

# A 500 MHz Random Cycle, 1.5 ns Latency, SOI Embedded DRAM Macro Featuring a Three-Transistor Micro Sense Amplifier

John Barth, William R. Reohr, Paul Parries, Gregory Fredeman, John Golz, Stanley E. Schuster, *Life Fellow, IEEE*, Richard E. Matick, *Life Fellow, IEEE*, Hillery Hunter, Charles C. Tanner, III, Joseph Harig, Hoki Kim, Babar A. Khan, John Griesemer, Robert P. Havreluk, Kenji Yanagisawa, Toshiaki Kiriata, *Senior Member, IEEE*, and Subramanian S. Iyer, *Fellow, IEEE*

**Abstract**—As microprocessors enter the highly multi-core/multi-threaded era, higher density, lower latency embedded memory will be required to meet cache design needs. This paper describes a 500MHz random cycle Silicon on Insulator (SOI) embedded DRAM macro which features a three-transistor micro sense amplifier, realizing significant performance gains over traditional array design methods. To address the realities of process integration, we describe the features and issues associated with integrating this DRAM into SOI technology, including deep trench processing and floating body effects. After a brief description of the macro architecture, details are provided on the three-transistor micro sense amplifier scheme, which is key to achieving a high transfer ratio with minimal area overhead. The paper concludes with hardware results and a summary.

**Index Terms**—DRAM chips, FET amplifiers, memory architecture, microprocessor chips, Silicon on Insulator.

## I. MOTIVATION

SYSTEM-LEVEL processor simulations show that doubling cache capacity results in respectable double-digit percentage gains for cache-constrained commercial applications [1]. Improving cache latency also has an impact on system performance, and integrating the cache on-chip eliminates delay, power and area penalties associated with high frequency off-chip I/O channels. Moving forward, trends in virtual machine technology, multi-threading and multi-core processors will continue to stress cache sub-systems [2], [3]. The availability of a high-performance embedded DRAM macro, offering a 3x density advantage over SRAM, would allow cache-dominated chips to double their cache capacity, while actually fitting in a smaller footprint. Smaller cache size not only reduces chip cost, but offers latency reduction from shorter wire run lengths to retrieve data. For power-constrained applications, eDRAM consumes one-fifth the keep-alive power of SRAM, while its high capacitance and small collector area

offer a soft error rate three orders of magnitude lower than SRAM.

## II. CURRENT LEVEL OF INTEGRATION

Logic-based embedded DRAM has matured into a wide range of applications (Fig. 1), most notably the on-chip cache for IBM's BlueGene/L supercomputer [4], currently ranked #1 on the top 500 supercomputer list [13]. Integration of embedded DRAM for caching enabled BlueGene/L to achieve world class performance, while utilizing only a modestly performing bulk processor technology. IBM's dual-core Power5 microprocessor [5], in contrast, leverages SOI technology to achieve multi-gigahertz CPU frequencies, and exploits bulk technology embedded DRAM to provide a massive single-chip 36 MByte external L3 cache. Imagine the performance attainable if embedded DRAM were available on-chip with the high-performance SOI microprocessor. In order to achieve this next level of integration, two requirements must be met: first, embedded DRAM technology must be integrated on the same high-performance SOI technology platform currently utilized by microprocessors; second, DRAM performance must be enhanced to service the aggressive latency requirements of multi-gigahertz processors.

## III. TECHNOLOGY FEATURES

To satisfy the first requirement, DRAM has been embedded on a 65 nm, partially depleted SOI, high-performance technology platform [Fig. 2(a)] [6]. The technology offers both 1.12 nm thin oxide and 2.35 nm thick oxide devices. The DRAM cell is fabricated on SOI and utilizes the technology's standard 2.35 nm thick oxide floating body NFET. A silicide strap forms a low resistance connection to a 20 fF deep trench capacitor. High trench capacitance is achieved by digging through the SOI and buried oxide (BOX), 4.5  $\mu\text{m}$  into the substrate. Conveniently, the buried oxide provides isolation between the SOI and substrate, eliminating the oxide collar required by bulk technologies to protect against vertical parasitic leakage. The details of this mechanism are described in the following section.

## IV. COLLAR PROCESS ELIMINATION

Fig. 2(b) shows the bulk and SOI embedded DRAM process cross sections. Bulk trench technologies must deal with parasitic leakage created by the vertical device formed between the node and N-Band, gated by the trench sidewall. To mitigate this

Manuscript received April 4, 2007; revised July 10, 2007.

J. Barth, C. C. Tanner, and J. Harig are with IBM Burlington, Essex Junction, VT 05452 USA (e-mail: jbarth@us.ibm.com).

W. R. Reohr, S. E. Schuster, R. E. Matick, and H. Hunter are with IBM T. J. Watson Research Center, Yorktown Heights, NY 10598 USA (e-mail: schustr@us.ibm.com).

P. Parries, G. Fredeman, J. Golz, H. Kim, B. A. Khan, J. Griesemer, R. P. Havreluk, K. Yanagisawa, T. Kiriata, and S. S. Iyer are with IBM, Hopewell Junction, NY 12533 USA.

Digital Object Identifier 10.1109/JSSC.2007.908006

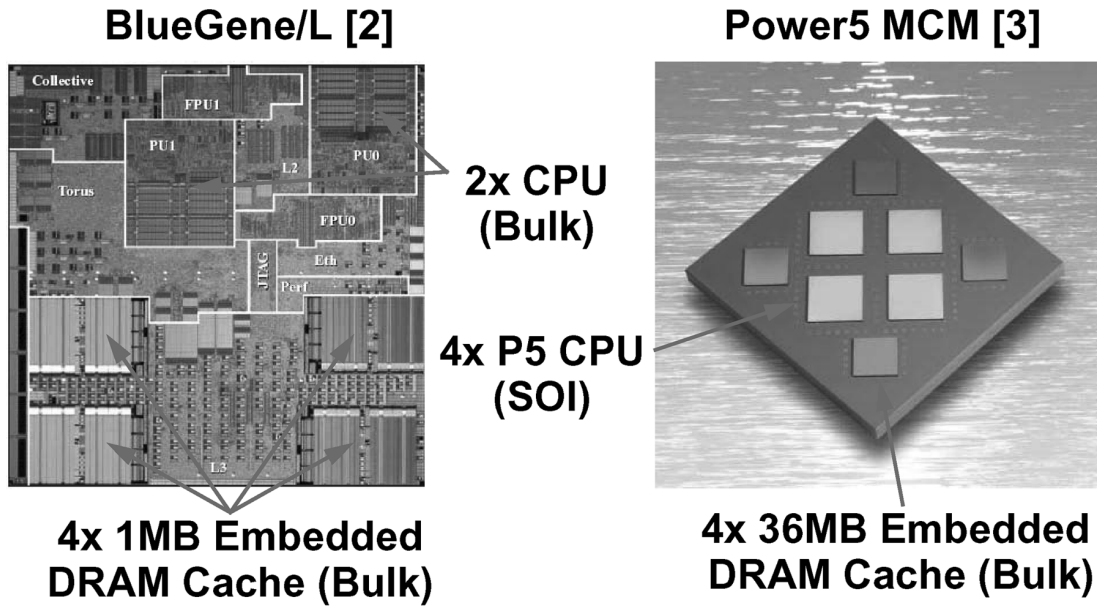


Fig. 1. BlueGene/L and Power5 MCM.

