# A 110-MHz/1-Mb Synchronous TagRAM

Yasuo Unekawa, Tsuguo Kobayashi, Tsukasa Shirotori, Yukihiro Fujimoto, Takayoshi Shimazawa, Kazutaka Nogami, Takehiko Nakao, Kazuhiro Sawada, Masataka Matsui, Takayasu Sakurai, *Member, IEEE,* Man Kit Tang, and William A. Huffman

***Abstract*— A 4-way set associative TagRAM with 1.189-Mb capacity has been developed which can handle a secondary cache system of up to 16 Mbytes. A 9-ns cycle operation and clock to $D_{out}$ of 4.7 ns are achieved by use of circuit techniques such as a pipelined decoding scheme, a single PMOS load BiCMOS main decoder, a BiCMOS sense-amplifying comparator, doubly placed self-timed write circuits, and highly linear VCO for a PLL. The device is successfully implemented with 0.7-$\mu$m double polysilicon double-metal BiCMOS technology.**

## I. Introduction

The performance of recent microprocessors is remarkable due to the progress in VLSI technology and the innovation of computer architecture. The operating frequency of most microprocessors will soon reach 200 MHz. A cache memory is indispensable for recent computer systems to fully utilize such high-speed microprocessors by increasing memory bandwidth [1]–[5],

To date, microprocessors with on-chip cache have been popular in high performance computer systems. However, the capacity of such an on-chip cache is limited up to 32 K bytes by the chip size constraint. This size of capacity is not sufficient to handle data for really high-end applications such as image processing and graphics. A large off-chip secondary cache should be introduced to build a hierarchical cache system in order to reduce the cache miss ratio of the smaller on-chip cache [4], [5].

The synchronous TagRAM described in this paper stores address tags and status bits of cached data and can be used to build the secondary cache system of up to 16 Mbytes. In order to handle the large secondary cache, the present TagRAM contains 1.189 Mb of four-transistor (4T_ SRAM cells which is the largest capacity ever reported for TagRAM's

In Section II, conflicting requirements on cache memories are described and a new architecture called "data streaming cache architecture" is proposed. The section also touches on the role and the required characteristics of the TagRAM used in the data streaming cache architecture. In Section III, an overview of the TagRAM features and its memory core architecture for a Tag look-up operation are shown, In Section IV, circuit design details of the TagRAM are explained in relation to the characteristics required for the TagRAM. Process technology and performance of the TagRAM are described in Section V. Section VI summarizes this work.

Manuscript received August 18, 1993; revised October 15, 1993.

Y. Unekawa, T. Kobayashi, Y. Fujimoto, T. Shimazawa, K. Nogami, T. Nakao, K. Sawada, M. Matsui, and T. Sakurai are with the Semiconductor Device Engineering Laboratory, Microelectronics Center, Toshiba Corporation, Kawasaki, 210 Japan.

T. Shirotori is with Toshiba Microelectronics Corporation, Kawasaki, Japan.

M. K. Tang and W. A. Huffman are with Silicon Graphics Inc., Mountain View, CA.

IEEE Log Number 9216553.

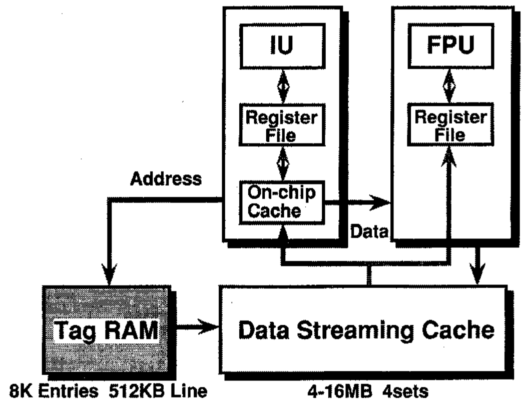



Fig. 1. Data streaming cache architecture.

## II. Data Streaming Cache Architecture

From the standpoint of advanced computer architecture, there exists conflicting requirements on the performance of cache memories. In processing integer data, it is important to eliminate the wait cycle using a fast cache. For this purpose a large capacity cache is not suitable because the integer data tends to have fairly high locality and the large capacity memory tends to be slow in nature. On the other hand, in floating-point processing, for example, in processing image and/or graphics data, the locality in time and space is much lower than the integer data. This means that the larger capacity is essentially important for floating-point data. The high speed requirement for the cache, however, is not so critical for these kind of data because the processing time itself is time consuming. To improve the computer performance for real applications, both classes of data must be handled properly.

