

A 45 nm SOI Embedded DRAM Macro for the POWER™ Processor 32 MByte On-Chip L3 Cache

John Barth, *Senior Member, IEEE*, Don Plass, Erik Nelson, Charlie Hwang, Gregory Fredeman, Michael Sperling, Abraham Mathews, Toshiaki Kirihaata, *Senior Member, IEEE*, William R. Reohr, Kavita Nair, and Nianzheng Cao

Abstract—A 1.35 ns random access and 1.7 ns-random-cycle SOI embedded-DRAM macro has been developed for the POWER7™ high-performance microprocessor. The macro employs a 6 transistor micro sense-amplifier architecture with extended precharge scheme to enhance the sensing margin for product quality. The detailed study shows a 67% bit-line power reduction with only 1.7% area overhead, while improving a read zero margin by more than 500ps. The array voltage window is improved by the programmable BL voltage generator, allowing the embedded DRAM to operate reliably without constraining of the microprocessor voltage supply windows. The 2.5nm gate oxide (G_{OX}) transistor cell with deep-trench capacitor is accessed by the 1.7 V wordline high voltage (VPP) with -0.4 V WL low voltage (VWL), and both are generated internally within the microprocessor. This results in a 32 MB on-chip L3 on-chip-cache for 8 cores in a 567 mm² POWER7™ die.

Index Terms—DRAM Macro, embedded DRAM Cache.

I. MOTIVATION

FOR several decades, the miniaturization of CMOS technology has been the most important technology requirements for increasing developing microprocessor performance and Dynamic Random Access Memories (DRAMs) density. However, the performance of the high-density DRAM has not kept pace with the high-performance microprocessor speed, hindering a system performance improvement. To address this performance gap, a hierarchical memory solution is utilized, which includes high speed Static Random Access Memories (SRAMs) as cache memories between a high-performance microprocessor and high density DRAM main memory.

Manuscript received April 16, 2010; revised June 29, 2010; accepted August 08, 2010. Date of publication November 22, 2010; date of current version December 27, 2010. This paper was approved by Guest Editor Ken Takeuchi. This material is based upon work supported by the Defense Advanced Research Projects Agency under its Agreement No. HR0011-07-9-0002.

J. Barth is with the IBM Systems and Technology Group, Burlington, VT 05452 USA, and also with IBM Microelectronics, Essex Junction, VT 05452-4299 USA (e-mail: jbarth@us.ibm.com).

E. Nelson is with the IBM Systems and Technology Group, Burlington, VT 05452 USA.

D. Plass, G. Fredeman, C. Hwang, M. Sperling, and K. Nair are with the IBM Systems and Technology Group, Poughkeepsie, NY 12601 USA.

A. Mathews is with the IBM Systems and Technology Group, Austin, TX 78758 USA.

T. Kirihaata is with the IBM Systems and Technology Group, Hopewell Junction, NY 12533 USA.

W. R. Reohr and N. Cao are with the IBM Research Division, Yorktown Heights, NY 10598 USA.

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/JSSC.2010.2084470

As technology is scaled in a nanometer generation, it is becoming significantly more difficult to enjoy a device scaling advantage, in part, due to increasing lithography challenges, as well as fundamental device physics limitation. Furthermore, it is even more important to improve the system performance to enable super-computing, which demands significantly larger cache memories with lower latencies. This results in a larger chip size with more power dissipation, where the embedded SRAM macro is one of the most significant area and power hungry elements. The first and second level cache memories have already been integrated in high-performance microprocessors [1], however, even with this approach it is difficult to meet the increasing system performance requirements. As a result, larger L3 cache integration [2] is the most important element for multi-thread, multi-core, next generation microprocessors.

High-performance and high-density DRAM cache integration with high performance microprocessor has long been desired, because the embedded DRAM 3x density advantage and 1/5 of the keep-alive-power compared to embedded SRAM. With on-chip integration, the embedded DRAM allows for the communication with the microprocessor core with significantly lower latency and higher bandwidth without a complicated and noisy off-chip IO interface [3]. The smaller size not only reduces chip manufacturing cost, but also achieves a faster latency from shorter wiring run length. In addition to the memory density and performance advantages, the embedded DRAM realizes 1000X better soft error rate than the embedded SRAM, and also increases the density of decoupling capacitors by 25X, using the same deep-trench capacitors to reduce on-chip voltage island supply noise.

Historically, integration of high density DRAM in logic technology started with ASIC applications [4], SRAM replacements [5], and off-chip high density cache memories [6], which have been already widely accepted in the industries. High density on-chip cache memory with embedded DRAM [7] was then employed in moderate performance bulk technology, which has leveraged supercomputers such as IBM's BlueGene/L [8]. As a next target, integration of high density embedded DRAM with a main-stream high-performance microprocessor is a natural step, however, because of ultrahigh performance requirement and SOI technology, it has not yet been realized.