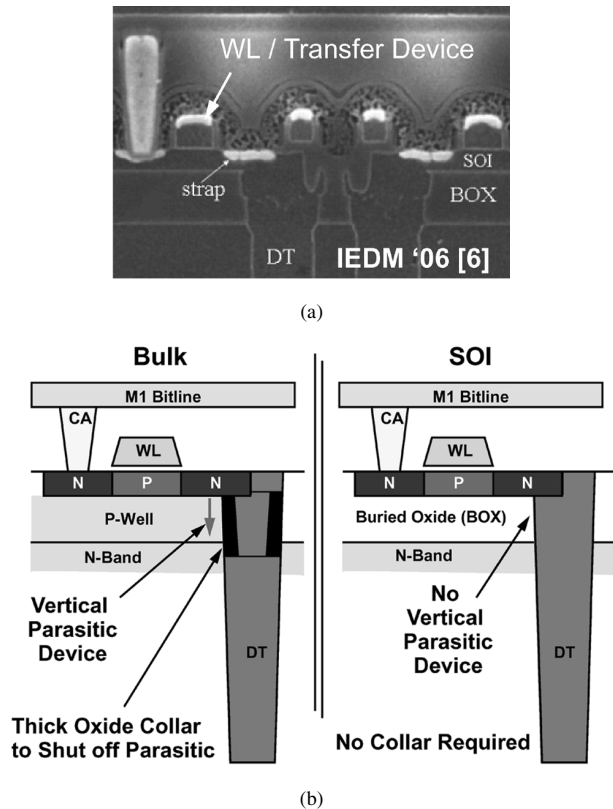


Fig. 2. SOI eDRAM process cross sections.

leakage, bulk technologies must form a thick oxide collar at the top of the trench. In SOI, the bulk P-Well is replaced by the buried oxide, isolating the node from the N-Band, and precluding the need for a collar. Eliminating the collar not only

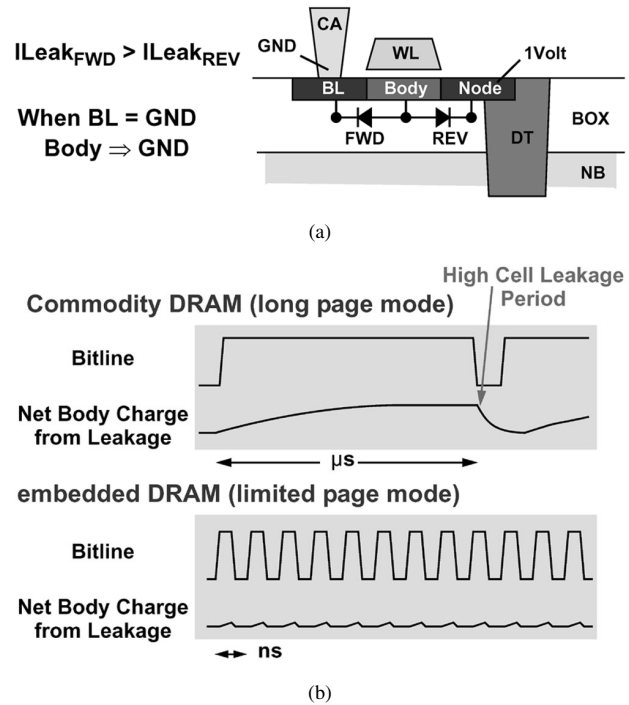


Fig. 3. SOI floating body effects.

reduces process complexity, but also increases the cross-sectional area of the trench, further reducing parasitic resistance to the storage capacitor. These factors combine to drop the overall eDRAM process cost adder from 15% for bulk to 7% for SOI. In addition, the absence of wells in SOI realizes lower source-drain capacitances and enables layout area reduction by eliminating well tie-downs and allowing smaller NFET-to-PFET spacing.

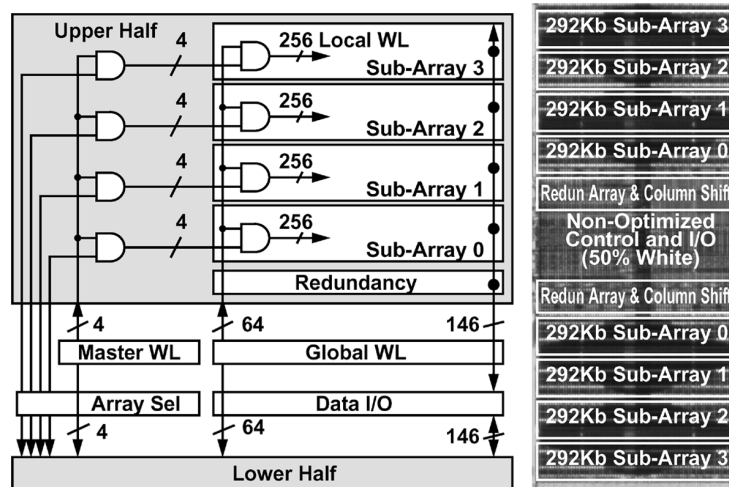


Fig. 4. Macro architecture.

## V. FLOATING BODY EFFECTS

Integrating a DRAM cell into SOI technology does, however, introduce a concern: without a well, the body of the eDRAM cell's SOI device is floating. The potential of the body is determined by a number of mechanisms. Primarily, these mechanisms are: coupling, junction leakage, impact ionization, thermal generation, and recombination. Some of these can be exploited; others must be carefully managed. During DRAM cell write-back, the array's wordline and bitline are both driven high, coupling the body of the transfer device high, acting to lower its threshold, and enabling a higher voltage to be written into the cell. When write-back is complete, the wordline is lowered, trapping charge on the storage capacitor. At this time, leakage mechanisms take over, specifically junction leakage at the forward and reverse bias diodes formed between the body and source drain regions [Fig. 3(a)]. If the device's body floats high, its  $V_t$  will be reduced, degrading the cell's retention characteristics due to subthreshold leakage. A high body potential also increases vulnerability to bipolar current when the source is quickly discharged.

For these reasons, it is desirable to prevent a high body situation on the cell access device. Recognizing that a forward-biased junction has higher leakage than a reverse-biased junction, holding the bitline low will cause the body to float closer to ground, thus reducing subthreshold leakage. To realize this scenario, bitline ground pre-charge is implemented to maintain a low body potential. Additionally, periods of bitline high are limited by the short page depth associated with this embedded DRAM's wide-I/O design.