To meet the completely different requirements mentioned above, the data streaming cache architecture shown in Fig. 1 is proposed. This architecture is basically a split level cache. A small on-chip cache provides fast access for integer data including addresses, and on the other hand, a large off-chip cache provides sustained high bandwidth for floating-point data. Then, the integer load/store unit in the integer unit IU uses the off-chip cache as a second level cache while the floating-point load/store unit in the floating-point unit FPU bypasses the on-chip cache and uses the off-chip cache as the first level cache. There is no contention between the IU and the

TABLE I
TAG RAM FEATURES

| | |
|---|---|
| Host CPU | TFP processor |
| Mapping | 4-way set-associative |
| Entry size | 8K lines |
| Line size | 512 bytes |
| Tag size | 8K entries × 4 ways × 20 bits |
| State size | 8K entries × 4 ways × 12 bits |
| Dirty size | 8K entries × 4 ways × 4 bits |
| Total bit count | 1.189 M bits |
| Redundancy | 8 rows |
| Other features | Integrated Dirty bit logic |
| | Self-timed write |
| | On-chip PLL |
| | JTAG supported |

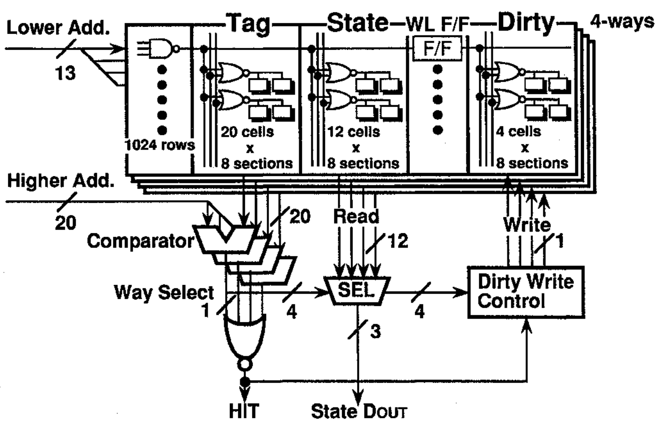



Fig. 2. Memory core architecture for Tag look-up operation.

FPU even when an on-chip cache miss occurs. This is because the IU does not continue parsing instructions while it handles an on-chip cache miss. The FPU also stops parsing instructions while the IU is handling the on-chip cache miss. When the on-chip cache has been filled, the IU again begins parsing instructions including the floating-point memory operations which access the off-chip cache.

Thus the requirements of both IU and FPU can be fulfilled at once in the hierarchical memory system. System performance improvement realized by the data streaming cache architecture over the conventional cache architecture mostly depends on the types of code which run. If the data would fit in the on-chip cache, the improvement will be none. However, if the on-chip cache miss would occur for every floating-point access, the improvement will reach a factor of twelve.

In the system under consideration, the TagRAM is used to support the off-chip cache of up to 16 Mbytes built with commodity synchronous SRAM's. The main function of the TagRAM is to generate a hit-miss indication of whether the desired data exists in the SRAM after comparing the incoming addresses to the read-out address tags. Hence, the TagRAM is a key device in achieving a high-speed cache system.

In general, the characteristic required for a TagRAM to build a high performance cache for high-speed microprocessors is

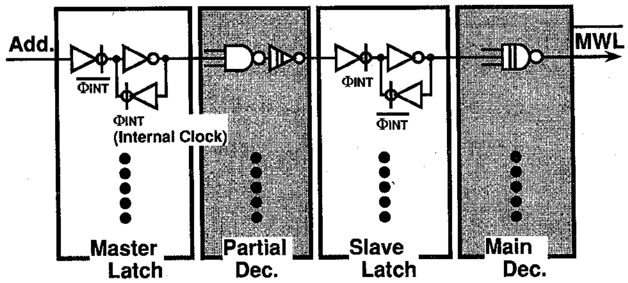



Fig. 3. Pipelined partial decoding scheme.

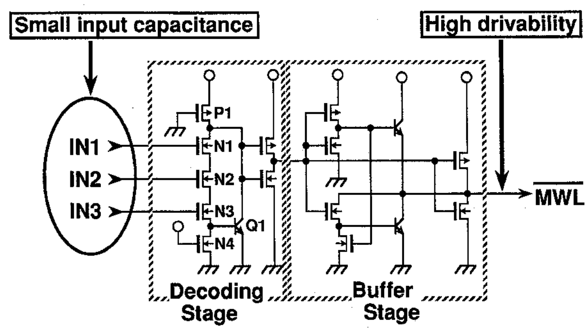



Fig. 4. Single PMOS load BiCMOS main decoder.

a short cycle time. In addition to this characteristic, small clock-to-output delay is also an important parameter since the TagRAM has to drive large capacity SRAM's as mentioned above.

## III. TAGRAM FEATURES

Table I summarizes the features of the designed TagRAM. Host processor of the TagRAM under consideration is TFP processor described in [6]. The TFP is a 300 MIPS, 300 MFLOPS, 4–issue superscalar RISC processor. The TagRAM is made to support a 4–way set associative cache. In order to handle the large secondary cache, the TagRAM contains 8K entries × 4 ways × 20 b of Tag memory, 8K entries × 4 ways × 12 b of memory for State bits and, 8K entries × 4 ways × 4 b of Dirty bits. The total number of memory cells sum up to 1.189 Mb. The TagRAM also contains a comparator to compare read-out address tags with higher physical addresses. The State and the Dirty bits are the state of cached data which are used to maintain cache coherency. In the State bits, Virtual Synonym bits are included which are used to resolve the first level cache synonym problem.