This paper describes a 1.35 ns random access, and 1.7 ns random cycle embedded DRAM macro [9] developed for the POWER7™ processor [10] in 45 nm SOI CMOS technology. The high performance SOI DRAM macro is used to construct a large 32 MB L3 cache on-chip, eliminating delay, area, and power from the off-chip interface, while simultaneously improving system performance, reducing cost, power, and soft error vulnerability.

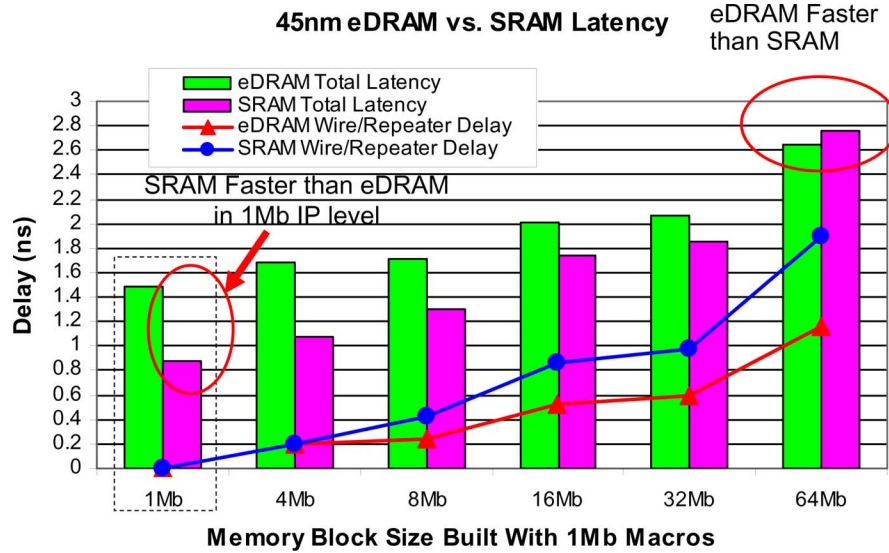


Fig. 1. 45 nm embedded DRAM versus SRAM latency.

Section II starts with the discussion with the density and access time trade-off between embedded DRAM and SRAM. Section III describes the embedded DRAM architecture. The discussion in Section IV moves into the details of the evolution for micro-sense amplifier designs and then explores the bitline high voltage generator design in Section V. To conclude this paper, Section VI shows the hardware results followed by a summary in Section VII.

II. EMBEDDED DRAM AND EMBEDDED SRAM LATENCY AND SIZE

The system level simulation shows that doubling the cache size results in respectable double digit percentage gains for cache-constrained commercial applications. Improving cache latency also has an impact on system performance. Placing the cache on-chip eliminates delay, power and area penalties associated with high frequency I/O channels required to go off-chip. Trends in virtual machine technology, multi-threading and multi-core processors further stress on the over taxed cache sub-systems.

Fig. 1 shows the total latency and the total area for embedded DRAM cache and embedded SRAM cache memories in a microprocessor. The latency and the size were calculated on the basis of the existing embedded DRAM and SRAM macro IP elements having 1 Mb building unit both in 45 nm SOI CMOS technology. Although embedded DRAM performance has been significantly improved over the past 5 years, embedded SRAM still holds a latency advantage at the 1 Mb macro IP level, showing approximately half that of DRAM macro. However, if one takes a system level perspective when building a large memory structure out of discrete macros, one realizes that wire and repeater delays become a significant component as shown. As the memory structure becomes larger and the wire delay becomes dominant, the smaller of the two macros will have the lower total latency. The cross-over point of the latency is approximately 64 Mb, where embedded DRAM realizes a lower total latency than embedded SRAM.

III. MACRO ARCHITECTURE

Fig. 2 shows the architecture of this embedded DRAM macro [9]. The macro is composed of four 292 Kb arrays and input/output control block (IOBLOCK), resulting in a 1.168 Mb density. The IOBLOCK is the interface between the 292 Kb arrays and processor core. It latches the commands and addresses, synchronizing with the processor clock, and generates sub-array selects, global word-line signals. It also includes a concurrent refresh engine [11] and a refresh request protocol management scheme [12] to maximize the memory availability. A distributed row redundancy architecture is used for this macro, resulting in no redundancy array.