## VI. ARRAY FLOATING BODY

Fig. 3(b) shows two graphs of net body charge, with coupling effects removed. The top graph shows the weak reverse bias leakage slowly charging the floating body high on the eDRAM cell's NFET. This would occur for a design like that of commodity DRAM, where the bitline is kept high for a long period of time, to keep a "page" of data open. When the bitline is lowered, high subthreshold leakage and bipolar current then rapidly

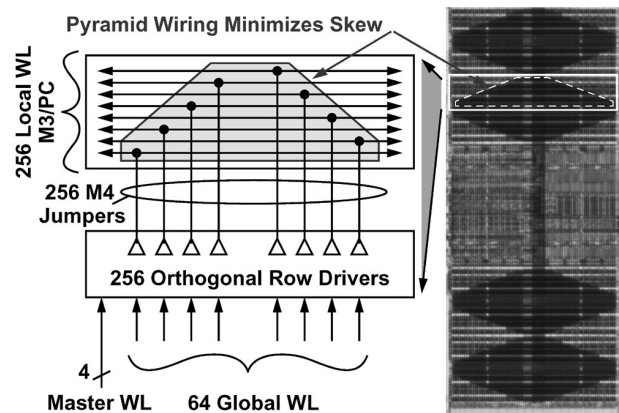


Fig. 5. Pyramid wiring scheme.

drain charge stored in the DRAM cell. The stronger forward bias leakage can also be seen discharging the body. The bottom graph in Fig. 3(b) shows frequent bitline grounding, as used in this paper's design; this maintains the cell device's body charge desirably near ground.

## VII. MACRO ARCHITECTURE

Architecturally the macro is divided into upper and lower halves, each half consisting of four 292 Kb sub-arrays (Fig. 4). In the central region, addresses are decoded into array selects, master wordlines and global wordlines, which are delivered over the sub-arrays on Metal-4. Locally, master wordlines are enabled by a single array select and combined with the global wordlines to select one of 256 local wordlines. One-hot late-select column select signals support 16-way cache associativity. Eight are driven to the upper array, eight to the lower array, selecting one of 16 columns. The central region also includes control logic and the 146 bit data I/O. It should be noted that this region reused an existing SRAM I/O design and is not area optimized.

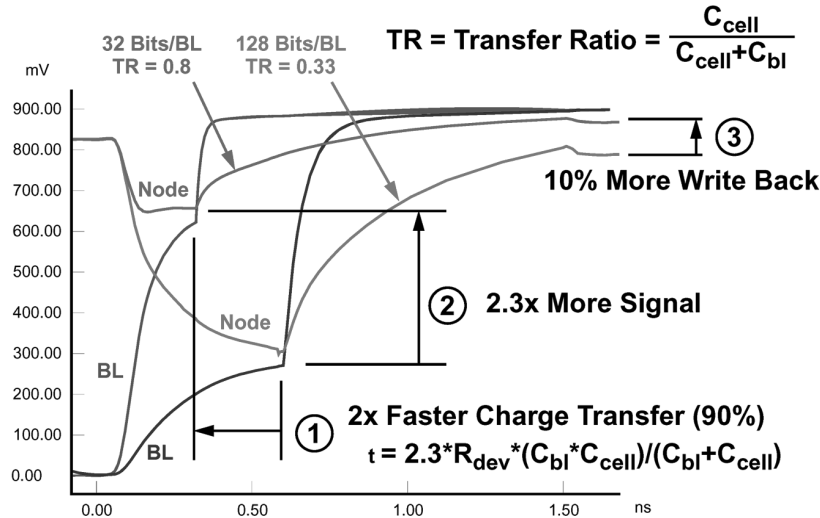


Fig. 6. High transfer ratio benefits.

VIII. PYRAMID ROW

Fig. 5 shows the unconventional orthogonal location of the wordline driver as introduced by Wiess *et al.*, at ISSCC 2002 [7]. Metal-4 is used in a pyramid-like pattern to jumper the orthogonal wordline driver to the east/west Metal-3 wordline strap. The pyramid pattern, easily distinguishable in the micrograph on the right side of the figure, minimizes wordline skew by using the shortest Metal-4 jumper on the longest Metal-3 runs. The orthogonal wordline configuration saves area by decoupling the wordline driver layout from the array wordline pitch and avoids replicating wordline pre-decode in each sub-array.

IX. SHORT BITLINE

To satisfy the second requirement for on-chip embedded DRAM caches, macro performance must be improved to meet the latency requirements of a multi-gigahertz microprocessor. Where area overhead is not a concern, reducing bitline length is an accepted method for improving DRAM performance. Fig. 6 compares two different eDRAM array designs: one with 128 cells per bitline, and a second with 32 cells per bitline. This figure highlights three performance gains: first, the lower bitline capacitance of the 32 cell bitline results in a reduced read time-constant and faster signal development; second, the high transfer ratio of the short bitline design creates a large bitline voltage swings, providing more signal; and third, reduced cell charge transfer during a read minimizes charging time to replenish the cell during write-back.

X. OVERHEAD

Short bitline performance is attractive, but must be weighed against increased area overhead. Each time a bitline is cut in half, sense circuits and data buffers must be doubled. The high-performance direct write embedded DRAM presented at ISSCC 2005 [8] requires 11 transistors for each sense amplifier, plus reference cells and bitline twisting, totaling a 14% area overhead. A 128-bits-per-bitline version would suffer 27% overhead; extrapolating to 32 cells per bitline would result in an

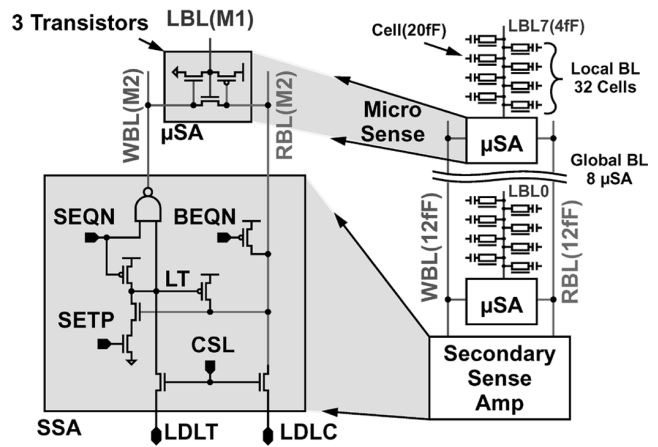


Fig. 7. Hierarchical micro sense amp.

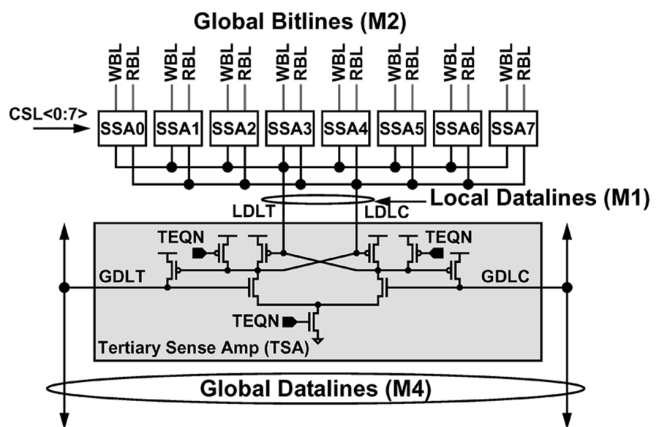


Fig. 8. Tertiary sense amp.

unacceptable 80% overhead. To make matters worse, implementing the sense amplifier design of [8] in SOI would require further overhead for body-tied SOI devices, required to prevent history-induced sense amplifier mismatch. To extract overall