There are five operation modes for the TagRAM—Tag look-up, Tag read, Tag write, State read, and State write—and the operations can be executed in a pipelined manner. The memory core architecture for a Tag look-up operation are shown in Fig. 2. In a Tag look-up operation, the physical address is provided to the TagRAM from the microprocessor. The physical address is split into higher and lower address. The Tag memory is accessed by the lower address and read-out tag is compared with the higher address. Consequently, indications of whether an address match occurred is provided as the HIT signal. In the case of a match, State bits corresponding to the way of the match are output in synch with HIT. The Dirty bits are

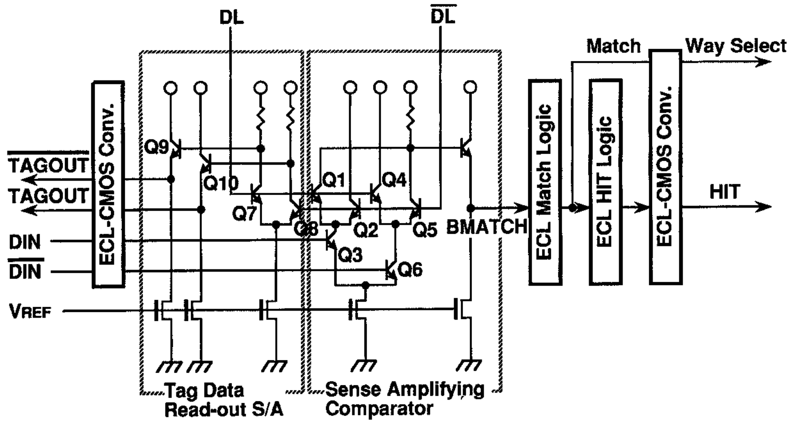



Fig. 5. Sense-amplifying comparator (SAC).

conditionally set to the way at the next cycle. The write control logic for Dirty bits is also integrated on chip.

Since the Dirty bit is written with the address received in the preceding cycle, a flip-flop is inserted between a State memory word line and a Dirty memory word line to hold the word line information for one cycle. An additional advantage of this configuration is the reduction of main word line capacitance which accelerates the Tag operation.

Modified double-word line structure [7] is adopted to unify the three kinds of memory subarrays—Tag, State, and Dirty—and to reduce the memory cell power consumption and word line delay. Each row has eight sections. and one section is activated at a time. Joint test action group (JTAG) boundary scan circuitry is also implemented to increase on-board testability.

## IV. Circuit Design Detail

The TagRAM is based on a fully synchronous design as opposed to the asynchronous designs seen in high-speed and large memory capacity requirements, several novel circuit techniques were introduced.

To achieve short cycle time, circuit techniques in address decoding such as pipelined partial decoding and single PMOS load BiCMOS main decoder were used. A sense amplifying comparator and ECL-based HIT signal generator were used to effectively reduce critical path delay.

In order to realize small clock to output delay, circuit techniques such as an on-chip PLL with a highly linear VCO and doubly placed write circuits were used.

### *A. Pipelined Partial Decoding Scheme*

In conventional partial decoding schemes, the address flip-flop is placed only at the address input. Partial decoding is started after the address is latched by the internal clock. So, the partial decoding time is included in the cycle time.

Fig. 3 shows a pipelined decoding scheme used to reduce the partial decoding time. In this scheme, partial decoders are placed in between master and slave transparent latches. Since the slave latch is placed after the partial decoders, The partial decoding can be done during address setup time. So, the partial decoding time can be invisible in the cycle time. If the master latch is also placed after the partial decoder, the master incorrectly latches the address of the succeeding cycle due to the internal clock delay and rather short address setup time. This scheme achieves a gain of 2 ns over the conventional scheme.

### *B. Single PMOS Load BiCMOS Main Decoder*

Fig. 4 shows the proposed BiCMOS main decoder. It consists of the decoding stage and the buffer stage. High drivability is compatible with small input capacitance. A normally on PMOS load P1 is adopted to minimize the input capacitance of the main decoder so as to reduce the driver delay of the partial decoder. To shorten the fairly large delay caused by the serially connected structure consisting of $N1$, $N2$, and $N3$, a bipolar transistor Q1 is added. This bipolar transistor enhances the drivability of the pull-down part of the decoder. In consequence, the pull-up device $P1$ can be designed to have large W/L which in turn realizes high-speed pull-up. The present circuit reduces the address decoding time by 0.5 ns compared to the conventional full CMOS decoder plus BiCMOS buffer scheme.

### *C. Sense-Amplifying BiCMOS Comparator*

Conventionally, a comparator for a cache is built with a MOS comparator inserted between a bit line and a BiCMOS sense amplifier. The output of the sense amplifier is a bit-wise match signal.

The newly proposed sense-amplifying comparator (SAC), whose circuit diagram is shown in Fig. 5, replaces the conventional MOS comparator with bipolar comparator [8] composed of $Q3$ and $Q6$ merged into a bipolar sense amplifier composed of $Q1$, $Q2$, $Q4$, and $Q5$. If the higher address on the DIN is high, the sense amplifier composed of $Q1$ and $Q2$ is activated. And if the read-out tag on the data line (DL) is also high which is the case of match, $Q1$ turns on and bit-wise match signal (BMATCH) outputs an ECL-low level. On the other hand, in case of no match, ECL-high level appears. This configuration achieves 0.5 ns delay reduction compared with the conventional scheme where a MOS comparator is used together with a BiCMOS sense amplifier.

