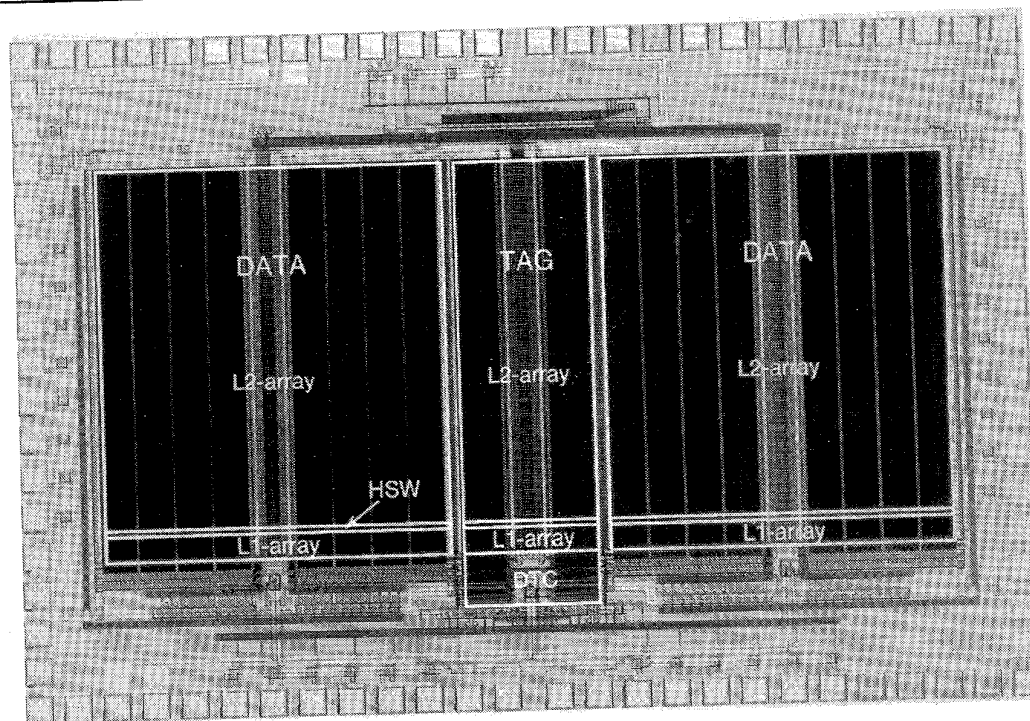


Figure 7: Chip micrograph.