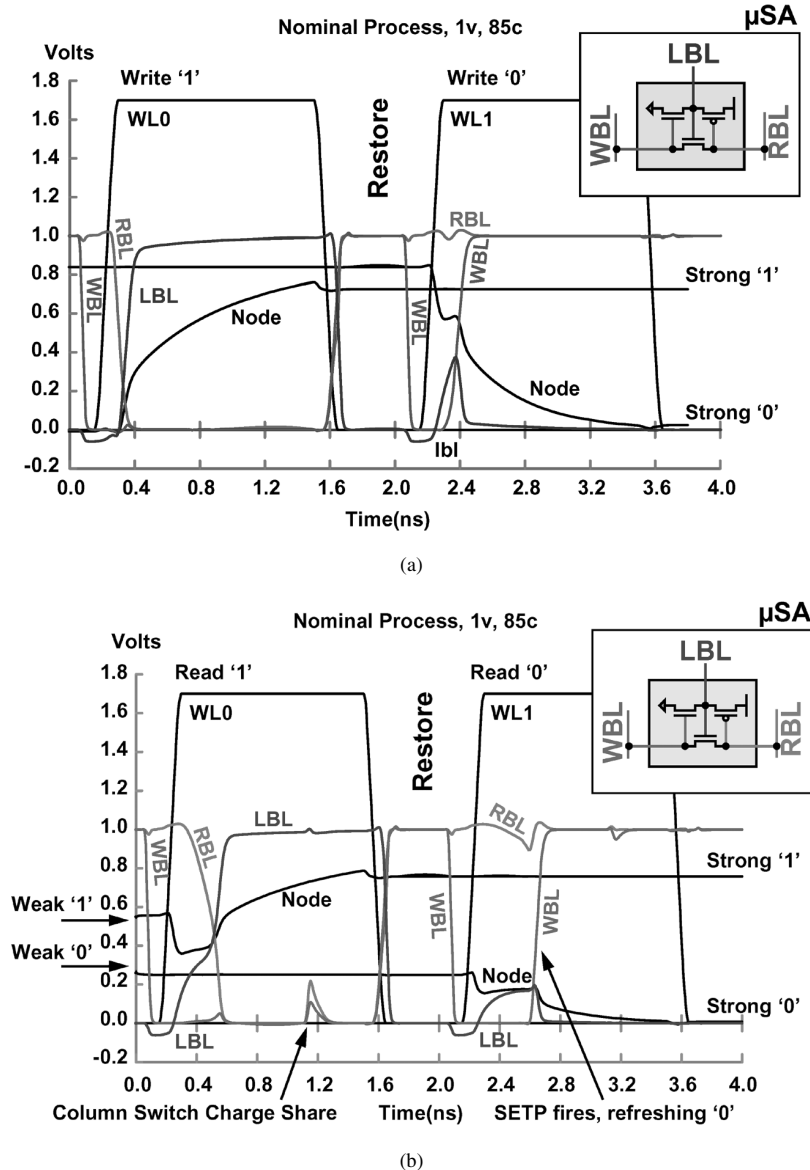


Fig. 9. Micro sense simulations.

benefit from a high cell-to-bitline transfer ratio, a low overhead sensing scheme is required.

#### XI. MICRO SENSE AMPLIFIER ( $\mu$ SA)

Fig. 7 proposes a three-transistor scheme to minimize sense amplifier overhead. The hierarchical scheme relies on the high transfer ratio during a read to create a large voltage swing on a local bitline; large enough, in fact, to be sampled with a single-ended amplifier. The amplifier has been minimized by decomposing its three basic operations—read, write '0', and write '1'—each being implemented with a single transistor. Each micro sense amplifier ( $\mu$ SA) services a short 32 cell local bitline, labeled LBL in Fig. 7. The single-ended nature of this sensing scheme allows a relaxed Metal-1 pitch, reducing the

line-to-line capacitance by a factor of three. The short bitline, relaxed pitch, and low SOI junction capacitance combine for a local bitline capacitance of only 4 fF. When coupled with a 20 fF cell, an 83% transfer ratio is realized, ideally producing over 800 mV of signal—more than enough to turn on a 250 mV  $V_t$  transistor.

Each three-transistor  $\mu$ SA is configured with an NFET for reading, a PFET for writing a full one, and an NFET for writing a full zero. The  $\mu$ SA transfers data to/from a secondary sense amplifier (SSA) via two global read/write bitlines, labeled RBL and WBL. The Metal-2 global bitlines, routed in parallel to the Metal-1 local bitlines, control the read/write operations to all  $\mu$ SAs in parallel. Each SSA services eight  $\mu$ SAs in a hierarchical fashion.

TABLE I

Operation	WBL	RBL	LBL
Pre-Charge	1	1	0
Write '1'	0	0	1
Write '0'	1	1	0
Read '1'	0	1Z	1Z
Refresh '1'	0	0	1
Read '0'	0	1Z	0Z
Refresh '0'	1	1	0

Z = Hi Impedance

SETP

## XII. TERTIARY SENSE AMPLIFIER

The  $\mu$ SA adds an extra level of hierarchy and necessitates a third-level (tertiary) sense amplifier. The bi-directional tertiary sense amplifier shown in Fig. 8 is responsible for transferring data between the Metal-4 global datalines and the selected SSA. One of eight secondary sense amplifiers is selected by the column address, which controls the one-hot column select lines (CSL). In a set-associative cache, the column address could be used as a late way select.

During a write, either the true or complement global dataline is remotely driven high. In turn, the tertiary sense amplifier pulls either the true or complement local dataline low, which is then passed to the selected SSA. During a read, the global data lines are left in the low pre-charge state until the selected SSA pulls either of the local datalines low. The data passes through the tertiary sense amplifier and pulls the corresponding global dataline high. It should be noted that a read or write to the sub-array is determined by the timing of the global datalines and does not require dedicated Read/Write control signals.

## XIII. OPERATION TRUTH TABLE

Table I shows the logical operation of the  $\mu$ SA. All cycles start in the pre-charge condition with read and write global bitlines held high, pre-charging the local bitline to ground. Sub-array selection triggers the sense amplifier equalize signal SEQN to go high (Fig. 7), forcing the write bitline low. This enables the  $\mu$ SA read transistor, and tri-states the local bitline, indicated by a 'Z' in the truth table. A write '1' is achieved by driving the read bitline low—as indicated in Table I—while a write '0' is achieved by driving the write bitline high. Logically, the local bitline remains low in the read '0' state and only transitions high in a read '1' state.