In the match logic, BMATCH signals from each bit are ORed to generate MATCH signal. HIT signal is generated from the MATCH signals of each way in the same manner. All circuits from the bit line through to the HIT signal generator are ECL-based to reduce critical path delay [9].

Fast tag data read-out is also an important requirement in the present TagRAM. For this purpose, a tag data sense amplifier composed of $Q7$ to $Q10$ is placed in parallel to the SAC. The addition of this extra sense amplifier is straight forward, because an input of the extra sense amplifier is the same as the SAC input. This is in contrast to the conventional circuit where comparator outputs connect to BiCMOS sense amplifiers and either extra signal lines or extra circuitry are needed to get access to the required input for the extra sense amplifier.

### D. On-Chip PLL

A phase locked loop (PLL) is integrated on chip. The on-chip PLL is used to shorten the clock-to-output delay as well as to cancel internal clock delay. By adjusting the feedback delay to the PLL reference input, the internal clock delay, even if negative, can be set to an arbitrary value. If the internal clock is generated by just buffering the external clock, the internal clock inevitably delays by 3 ns due to the heavy load connected to it. This causes the degradation of clock-to-output delay.

The voltage controlled oscillator (VCO) is a key component of the PLL. The linearity over wide oscillation frequency to the input voltage is important to obtain a large lock frequency range and stable operation of the PLL. In Fig. 6, the proposed VCO (a) and the conventional VCO (b) are shown together with a measured linearity comparison (c). Highly linear oscillation of the present VCO comes from the large variable range of the effective resistance of the transfer gate and the large control current compared with the conventional VCO. Due to the high linearity, the PLL stably locks frequencies from 50 to 150 MHz. A measured jitter is 0.4 ns. The inclusion of the PLL on a chip can reduce cycle time by 1 ns.

A clock skew is another important issue in designing synchronous TagRAM's. A hierarchical and balanced clock tree for clock distribution minimizes the clock skew to less than 0.5 ns.

### E. Doubly Placed Self-Timed Write Circuits

In order to minimize the clock-to-output delay, sense amplifiers should be placed near the pads. This rules out the possibility of bit line partitioning and the placement of sense amplifiers at the center of a memory array. However, each bit line is highly capacitive due to the 1024+8 memory cells that are connected to it. The $RC$ delay of the bit line amounts to 2.5 ns and this hinders fast write and write recovery. Nevertheless, because of the inherently small bit line swing of only 0.2 V needed for the BiCMOS sense amplifier, it does not cause a problem during read-out.

In order to reduce the bit line $RC$ delay, a write circuit and bit line precharge circuit are placed at both ends of each bit

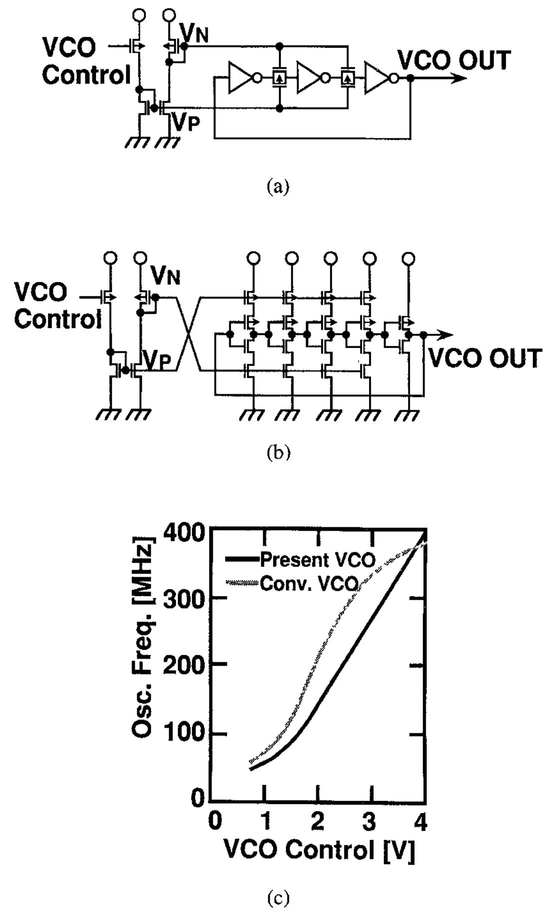



Fig. 6. Highly linear voltage control oscillator (VCO). (a) Present VCO. (b) Conventional VCO. (c) Measured property comparison.

line as shown in Fig. 7. This configuration reduces the write operation delay by 1 ns and eliminates the case where the write operation determines the cycle time. The write operation is controlled by a self-timed write pulse which is generated through a delay line. After receiving the write pulse, the bit line precharge takes place automatically.

## V. Simulated and Measured Results

The simulated waveforms of a Tag look-up cycle operated at 9 ns cycle time are shown in Fig. 8. This is an after-layout simulation where the precise values of wiring resistance and capacitance are extracted from the fixed layout data. To improve the accuracy of the simulation, coupling capacitances of adjacent pairs of bit lines were calculated and incorporated in the simulation.