Each 292 Kb array consists of 264 word-lines (WLs) and 1200 bit-lines (BL), including eight redundant word-lines (RWLs) and four redundant data-lines (RDLs). Orthogonally segmented word-line architecture [13] is used to maximize the data bus-utilization over the array. In this architecture, the global word-line-drivers (GWLDVs) are arranged in the IOBLOCK located at the bottom of the four arrays. The GWLDVs drive the global WLs (GWLs) over the four arrays using 4th metal layers (M4). The GWLs are coupled to the Local Word-Line Drivers (LWLDVs), located adjacent to the sense amplifier area in each array. This eliminates the necessity to follow the pitch limited layout requirement for the LWLDVs, improving the WL yield. Each LWLDV drives the corresponding WL by using vertically arranged metal 4 layers (M4) over the array. The M4 WLs are coupled to the 3rd metal layer (M3) WLs, which run horizontally, parallel to the on pitch WL. The WLs are finally stitched to the poly WL at every 64 columns to select the cells.

The 292 Kb array are also divided into eight 33 Kb micro-arrays for micro-sense-amplifier architecture [13]. 32 cells with an additional redundant cell (total 33 cells) are coupled to the Local Bit-Line (LBL). This enables a maximum of 8 row repairs in any array, however to save fuse latch area, only 16 repairs can be made per macro. Similar to the row redundancy array

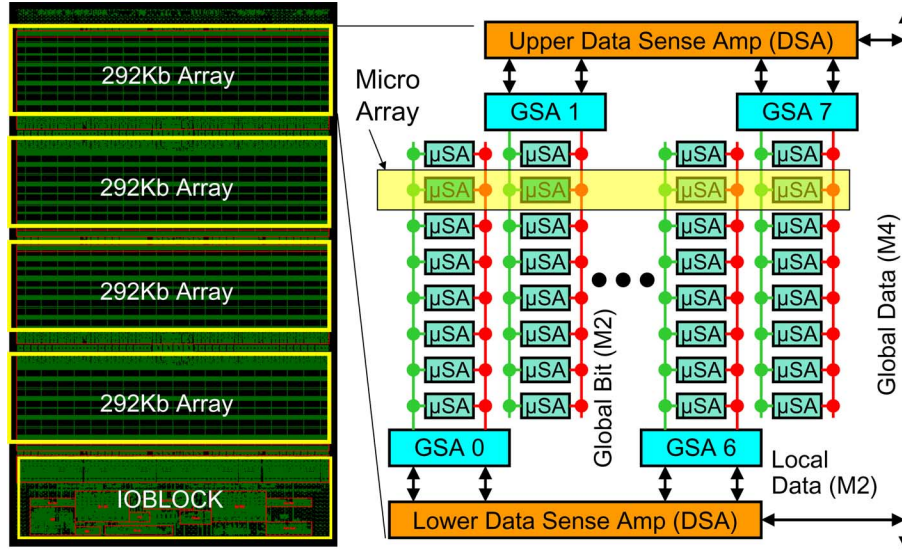


Fig. 2. Macro architecture.

architecture in 65 nm SOI embedded DRAM [13], this scheme offers a good tradeoff between area and repair region size.

The Micro sense architecture is a hierarchical scheme that relies on a high transfer ratio during a read to create a large voltage swing on a LBL; large enough to be sampled with a single-ended amplifier.

The micro sense amp (μ SA) transfers data to/from a global sense amplifier (GSA) along two global bit-lines labeled read bit-line (RBL) and write bit-line (WBL). The metal 2 global bit-lines, routed in parallel to the metal 1 LBL, control the read/write operations to the μ SAs. The uni-directional WBL controls write '0' while the bidirectional RBL manages both read and write '1'.

The μ SA adds an extra level of hierarchy and necessitates a third level data sense amp (DSA). The bidirectional DSA is responsible for transferring data between the Metal 4 global data lines and the selected GSA. One of eight Micro Arrays are selected horizontally in the Y dimension using the master word-line (MWL) decodes while one of eight GSA are selected in the X dimension by the column signals. In order to meet the tight pitch of the array, GSA are interleaved with 4 above the array, supported by an upper DSA and 4 below, supported by a lower DSA. Both DSAs share common metal 4 data lines.

The column select signal (CSL) selects one out of 8 GSAs such that the data bit in the selected column is transferred to the single-ended read dataline complement (RDC) or from write dataline true/complement (WDT/Cs) by the DSA. In a set associative cache, the one-hot column select would be used as a late way select, achieving a high speed latency for the POWER7™ microprocessor. A total of 146 RDCs and 146 WDT/C pairs are arranged using the 4th metal layers over four arrays. An additional four redundant RDCs and four WDT/C pairs support two out of 73 data-line redundancy repair on each left and right column domain, resulting in total 150 datalines per macro.

IV. EVOLUTION OF MICRO-SENSE AMPLIFIERS

Conventional techniques for improving DRAM performance involve reducing the bit-line (BL) length. This short-BL architecture [14], however, increases the area overhead due to additional sense amplifiers, bit-line twisting, and reference circuits. The area overhead is significantly increased when the BL length shorter than 128 cells/BL is employed, which makes the embedded DRAM cache solution less attractive. Sense amplifier area is further degraded with body-tied SOI devices, required to prevent history-induced sense amp mismatch on small signal, long bit-line architectures. The micro sense-amp (μ SA) architecture is introduced to provide high performance sensing, without incurring the overhead associated with conventional techniques.