## XIV. WRITE WAVEFORMS

Fig. 9(a) shows a simulation of write operation using the  $\mu$ SA. To write a '1', the selected SSA pulls the read bitline low, forcing the local bitline high, writing a '1' into the selected cell. To write a '0', the selected SSA drives both the read bitline and write bitline high, in turn driving the local bitline low and writing a '0' into the selected cell. It should be noted that due to the  $\mu$ SA's unique configuration, writing opposite data onto

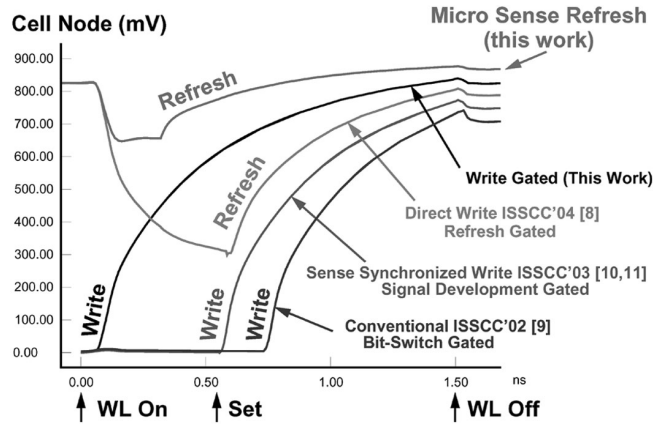


Fig. 10. Cell performance innovations.

the local bitline is accomplished without contention, eliminating any technology P-to-N ratio dependence.

## XV. READ WAVEFORMS

The process of reading using the  $\mu$ SA is also relatively straightforward, but contains one unexpected component. As shown in Fig. 9(b), the  $\mu$ SA is first taken out of pre-charge by lowering the write bitline. The read bitline, however, is held in the pre-charge state until an interlock circuit detects the write bitline is low. This extended pre-charge absorbs any line-to-line coupling from the write bitline to the read bitline, preventing any premature amplification. When a wordline is activated, read data is transferred from the cell to the local bitline. For a stored '1', the local bitline rises at least 1 threshold above the read transistor, weakly pulling the read bitline low. Here is the novel component of this operation: when the read bitline falls below the threshold of the  $\mu$ SA PMOS, feedback drives the local bitline to a full high level. This amplifies the local bitline, refreshes the cell, and strongly drives the read bitline low, which can be seen as a slope change on both the local bitline and the read bitline. It should be noted that the refresh of a '1' is self-timed, and requires no external timing control.

For a stored '0', the local bitline and cell are very close in potential, therefore the local bitline remains low, leaving the read transistor off and the read bitline high. The read bitline remains high until the external timing signal SETP triggers the SSA to evaluate the read bitline. With the read bitline still in the high state, the SSA drives the write bitline high, forcing the local bitline low, refreshing the '0'.

## XVI. CYCLE LIMITS

Historically, the cycle time of NFET DRAM arrays are determined by writing a high level into the cell. Fig. 10 shows the evolution of embedded DRAM architectural advances aimed at reducing cycle time. These advances include the sense-synchronized write [10], [11] and direct write [8] described at ISSCC 2003 and 2004, respectively. Similar to prior work, this paper's tertiary, secondary, and micro sense amplifiers are configured to be bi-directional and support direct write of data prior to wordline activation, without disrupting reads of adjacent sense ampli-

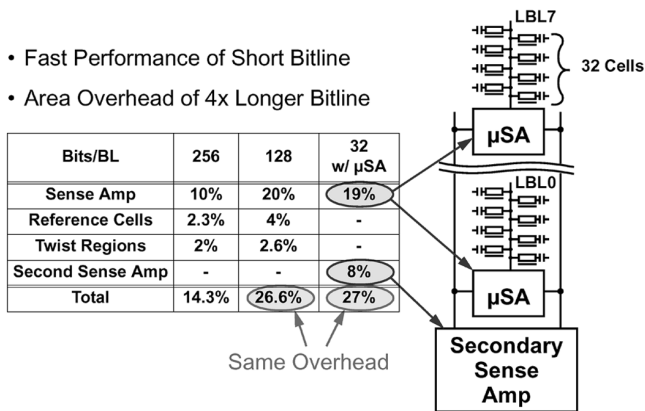


Fig. 11. Micro sense overhead.

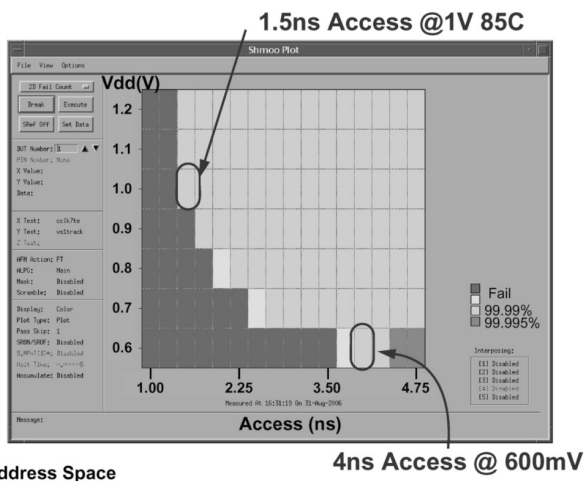


Fig. 12. Access time results.

fiers and data bits. By significantly improving the refresh cycle time, the micro sense architecture reduces cycle time down to the cell's fundamental ability to write opposite data.

### XVII. MICRO SENSE AMPLIFIER OVERHEAD

As mentioned previously, the performance of short bitline designs must be weighed against their increased area overhead. To quantify this tradeoff, Fig. 11 compares the area overheads of three different eDRAM array designs: 1) a direct write array with 256 cells/bitline [8]; 2) the same design style, but with 128 cells/bitline; and 3) the  $\mu$ SA-based array described in this paper, with 32 cells/bitline. As seen in the total area overhead row, the three-transistor  $\mu$ SA design achieves its breakthrough performance with the same 27% area overhead associated with a bitline four times its length.

### XVIII. LATENCY AND CYCLE TIME

The 2 Mb macro shmoo plot in Fig. 12 demonstrates 1.5 ns access time with 99.99% bit yield across the full address space

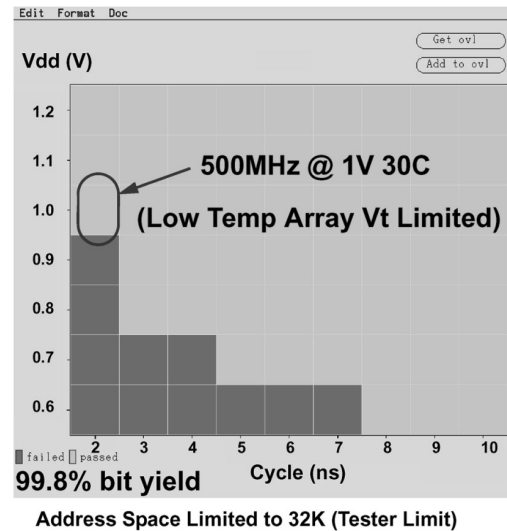


Fig. 13. Cycle time results.

at 1 V and 85 °C. The design achieves a respectable 4 ns access at 600 mV. As mentioned in description of the sense amplifier's operation in write mode, in contrast to SRAM cells and cross-coupled DRAM sense amplifiers, there is no reliance on nMOS/pMOS strength ratios for writing to the opposite state. This enables very low 600 mV operation, limited only by the highest threshold read transistor. All-good repairable 13.5 Mb chips have been successfully manufactured to a relaxed product-like specification.

The shmoo in Fig. 13 demonstrates 500 MHz operation at 1 V. Macro performance exceeds the cycle limit of the full address space tester used for this plot; for this reason, 500 MHz verification was limited to the 32 k address space of the high-speed tester.