The delay time distribution of a Tag look-up cycle operated at 9 ns cycle time is shown in Fig. 9. By the use of the pipelined partial decoding scheme, partial decoding was completed before the rising edge of the internal clock. So, the internal cycle starts from the main decoding. The single PMOS load BiCMOS main decoder and the sense amplifying comparator contributed to shortening the main decoding time and BMATCH signal generation respectively. The on-chip PLL was effectively used to reduce the clock-to-output delay.

The minimum clock cycle time is 9 ns in typical conditions, which corresponds to 110 MHz clock frequency. If the

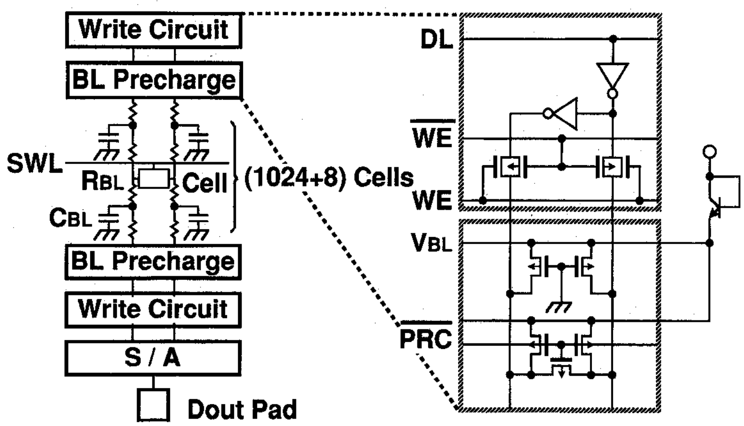
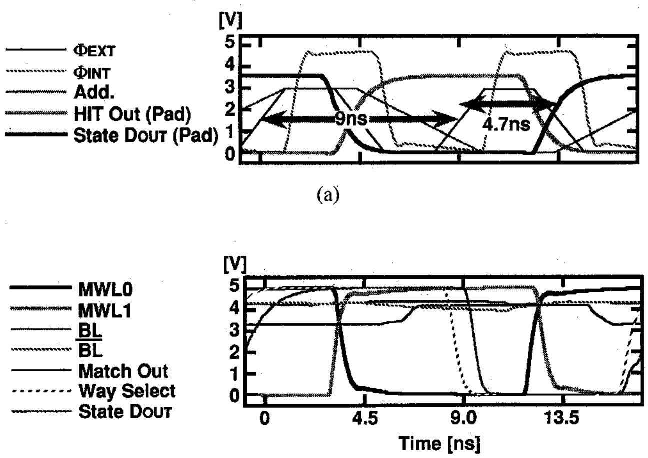



Fig. 7. Doubly placed precharge and write control circuit.




(a)

(b)

Fig. 8. Simulated waveforms of a Tag look-up cycle.

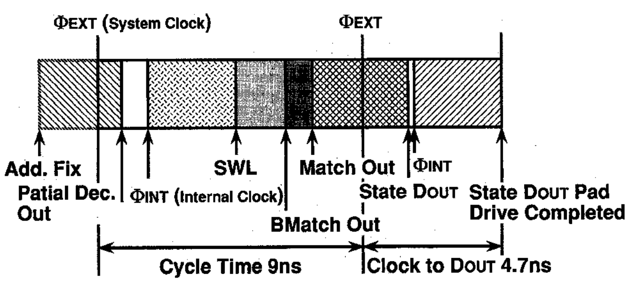



Fig. 9. Distribution of delay time in a Tag look-up cycle.

TagRAM is designed with a pure CMOS technology without the circuit ideas mentioned before, the clock cycle time is estimated to be 18 ns. If the RAM is designed with BiCMOS technology without the circuit ideas described in Section IV, the clock cycle time is estimated to be 13 ns. So, the new circuit ideas give rise to the delay improvement of 4 ns.

Fig. 10 shows the chip microphotograph. The chip size is 14.8 mm × 14.8 mm, and 5.034M transistors are on the chip. Peripheral circuits surrounding the memory core macro were designed using standard cell methodology. The device was implemented with 0.7–um double-polysilicon and double-

TABLE II
PROCESS TECHNOLOGY AND PERFORMANCE

| | |
|---|---|
| **Process Technology** | |
| Technology | 0.7-$\mu$m double-polysilicon/double-metal BiCMOS process |
| Memory cell | Highly resistive polysilicon load 4T SRAM cell |
| Cell size | 8.0 $\mu$m × 4.8 $\mu$m |
| Chip size | 14.8 mm × 14.8 mm |
| Cell occupancy | 20.8% |
| Package | 155-pin ceramic PGA |
| **Performance** | |
| Supply voltage | 5.0 V |
| Cycle time | 9.0 ns |
| Clock to $D_{out}$ | 4.7 ns |
| Power dissipation | 3.0 W (at 75 MHz) |

metal BiCMOS technology. The memory cell is a highly resistive polysilicon load four-transistor (4T) SRAM cell, and the cell size is 8.0 um × 4.8 um. Test circuit area overhead amounts to 17% of the total chip. Power dissipation at 75 MHz operation was estimated to be approximately 3.0W. The process technology and performance are summarized in Table II.

The measured scheme plot of cycle time versus supply voltage is shown in Fig. 11. Cycle operation of was 9 ns achieved at the supply voltage of 5 V and room temperature.