A. Three Transistor μ SA (3T μ SA)

Fig. 3(a) shows the array design featuring 3 transistor μ SA architecture (3T μ SA) [13]. In this approach, only a small number of cells (32) are connected to the LBL for each column in a sub-array. A 18 fF deep-trench capacitor (DT) is used in combination with a 3.5 fF LBL, resulting in 84% transfer ratio. The LBL is coupled to the gate of the NFET read head transistor (RH). The sensing operation relies on the ultrahigh transfer ratio during a read to create a large voltage swing on a LBL, large enough to turn on the NFET RH as a single-ended BL sense-amplifier. The single-ended LBL arrangement enables relaxed 1st metal layer (M1) pitch, increasing line to line space by 3X. The LBL is supported by the PFET feedback device (FB) and NFET pre-charge/write 0 device (PCW0). The 3T μ SA transfers data to/from a GSA via two Global Read/Write Bit-Lines (RBL/WBL) using second metal layers (M2). The RBL and WBL wires are arranged over the M1 LBLs. Each GSA, in turn, services 8 μ SAs, supporting 256 cells for each column in an array as a hierarchical fashion.

During pre-charge, WL is low, turning all cell transfer devices off. The WBL and RBL are pre-charged to high, which turns on

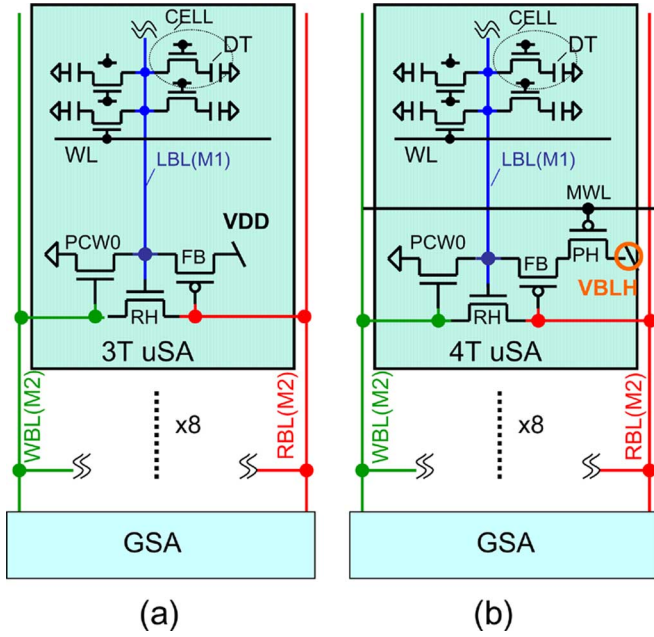


Fig. 3. 3T and 4T micro-sense amplifiers.

the NFET PCW0 and turns off the PFET FB. The LBL is therefore precharged to GND through the PCW0. Prior to WL activation, WBL is driven low by the GSA, which turns off the PCW0. As a result, the LBL will be floated at GND level, waiting for the signal transfer from the cell. When WL rises to 1.7 V, the signal development starts on the LBL.

When the cell stores “0” data, the LBL remains at low voltage, keeping the RBL at high level. The GSA senses a “1” at RBL, interpreting this to be a “0” value. Upon Detection of “0”, the GSA drives the WBL high, forcing the LBL to low with NFET PCW0, writing back “0” data to the memory cell.

When the cell stores “1”, the LBL is pulled high by the charge sharing between high storage node and GNDed BL. The 32 cells/LBL allows the large LBL swing, turning on the NFET RH. This results in discharging the RBL through the WBL, to GND through the NFET RH. When the RBL drops a PFET threshold below the PFET FB source voltage, the PFET FB turns on, driving the LBL high, providing a positive feedback to the NFET RH, further accelerating the low-going RBL swing. As a result, the cells are naturally restored to the full VDD in a short period without having any additional timing requirement.

When a 1 is to be written, RBL is pulled to ground by the GSA. This allows PFET FB to turn on, making the LBL high. The operation will occur at the same time when the WL rises. This results in writing the high voltage to the cell very early in the cycle as a direct write [4]. For writing “0” to the cell, the WBL stays high, which keeps the NFET PCW0 on. This clamps the LBL at GND during the WL activation, allowing the writing of the 0 data to the corresponding cells.