### XIX. FEATURES

Fig. 14 shows a photomicrograph of the 13.5 Mb test chip, containing two 2 Mb macros and two 4 Mb embedded DRAM macros. Fig. 14 also shows a features table, including the array's 40  $\mu$ s retention specification. Using concurrent refresh (described by Kiriata, *et al.* at ISSCC 2004 [9]), array availability is 98%. As listed in the features table, typical keep-alive power (static power plus the power cost of periodic refresh) is 45 mW—one-fifth the power of SRAM in the same technology.

### XX. SUMMARY

This paper has described a prototype embedded DRAM macro developed in a 65 nm SOI process. SOI enables DRAM process simplification, resulting in a low 7% cost adder. The paper proposed a hierarchical three-transistor micro sense amplifier, offering significant performance improvements compared to cross-coupled amplifiers with similar area overhead. By maintaining a 3x density advantage over SRAM, we feel this would be an ideal on-chip cache

Technology	65nm PD-SOI (Tox=1.12nm/2.35nm)
Cell Size	0.23 $\mu$ m $\times$ 0.55 $\mu$ m (0.1265 $\mu$ m <sup>2</sup> )
Retention	40us @ 105c, 98.7% Avail w/ CR [12]
Macro Size (Calculated)	568 $\mu$ m $\times$ 1170 $\mu$ m 0.665mm <sup>2</sup> (0.317x SRAM)
Organization	1K Row x 16 Col x 146 (2Mb)
Sub-Array	Hierarchical w/ Micro Sense 32 Rows/Local BL 8 Local BL/Global BL 1200 Bits/WL
Redundancy	2 - 32 WL Row Arrays Static Column Shifting
Supply	1.0V / 1.7V (WL High)
Performance	2ns Random Cycle 1.5ns Random Access
Power(Nom)	AC only - 76mW (0.84x SRAM) Stdbby + Refresh - 42mW (0.18x SRAM)

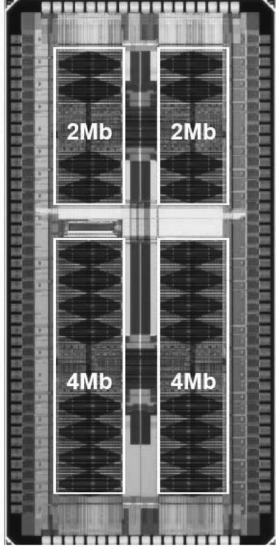


Fig. 14. Technology summary and die photo.

memory for high-performance microprocessor and gaming applications. We are confident this technology is scalable as we progress with our 45 nm SOI embedded DRAM product development.

#### REFERENCES

- [1] R. Matick *et al.*, "Logic-based eDRAM: Origins and rationale for use," *IBM J. Res. Devel.*, vol. 49, no. 1, pp. 145–65, Jan. 2005.
- [2] E. Cohen *et al.*, "A 64B CPU pair: dual and single-processor chips," in *IEEE ISSCC Dig. Tech. Papers*, 2006, pp. 333–342.
- [3] S. Naffziger *et al.*, "The implementation of a 2-core multi-threaded Itanium-family processor," in *IEEE ISSCC Dig. Tech. Papers*, 2005, pp. 182–183.
- [4] S. Iyer *et al.*, "Embedded DRAM: Technology platform for Blue Gene/L Chip," *IBM J. Res. Devel.*, vol. 49, no. 2/3, pp. 333–50, Mar./May 2005.
- [5] J. Clabes *et al.*, "Design and implementation of the POWER5 microprocessor," in *IEEE ISSCC Dig. Tech. Papers*, 2004, pp. 56–57.
- [6] G. Wang *et al.*, "A 0.127  $\mu$ m<sup>2</sup> high-performance 65 nm SOI based embedded DRAM for on-processor applications," in *IEDM Tech. Dig.*, 2006, pp. 1–4.
- [7] D. Weiss *et al.*, "The on-chip 3 MB subarray based 3rd level cache on an Itanium microprocessor," in *IEEE ISSCC Dig. Tech. Papers*, 2002, pp. 112–113, 451.
- [8] J. Barth *et al.*, "A 500MHz multi-banked compilable DRAM macro with direct write and programmable pipelining," in *IEEE ISSCC Dig. Tech. Papers*, 2004, pp. 204–205, 523.
- [9] J. Barth *et al.*, "A 300 MHz multi-banked eDRAM macro featuring GND sense, bit-line twisting and direct reference cell write," in *IEEE ISSCC Dig. Tech. Papers*, 2002, pp. 156–157.
- [10] H. Pilo *et al.*, "A 5.6-ns random cycle 144-Mb DRAM with 1.4 Gb/s/pin and DDR3-SRAM interface," *IEEE J. Solid-State Circuits*, vol. 38, no. 11, pp. 1974–1980, Nov. 2003.
- [11] Y. Taito *et al.*, "A high density memory for SoC with a 143MHz SRAM interface using sense-synchronized-read/write," in *IEEE ISSCC Dig. Tech. Papers*, 2003, pp. 306–307.
- [12] T. Kirihaata *et al.*, "An 800MHz embedded DRAM with a concurrent refresh mode," in *IEEE ISSCC Dig. Tech. Papers*, 2004, pp. 206–207, 523.
- [13] Top 500 Supercomputer Sites TOP 500 List. Nov. 2006 [Online]. Available: <http://www.top500.org/list/2006/11/100>



**John Barth** received the B.S.E.E. degree from Northeastern University, Boston, MA, in 1987, and the M.S.E.E. degree from National Technological University (NTU), Fort Collins, CO, in 1992.

He works on embedded DRAM, macro architecture and core design for IBM Systems and Technology Group, Burlington, VT. Currently, he is developing SOI embedded DRAM macros for high-performance microprocessor cache applications.

During his BS degree work, he was employed at the Timeplex Development Laboratory, Rochelle Park, NJ, from 1984 to 1985, where he wrote data communications and network monitoring software. He also co-oped at the IBM Development Laboratory, Essex Junction, VT, in 1986, where he was involved in the design and characterization of the 1 Mb DRAM product. After completing the BS degree, he returned to the IBM Development Laboratory, Essex Junction, VT, during which time he was involved in the design of a 16 Mb DRAM product featuring embedded ECC and SRAM cache. Following this, he was involved in the array design for the 16/18 Mb DRAM products and in 1994, started work on wide I/O, high-performance, DRAM Macros for embedded general purpose ASIC applications. Utilization of these macros expanded into network switching, stand-alone caches for P5 and P6 microprocessors and embedded caches for the Blue Gene/L super computer. Currently, he holds 25 U.S. patents with eight pending and has co-authored nine IEEE papers.

Mr. Barth received the IEEE JOURNAL OF SOLID-STATE CIRCUITS 1989–1990 Best Paper Award. In 2002, he was co-recipient of the ISSCC Beatrice Award for Editorial Excellence for a 144 Mb DRAM targeted for stand-alone SRAM cache replacement. From 2002 through 2007, he was a member of the ISSCC Memory Subcommittee responsible for conference organization and paper selection.



**William R. Reohr** received the Bachelor of Science degree in electrical engineering from the University of Virginia, Charlottesville, in 1987, and the Master of Science degree in electrical engineering from Columbia University, New York, NY, in 1990.