## VI. CONCLUSION

A 110 MHz /1Mb synchronous TagRAM was developed. It can be used to build a secondary cache system of up to 16Mbytes with commodity synchronous SRAM's. Data streaming cache architecture is proposed and used where integer data and floating point data are handled differently to optimize the image and graphics processing speed.

Cycle operation of 9 ns and clock-to-output delay of 4.7 ns in typical conditions were achieved by he use of circuit techniques such as a pipelined decoding scheme, a single PMOS load BiCMOS main decoder, a BiCMOS sense-amplifying comparator, a highly linear VCO for the PLL, and doubly

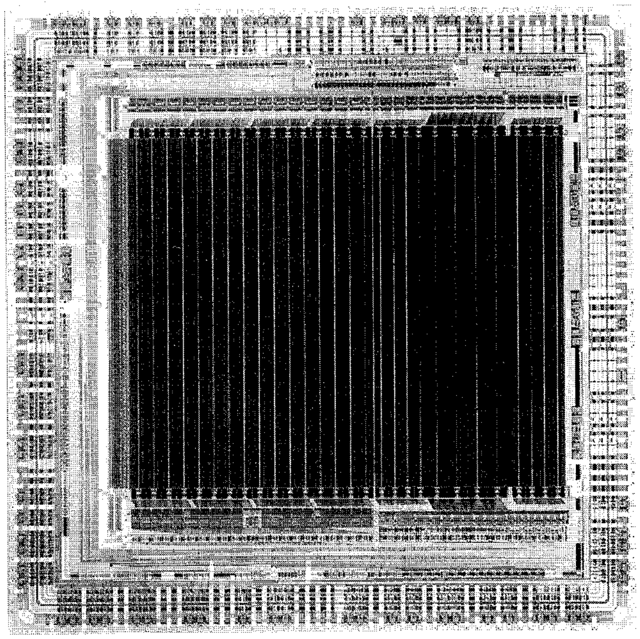


Fig. 10. Chip microphotograph. Chip size is 14.8 mm × 14.8 mm.

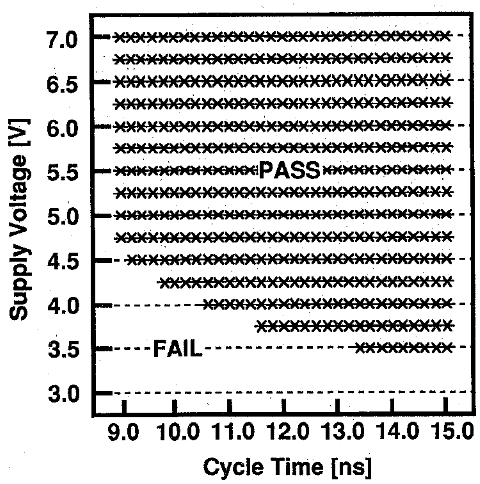



Fig. 11. Schmoo plot. The 9.0-ns cycle operation is achieved at the supply voltage of 5.0 V.

placed self-timed wire circuits. The device was successfully implemented with 0.7–$\mu$m double-polysilicon and double-metal BiCMOS technology. he circuit ideas proposed in the paper and the data streaming cache architecture are promising in the forthcoming age of 200 MHz computing systems.

## Acknowledgment

The encouragement by Y. Un'no, K. Kanzaki, and K. Maeguchi throughout the work is appreciated.

## References

[1] T. Watanabe, "An 8-kbyte intelligent cache memory," in *ISSCC Dig. Tech. Papers,* Feb. 1987, pp. 266–267.

[2] T. Sakurai *et al.,* "A circuit design of 32Kbyte integrated cache memory," in *Proc. Symp. VLSI Circ.,* Tokyo, Aug. 1988, pp. 45–46.

[3] K. Sawada *et al.,* "32Kbyte integrated cache memory," *IEEE J. Solid-State Circuits,* vol. 24, pp. 881–888, Aug. 1989.

[4] K. Uchiyama *et al.,* "Architecture and design of a second-level cache chip with copy-back and 160MB/s burst-transfer features," in *Symp. VLSI Circ.,* Honolulu, June 1990, pp. 115–116.

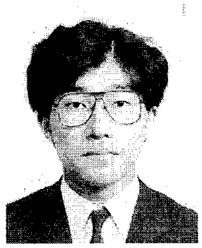
[5] K. Nogami *et al.,* "64Kbyte snoopy cache memory with flexible expandability," in *ISSCC Dig. Tech. Papers,* 1991, pp. 266–267.

[6] Yan-Tek Hsu, "Silicon graphics TFP micro supercomputer chipset," in *Hot Chips V Symp. Rec.,* Stanford, Aug. 1993, pp. 8.3.1–8.3.9.

[7] T. Sakurai *et al.,* "A low power 46ns 256Kbit CMOS static RAM with dynamic double word line," *IEEE J. Solid-State Circuits,* vol. SC-19, pp. 578–585, Oct. 1984.

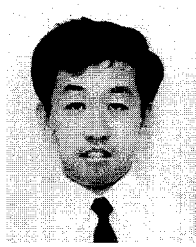
[8] G. Kitsukawa *et al.,* "Logic-in-memory VLSI," in *1983 Joint Convention Rec., Four Institutes of Engineers Japan,* 1983, pp. 5–80.