B. Four Transistor μ SA (4T μ SA)

The PFET FB in 3T μ SA is the key device to amplify the LBL as fast as possible. This is achieved by giving a positive feedback to the LBL when the LBL swings below the threshold of the PFET FB. However, a leakage current or noise coupling which

makes RBL go low, also contributes to inadvertently enabling the PFET FB. This may result in a false amplification when “0” data are read. In fact, the RBL will go low regardless of the cell data as time goes by, due to the leakage path to GND through the NFET RH whose source is at GND in a read mode. The RBL is connected in parallel to all 8 NFET RH devices in 8 sub-arrays, resulting in 8X leakage path to GND. To complicate matters, the small RBL discharge event also amplifies the LBL in unselected segments. Because the LBLs in unselected sub-arrays are lightly loaded, (not coupled to the cell capacitor), the positive feedback to the LBLs in the unselected segments are much faster than that of the selected sub-array coupling to the cell. As a result, high-going LBL feedback by the FB PFET in the any unselected sub-array also contributes to discharging the RBL.

Fig. 3(b) shows the four transistor μ SA design. In order to save power and overcome the PFET leakage problem, the PFET header (PH) is introduced. The gate of the PFET PH is controlled by the master-wordline signal (MWL), which runs perpendicular to the LBLs. The signal MWL goes low when a WL in the corresponding sub-array is activated. This enables a positive PFET FB feedback only to the selected LBL. The MWL signals in unselected sub-arrays stay high, preventing a positive feedback to the LBLs in the unselected sub-arrays.

In addition to the FET PH inclusion in μ SA, the source of the PH is coupled to the BL high voltage supply (VBLH). The VBLH voltage is generated by VBLH generator located on the top of the embedded DRAM macro, and optimized for the μ SA operation during the BIST test. The 4T μ SA design with PFET PH and VBLH supply reduces not only the stand-by power by the PFET PH and FB stacking, but also AC power by preventing the unnecessary transition on unselected LBLs in an active state.

C. Line-to-Line Coupling in 3T and 4T μ SA

The tight M2 RBL and WBL in 3T and 4T μ SA architecture create a high coupling ratio and disturbance effects during a sense and write operations. Fig. 4 describes three potential coupling mechanisms for the adjacent three columns, creating during a write to the center column, labeled as the write aggressor. The first mechanism involves writing a ‘1’, where the center RBL (RBL_C) falls, coupling the right WBL (WBL_R) below ground. The WBL_R bounces the source of the read head device (RH) to below ground, increasing the RBL leakage. This reduces the 0 sensing margin. The second mechanism involves writing a ‘0’, when the center WBL (WBL_C) rises, it couples the left RBL (RBL_L) above VDD, delaying the positive feedback and refresh of ‘1’ resulting in a performance degradation. The third mechanism, involves the half-selected LBL that share the same global. When the RBL falls, it couples the floating half selected LBL below ground, effectively increasing V_{gs} and array device leakage. The cells in unselected sub-arrays see more negative coupling than the selected sub-array, because the LBLs in the unselected sub-arrays are lightly loaded (not coupled to the cell capacitor). This results in retention time degradation. These three coupling mechanisms may create various pattern sensitivities, which not only reduce the yield of the embedded DRAM macro, but also make it difficult to find the weak cells. Six transistor μ SA (6T) is introduced to overcome

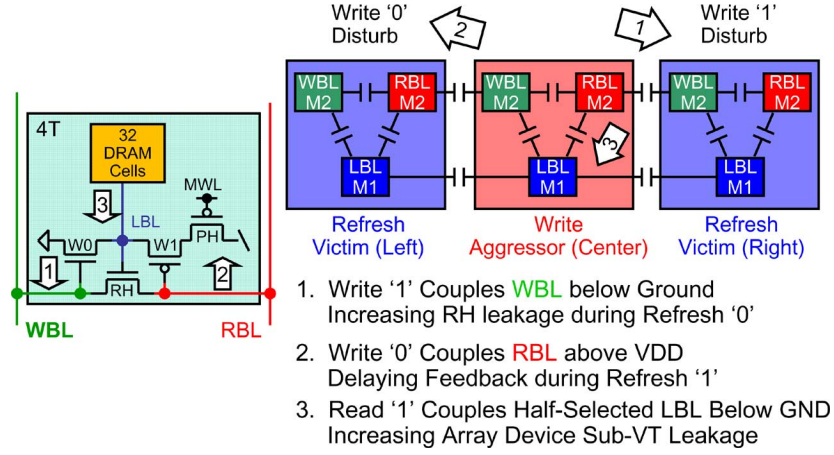


Fig. 4. Line-to-line coupling mechanisms.