He joined IBM in 1987 and presently holds the position of Research Staff Member at the T. J. Watson Research Center, Yorktown Heights, NY. He has contributed to the design of high-performance processors, primarily their cache memories, for IBM's S/390 servers, Sony's PlayStation 3, and Microsoft's

XBOX 360. Additionally, as a principle investigator on a DARPA contract, he was responsible for the early development of a novel memory known as MTJ MRAM, which has been commercialized by Freescale, Inc., and Honeywell, Inc. Most recently, he has participated in the development of an embedded DRAM macro for Silicon-On-Insulator (SOI) technology. He holds over 25 patents covering various areas of VLSI logic, circuits, and technology.



Mr. Reohr received the Jack Raper Award for the Outstanding Technology Directions Paper at ISSCC 2000 and a Best Paper Award at the 1993 IEEE VLSI Test Symposium.



**Paul Parries** received the B.S. and Ph.D. degrees in electrical engineering from the University of Minnesota.

He joined IBM, Burlington, VT, in 1985. He is currently a Senior Technical Staff Member working in embedded DRAM technology and product development in IBM, East Fishkill, NY. He has had nearly 20 years experience working on commodity DRAM and embedded DRAM technologies and has expertise in process development, DRAM process integration, circuit design and characterization.



**Gregory Fredeman** received the B.S. degree in electrical engineering technology from SUNY Utica/Rome, NY, in 1996. He received the M.S. degree in electrical engineering from Walden University in 2005.

He joined IBM Microelectronics in Vermont in 1996, where he worked on test development and design verification for stand-alone synchronous DRAM. He transferred to IBM Microelectronics, East Fishkill, NY, in 2000, where he worked on the design of high-performance embedded DRAM and

EFUSE products. He is currently working on SOI embedded DRAM products.



**John Golz** received the B.A. degree in physics from the University of California at Berkeley in 1989, and the M.S. degree in electrical engineering from Yale University, New Haven, CT, in 1992.

He joined IBM Microelectronics in 1996, where he has worked on development of stand-alone and embedded DRAM products. His current activities include the optimization of memory array cores for performance, cost, and yield, for single-chip integration of DRAM technology with high-performance logic.



**Stanley E. Schuster** (M'65-SM'89-F'90-LF'06) received the B.S. and M.S. degrees in electrical engineering from New York University, New York, NY, in 1962 and 1969, respectively.

In 1965, he joined the IBM Research Division, Thomas J. Watson Research Center, Yorktown Heights, NY, where he currently is a Research Staff Member working in the VLSI Design department in the area of low-power high-speed digital circuits and logic-based eDRAM for memory hierarchies. His work on high-performance logic-based eDRAM

include pioneering effort and analysis to demonstrate the value and feasibility of logic-based embedded DRAM for system products.

Initially upon joining IBM, he worked on nMOS device characterization and circuit design for logic and memory. This work was part of the effort that led to IBM's nMOS technology for main memory. He was also involved in the application of semiconductor technology to communication systems. During the early 1970s, he did extensive work that demonstrated the leverage of word-line and bit-line redundancy for semiconductor memory yield. His work on redundancy included the development of laser personalization techniques for integrated circuits. This technique was also used to achieve fast turnaround times in the personalization of PLAs. In the late 1970s, he started a research effort on the design and application of a series of nMOS and CMOS memory chips for very high-speed operation. This work included the design of very high-speed CMOS ECL-compatible SRAMs with cycle time less than access time. A paper describing this work received the 1992 IEEE International Solid-State Circuits Conference Lewis Winner Award for Outstanding Paper. This work was instru-

mental in demonstrating that CMOS could move into the high-speed arena that was previously the domain of bipolar technology.

Mr. Schuster was an associate editor for the IEEE JOURNAL OF SOLID-STATE CIRCUITS from 1988 to 1992, and was guest editor for the Special Issue on Logic and Memory in October 1986. He served on the ISSCC program committee from 1985 to 1988. He was co-chairman of the IEEE Solid-State Circuits and Technology Workshop Committee and a member of the AdCom of the IEEE Solid-State Circuits Society from 1994 to 2003. He has 51 issued patents and over 56 published papers. He is a Fellow of the IEEE and a member of the IBM Academy of Technology. He has received eight IBM Outstanding Invention, Technical Achievement and Contribution Awards and an Eighteenth Level IBM Invention Achievement Award. He is also the recipient of an IBM Corporate Award for his work on high-speed CMOS SRAMs.



**Richard E. Matick** (M'65-SM'86-F'90-LF'00) received the B.Sc., M.Sc., and Ph.D. degrees in electrical engineering from Carnegie Mellon University, Pittsburgh, PA, in 1955, 1956, and 1958, respectively. After completing his Ph.D., he joined the IBM Research Division and worked in the areas of thin magnetic films, memories, and ferroelectrics.

He is currently a Research Staff Member in Systems Technology and Microarchitecture at IBM Research Division, Thomas J. Watson Research Center, Yorktown Heights, NY. As Manager of the Magnetic

Film Memory group from 1962 to 1964, he received an IBM Outstanding Invention Award for the invention and development of the thick-film read-only memory. He joined the Technical Staff of the IBM Director of Research in 1965 and remained until 1972, serving in various staff positions and as Technical Assistant to the Director of Research. In 1972, he took a one-year sabbatical to teach at the University of Colorado and IBM in Boulder, CO. He spent the summer of 1973 teaching and doing research at Stanford University. In 1986, he received an IBM Outstanding Innovation Award and in 1999 an IBM Corporate Patent Portfolio Award as co-inventor of the industry-standard video RAM memory chip, used to provide the high-speed, high-resolution display bit buffer in personal computers and workstations. His work in high-density CMOS cache memory design, for which he received an IBM Outstanding Technical Achievement Award in 1990, served as the foundation for the high-speed cache system used in the IBM RISC/6000 series processors. He is co-initiator of the concept of logic-based embedded DRAM, which has become a key IBM strategy for systems-on-a-chip. It was initially conceptualized in 1990 and became a reality in the late 1990s. He is the author of the books *Transmission Lines for Digital and Communication Networks*, McGraw-Hill, 1969 (reprinted as an IEEE Press Classic Reissue in 1995 and in paperback in 2001), and *Computer Storage Systems and Technology*, Wiley, 1977. He is also the author of chapters on memories in *Introduction to Computer Architecture* (H. Stone, Editor), SRA 1975 (1st Edition), 1980 (2nd Edition), and in *Electronics Engineers' Handbook*, 2nd and 3rd Editions, McGraw-Hill, 1982 and 1989. He also contributed the "Cache Memory" entry in the *Encyclopedia of Computer Science*, 3rd Edition, Van Nostrand Reinhold, 1993.

Dr. Matick is a member of Eta Kappa Nu and an IEEE Fellow.



**Hillery Hunter** received the Ph.D. degree in electrical engineering from the University of Illinois at Urbana-Champaign in 2004.