[9] H. Hara *et al.,* "0.5-um 3.3V BiCMOS standard cells with 32Kbyte cache," *IEEE J. Solid-State Circuits,* vol. 27, pp. 1579–1584, Nov. 1992.



**Yasuo Unekawa** was born in Hiroshima, Japan, on December 10, 1963. He received the B.S., M.S., and Ph.D. degrees in electronic engineering from University of Tokyo, Tokyo, Japan, in 1986, 1988, and 1991, respectively. His Ph.D. work was on thin film fabrication of high $T_c$ superconducting oxides.

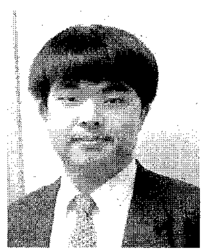
In 1991 he joined the Semiconductor Device Engineering Laboratory, Toshiba Corporation, Kawasaki, Japan, where he was engaged in the development of 1Mbit synchronous TagRAM. He is presently involved in the development of telecommunication LSI's.



**Tsuguo Kobayashi** was born in Hyogo, Japan, on November 9, 1963. He received the B.S. and M.S. degrees in electrical engineering from Keio University, Yokohama, Japan, in 1986 and 1988, respectively.

In 1988 he joined the Semiconductor Device Engineering Laboratory, Toshiba Corporation, Kawasaki, Japan, where he has been engaged in the research and development of on-chip cache macros, integrated cache memories, and high-end RISC processors.

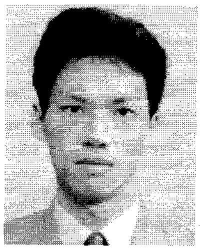
Mr. Kobayashi is a Member of the Institute of Electronics, Information, and Communication Engineers of Japan.



**Tsukasa Shirotori** was born in Nagano, Japan, on December 27, 1963. He received the B.S. degree in chemistry from Kanagawa Institute of Technology, Kanagawa, Japan, in 1986.

In 1986 he joined Toshiba Microelectronics Corporation, Kawasaki, Japan. He has been engaged in the research and development of high-end microprocessors at Semiconductor Device Engineering Laboratory, Toshiba Corporation, since 1986.

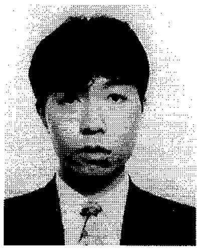
Mr. Shirotori is a Member of the Institute of Electronics, Information, and Communication Engineers of Japan.



**Yukihiro Fujimoto** was born in Shiga, Japan, on February 14, 1967. He received the B.S. degree in physics from Nagoya University, Aichi, Japan, in 1989.

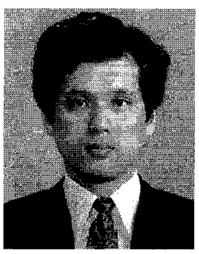
In 1989 he joined the Semiconductor Device Engineering Laboratory, Toshiba Corporation, Kawasaki, Japan, where he has been engaged in the research and development of high-end microprocessors.

Mr. Fujimoto is a member of the Institute of Electronics, Information and Communication Engineers of Japan.



**Takayoshi Shimazawa** was born in Saitama, Japan, on April 27, 1965. He received the B.S. and M.S. degrees in electric engineering from Keio University, Japan, in 1989, and 1991, respectively.

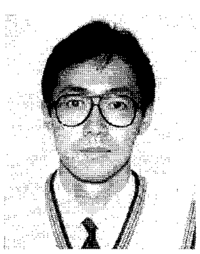
He joined the Semiconductor Device Engineering Laboratory, Toshiba Corporation, Kawasaki, Japan, in 1991. He has been engaged in the research and development of CMOS logic LSI's. At present, he is involved in the design of the video compression/decompression LSI's.



**Kazutaka Nogami** was born in Oita, Japan, on May 9, 1959. He received the B.S. and M.S. degrees in applied physics from the University of Tokyo, Tokyo, Japan, in 1982 and 1984, respectively.

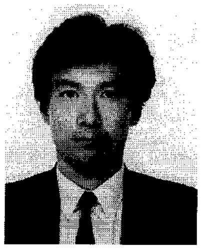
In 1984 he joined the Semiconductor Device Engineering Laboratory, Toshiba Corporation, Kawasaki, Japan, where he was engaged in the research and development of the 1-Mb VSRAM, integrated cache memories, and on-chip cache macros. He also worked on low-power circuits, drivability-controlled output circuits, and redundancy for FPGA's. He spent a year and a half on leave at the Information Systems Laboratory, Stanford University, Stanford, CA. His current interests include application-specific memories, reconfigurable LSI's, and interface circuits.

Mr. Nogami is a Member of the Institute of Electronics, Information, and Communication Engineers of Japan.



**Takehiko Nakao** was born in Tokyo, Japan, on September 19, 1964. He received the B.S. and M.S. degrees in electric engineering from Keio University, Japan, in 1988 and 1990, respectively.