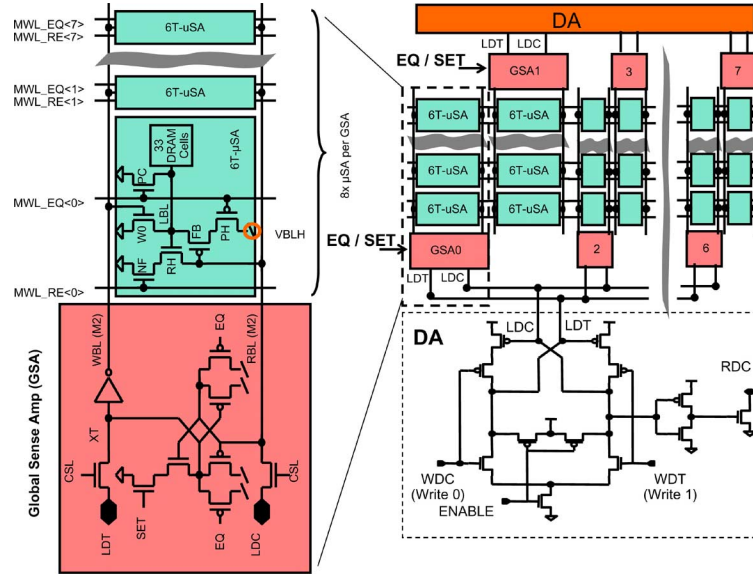


Fig. 5. 6T macro-sense amplifier architecture.

the line-to-line problem, improving the performance while reducing power dissipation.

D. Six Transistor μ SA (6T μ SA)

Fig. 5 shows the detailed 6T μ SA array architecture, which includes a 6T μ SA, a GSA, and DSA. The NFET pre-charge/write0 device (PCW0) in 3T/4T μ SA design is split into a precharge NFET (PC) and a write-0 NFET (W0). In addition to the PC and W0 separation, the NFET footer device (NF) is included to enable the NFET read head (RH). The master-wordline equalization signal (MWLEQ) controls PC and PH devices while the master-wordline read-enable signal (MWLRE) controls the NF device. The MWL signals are routed on M3, perpendicular to the LBL, and activated only in the selected sub-array.

All cycles start in the pre-charge condition with GSA equalization signal (EQ) low, holding RBL high and WBL low. When the sub-array is selected, the signal MWLEQ goes low in the selected sub-array. This disables the NFET PC, floating the selected LBL. MWLEQs in the unselected sub-arrays stay

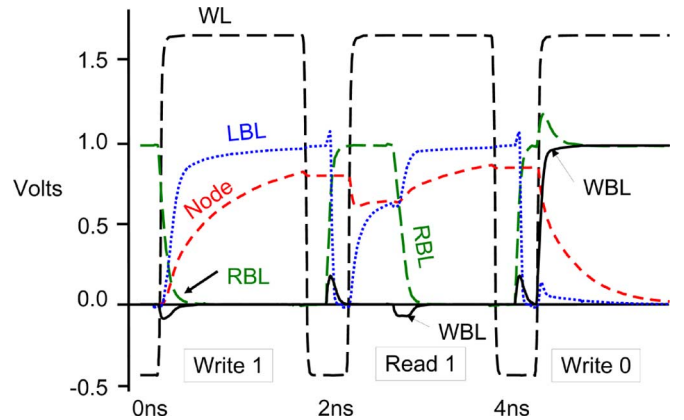


Fig. 6. Simulated Write 1 Read 1 Write 0 waveforms.

high, clamping the un-selected LBLs at GND. This overcomes the coupling mechanism-3 and eliminates the chance of an unselected LBL drifting high causing a read '0' failure. The low-going MWLEQ also turns on the pMOS head device