She is a Research Staff Member in the Exploratory System Architecture Department of IBM's T. J. Watson Research Center, Yorktown Heights, NY. She is interested in cross-disciplinary research, spanning circuits, microarchitecture, and compilers to achieve new solutions to traditional problems. Her current focus is on next-generation microprocessor caches and high-performance com-

mercial computing.

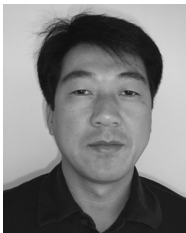


**Charles C. Tanner III** graduated from Technical Career Institute (TCI) of New York City in 1978.

He joined IBM Microelectronics in 1978, where he worked in a Future Logic Test System Design/Development Team for IBM products. From the mid-1980s to 1990, he was the Lead of the Lab Test Line Support Team. He has been doing SRAM/DRAM functional characterization from 1990 to the present. Currently, he is doing functional characterization of SOI embedded DRAM macros at IBM Systems and Technology Group, Burlington, VT.



**Joseph Harig** is a Senior Laboratory Specialist with IBM Systems and Technology Group, Burlington, VT. He joined IBM in Burlington as a memory tester technician in 1978. He has over 23 years experience in coding production DRAM test programs on Teradyne and Advantest testers. He currently is engaged in characterization and array pattern development for embedded DRAM products to support design verification.



**Hoki Kim** received the B.S. and M.S. degrees in electronics engineering from Yonsei University, Seoul, Korea, in 1988 and 1990, respectively, and the Ph.D. degree in electrical engineering from Texas A&M University, College Station, TX, in 2001.

He joined the Semiconductor Research and Development Center in Hyundai Electronics Industries (now Hynix Semiconductor), Icheon, Korea, and was with the company from 1990 through 1996 as a circuit designer. During his Ph.D. study, he worked for IBM T. J. Watson Research Center, Yorktown

Heights, NY, and Texas Instruments, Dallas, TX. Since 2001, he has been with the Autonomic Design Department of IBM System and Technology Group, East Fishkill, NY, working on embedded DRAM as an Architecture and Circuit Design Engineer. His research interests include high-speed circuit design in bulk and SOI technology, high-performance embedded memory design, cache memory design, and design for testability.



**Babar A. Khan** received the B.S., M.S., and Ph.D. degrees in electrical engineering from the Massachusetts Institute of Technology, Cambridge (MIT).

He was with Philips Research from 1984 to 1998, where his work included research on SOI films formed by laser and lamp recrystallization, polysilicon TFT displays, MOCVD of II-VI films, plasma addressed liquid crystal displays (LCDs), high-intensity discharge lamps formed by micromachining and reflective LCDs on silicon. Since 1998, he has been with IBM Microelectronics, Hopewell Junction, NY,

working on the development of over three generations of embedded DRAM technology. His interests include device design, process development, process integration, product debugging and product development.



**John Griesemer** received the B.S. degree in electrical engineering from Syracuse University, Syracuse, NY, in 1969, and the M.S. degree in materials science from Columbia University, New York, NY, in 1992.

He joined IBM Microelectronics SRDC, Hopewell Junction, NY, in 1996, where he worked in design post-processing and design integration of DRAM and LOGIC technology.

**Robert P. Havreluk** received the B.S. degree in electrical engineering from the New York Institute of Technology in 1974.

He joined IBM T. J. Watson Research Center, Yorktown Heights, NY, in 1974 and presently holds the position of Engineer. In the subsequent years, he

has contributed to the design of high-performance processors for IBM's S/390 servers and stand-alone DRAM for IBM systems. Recently, he has participated in the development of an embedded DRAM macro for Silicon-On-Insulator (SOI) technology and also participated in the early development of a novel memory known as MTJ MRAM, which has been commercialized by Freescale, Inc., and Honeywell, Inc.



**Kenji Yanagisawa** received the B.S. degree in electrical engineering from Rensselaer Polytechnic Institute, Troy, NY.

He joined IBM in 2005, joining the eDRAM team in East Fishkill, NY, as a Layout Design Engineer. He has since worked on creating custom layout for both eDRAM and eFUSE for 65 nm and 45 nm technologies.



**Toshiaki Kirihata** (A'92-SM'99) received the B.S. and M.S. degrees in precision engineering from Shinshu University, Nagano, Japan, in 1984 and 1986, respectively.

In 1986, he joined IBM Research, Tokyo Research Laboratory, IBM Japan Ltd., where he was involved in a 22 ns 1 Mb CMOS high-speed DRAM project. From 1989 to 1991, he was a Lead Engineer for a 14 ns 4 Mb high-speed DRAM circuit design. He then joined a low-power DRAM project at IBM Burlington Laboratory, IBM Microelectronics Division, Essex Junction, VT, working primarily on the 2/3 V<sub>dd</sub> sensing scheme

with a limited bitline swing. From January 1993 to June 1996, he worked on the 256 Mb DRAM joint development as a Lead Engineer with Toshiba and Siemens researchers at IBM, East Fishkill, NY. He then led a 256 Mb Synchronous DRAM design project, which was also a joint project with Toshiba and Siemens until March 1997. He transferred from the Tokyo Research Laboratory, IBM Japan, to IBM Research, T. J. Watson Research Center, Yorktown Heights, NY, in November 1996. From 1998 to 2000, he worked on a 390 nm 1 Gb DDR SDRAM design and a 512 Mb DDR2 SDRAM design as a Product Design Team Leader at IBM, East Fishkill, NY. He transferred to the IBM Microelectronics in 2000 as an IBM East Fishkill Design Department Manager, and led a 2.9 ns short random cycle embedded DRAM project with destructive read architecture. He is currently a Senior Technical Staff Member for IBM Systems and Technology Group, Hopewell Junction, NY, managing an embedded DRAM design department for high-performance embedded DRAM and electrical fuse development.

Mr. Kirihata presented papers at the ISSCC 1998, 1999, 2001, and 2004. He is a senior member of the IEEE.



**Subramanian S. Iyer** (M'77-SM'88-F'95) received the B.Tech. degree in electrical engineering from the Indian Institute of Technology, Bombay, in 1977, and the M.S. and Ph.D. degrees in electrical engineering from the University of California at Los Angeles in 1978 and 1981, respectively.

He is a Distinguished Engineer and Director of 45 nm CMOS Development at the Semiconductor Research and Development Center, IBM Systems and Technology Group, Hopewell Junction, NY. He joined the IBM T. J. Watson Research Center in 1981

and was manager of the Exploratory Structures and Devices Group until 1994, when he founded SiBond LLC to develop and manufacture Silicon-On-Insulator materials. He has been with the IBM Microelectronics Division since 1997. He has authored over 150 articles in technical journals and several book chapters and co-edited a book on bonded SOI. He has served as an Adjunct Professor of Electrical Engineering at Columbia University, New York, NY.

Dr. Iyer has received two Corporate Awards and four Outstanding Technical Achievement Awards at IBM for the development of the Titanium Salicide process, the fabrication of the first SiGe heterojunction bipolar transistor, and the development of embedded DRAM technology and the development of eFUSE technology. He holds over 35 patents and has received 19 Invention Plateau Awards. He received the Distinguished Alumnus Award from the Indian Institute of Technology, Bombay, in 2004. He is a Fellow of IEEE and a Distinguished Lecturer of the IEEE.