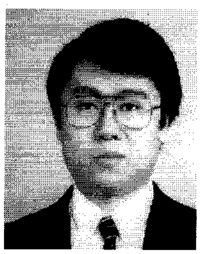
He joined the Semiconductor Device Engineering Laboratory, Toshiba Corporation, Kawasaki, Japan, in 1990. He has been engaged in the research and development of CMOS logic LSI's. At present, he is involved in the design of the LSI for communication.



**Kazuhiro Sawada** was born in Hyogo, Japan, on March 25, 1957. He received the B.S. and M.S. degrees in electrical engineering from Keio University, Yokohama, Japan, in 1980 and 1982, respectively.

In 1982 he joined the Semiconductor Device Engineering Laboratory, Toshiba Corporation, Kawasaki, Japan, where he has been engaged in the research and development of 256Kbit SRAM and many application specific memories, such as 1-Mbit virtual SRAM, 1-Mbit DRAM embedded 72K gate array, integrated cache memories, and on-chip large cache macro.

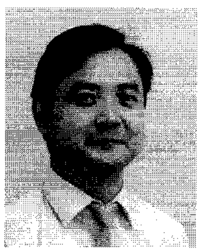
Mr. Sawada is a member of the Institute of Electronics, Information, and Communication Engineers of Japan.



**Masataka Matsui** was born in Tokyo, Japan, 1960. He received the B.S. and M.S. degrees in electronic engineering from the University of Tokyo, Tokyo, Japan, in 1983 and 1985, respectively.

In 1985 he joined the Semiconductor Device Engineering Laboratory, Toshiba Corporation, Kawasaki, Japan, where he has been engaged in the research and development of advanced logic and memory LSI's, including 1-Mbit CMOS and BiCMOS SRAM's and MPEG 1/2 video decoders. He is currently working with Prof. Peterson at the STAR Lab, Stanford University, Stanford, CA, as a Visiting Scholar, where he is studying low-power LSI design.

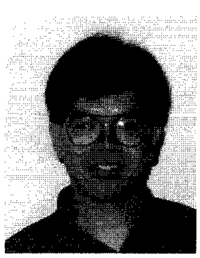
Mr. Matsui is a Member of the Institute of Electronics, Information and Communication Engineers of Japan.



**Takayasu Sakurai** was born in Tokyo, Japan, on January 10, 1954. He received the B.S., M.S., and Ph.D. degrees in electronic engineering from University of Tokyo, Tokyo, Japan, in 1976, 1978, and 1981, respectively. His Ph.D. work is on electronic structures of an Si-$SiO_2$ interface.

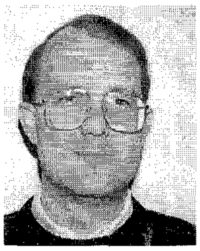
In 1981 he joined the Semiconductor Device Engineering Laboratory, Toshiba Corporation, Kawasaki, Japan, where he was engaged in the research and development of CMOS dynamic RAM and 64-Kbit, 256-Kbit SRAM, 1-Mbit virtual SRAM, cache memories, and BiCMOS ASIC's. During the development he also worked on the modeling of interconnect capacitance and delay, new memory architectures, hot-carrier resistant circuits, arbiter optimization, gate-level delay modeling, $n$th power MOS model, and transistor network synthesis. From 1988 through 1990, he was a Visiting Scholar at the University of California, Berkeley, doing research in the field of VLSI CAD. He is currently back in Toshiba and managing memory/logic VLSI development. His present interests include VLSI microprocessors, DSP's, FPGA's, and video compression/decompression LSI's. He is a Visiting Lecturer at Tokyo University and serves as a program committee member for the Symposium on VLSI Circuits, the CICC, and the ACM FPGA Workshop.

Dr. Sakurai is a Member of Institute of Electronics, Information, and Communication Engineers of Japan and the Japan Society of Applied Physics.



**Man Kit Tang** was born in Hong Kong on September 28, 1959. He received the B.S. and M.S. degrees in electrical engineering from California State University, Fresno, and Stanford University in 1984 and 1986, respectively.

In 1990 he joined Silicon Graphics Computer Systems, where he has been engaged in research and development of high-end superscalar RISC processors and symmetric multiprocessing computer systems. Prior to joining Silicon Graphics, he was engaged in research and development of symmetric multiprocessing systems and VLIW mini supercomputers for Olivetti Advanced Technology Center, Altos Computer System, and Cydrome Inc.



**William A. Huffman** was born in Wichita, KS, in 1952. He received the B.S. and M.S. degrees in physics from the Massachusetts Institute of Technology in 1974.

From 1974 to 1981 he participated in research in high energy physics and bio-physics at Harvard University, working in the area of particle physics, neutrino detectors, high vacuum, cryogenics, photon emission, and spectroscopy. From 1981 to 1983 he was with Computervision of Bedford, MA, where he wrote the transcendental function microcode and other microcode for a scientific CPU and wrote the back end of a hardware simulator. From 1983 to 1991 he was with Alliant Computer Systems, Littleton, MA, where he designed the instruction parser, the floating-point vector engines, and the transcendental microcode for the FX8 processor. Since 1991 he has been with Silicon Graphics Computer Systems, Mountain View, CA, where he is responsible for the design of the cache and multiprocessor coherence components of the TFP processor.