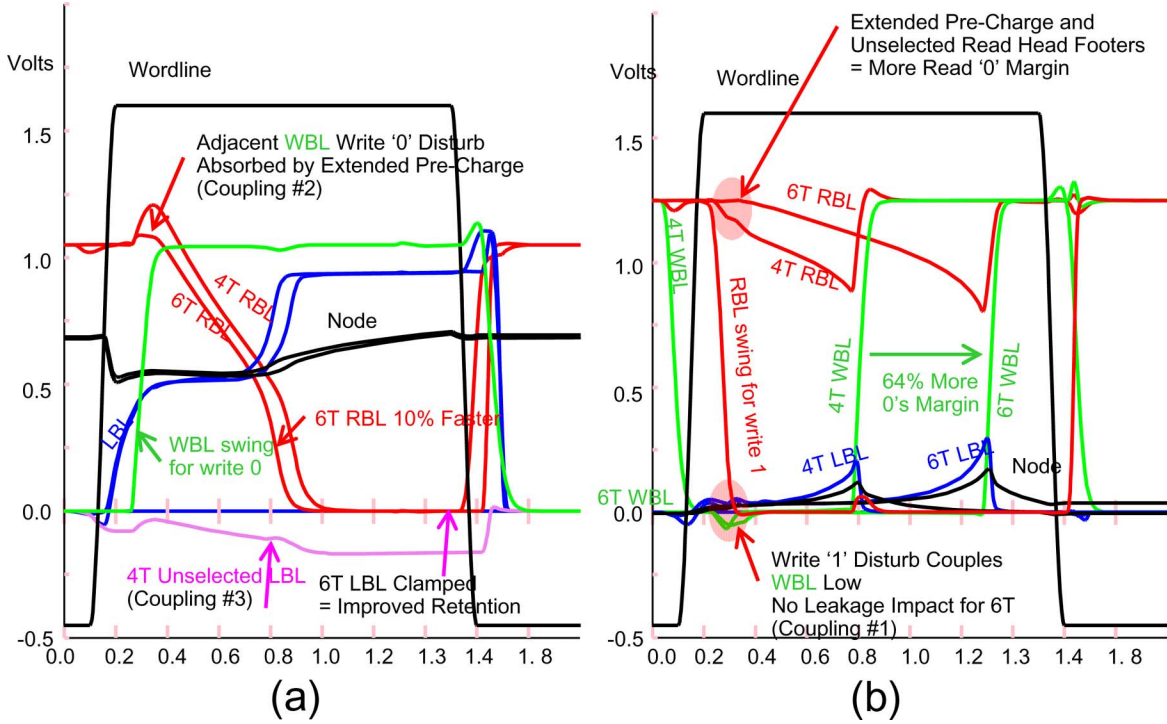


Fig. 7. Simulated line-to-line coupling effect for 4T and 6T μ SA sensing operations. (a) Read 1. (b) Read 0.

(PH), enabling the PFET feedback device (FB) for direct write. In order to manage the write coupling created by the tightly spaced M2 runs, the extended pre-charge scheme absorbs coupling caused by writing to an adjacent RBL/WBL pair. This is realized by controlling the signal EQ for the top and bottom GSA independently. When writing to the even column, the EQ controlling the even columns (lower GSAs) is released to enable direct data write, while the EQ controlling the odd columns (upper GSAs) is held in pre-charge until write transition is complete. The extended pre-charge scheme absorbs coupling mechanisms 2 and 3 without creating a write power burn or impacting refresh cycle time. The write power burn is avoided by delaying the activations of the master word line read enable signal (MWL_RE). This delay does not impact refresh cycle time because LBL signal is not fully developed until after the write data have been written. This delay also favors a read zero by reducing the amount of time RBL is exposed to leakage of the Read Head RH device.

Write data is delivered to the DSA, via M4 write data lines (WDT/WDC). Initially low, WDT or WDC is remotely driven high during a write. To write a '1', WDT is driven high, DSA will drive Local Data-line Complement (LDC) low, pass to the column selected by the corresponding CSL, and pull RBL low. This forces LBL high, writing a '1' into the node of the selected cell. To write a '0', WDC is remotely driven high, DSA will drive Local Data-line True (LDT) low and pass to XT of the GSA selected by the corresponding CSL. This forces WBL high, driving LBL low, writing a '0' into the selected cell.

When a read command is accepted, read data are transferred from the cell to the LBL upon activation of the selected word-line (WL). After a small delay to allow signal to develop, a master-word-line read enable signal (MWL_RE) is activated for

the μ SA in the selected sub-array. For a stored '1', LBL will rise at least 1 threshold above the read head NFET (RH), weakly pulling RBL low. When RBL falls below the threshold of the feedback PFET (FB), LBL is driven to a full high level. This amplifies LBL charge, refreshes the cell, and strongly drives RBL low. Note that refresh of a '1' is self-timed, requiring no external control. RBL falling will pass from the selected column to the DSA, driving M4 read data line RDC low. For a stored '0', LBL remains low, RBL remains high and WBL remains low until external timing signal SET triggers GSA to evaluate RBL. With RBL high, XT falls, driving WBL high, driving LBL low and refreshing the '0'. Fig. 6 shows the simulated waveforms for write 1, read 1, and write 0, demonstrating successful operation of 500 MHz random cycle operation.

E. Analysis

Fig. 7(a) shows simulated array waveforms for a refresh '1' in the 4T and 6T μ SAs. This simulation is done with a 5 sigma worst case array device and 4.5 sigma slow read head device (RH) at low temperature and low voltage. This extreme condition is meant to demonstrate the difference in the design points for analyzing sensing "1". As discussed in the previous section, the high-going WBL for the write 0 couples into the adjacent RBL (Write 0 Disturb). Large up-coupling on the floating 4T RBL can be seen to delay RBL discharge, resulting in a 10% delay increase from WL rise. Additionally the low-going RBL couples the Half Select LBL below Ground, creating a retention loss in 4T μ SA, however, the 6T successfully clamps the half selected LBL at Ground and sees no retention loss.

Fig. 7(b) shows simulated array waveforms for a refresh '0' in the 4T and 6T μ SAs. To simulate the worst case scenario for sensing "0", this simulation is done with 4.5 Sigma Fast Read

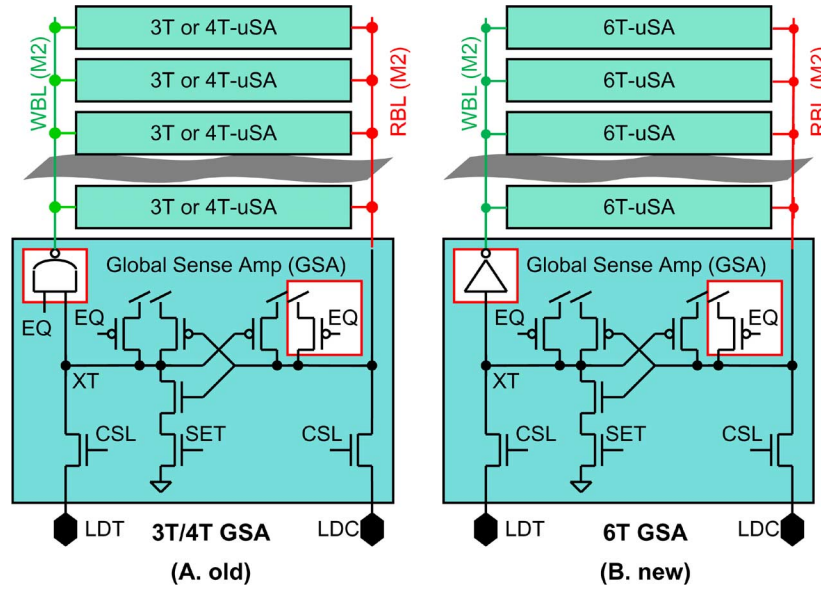


Fig. 8. Global sense amp simplifications.

	μ SA 4T $\Rightarrow$ 6T	GSA Simplification	Distributed Redundancy	Net
PC Tracks	+28	-16	-4	+8 (+1.7%)

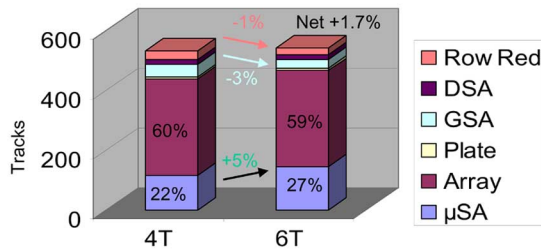


Fig. 9. Sense amp area.

Head device and operated at an elevated Voltage and Temperature to increase leakage and coupling. Once again this extreme condition is meant to demonstrate the difference in the design points. The low-going RBL couples the adjacent WBL below GND (Write 1 Disturb). Small down-coupling on the both 4T and 6T WBL can be seen when the adjacent RBL goes low. In the 4T Case, it creates Read Head Source Leakage through all 8 Read Head (RH) on the WBL. The 6T Read Head Source is not connected to WBL and is not subject to increased leakage. Additionally the 6T RBL is clamped high by the extended pre-charge. When the read is enabled, only 1 of 8 are enabled, further reducing leakage on RBL. The simulated waveforms show a slope difference in RBL leakage due to Half Select Read Head leaking in the 4T case and Footed off in the 6T case. The combinations of extended pre-charge and read footers increase read zero margin by more than 500 ps, or 64% as measured from Word-Line Rise, significantly improving the set timing window.

This 6T μ SA architecture does require more devices to support a LBL, however, the design of the GSA can be simplified as shown in Fig. 8. The 3T and 4T share the same GSA design

Refresh '1' Active Power Savings

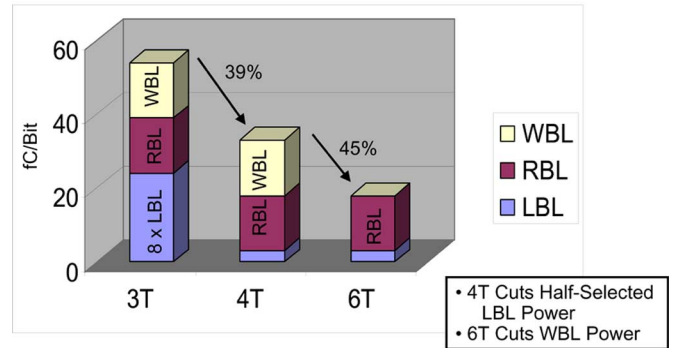


Fig. 10. Active power savings.

Fig. 8(a). Because WBL controls both Local Pre-Charge and Write zero, a NAND gate was required in the GSA to merge these functions. It should also be noted that the NAND gate needed to be large enough to sink the RBL discharge current during a read '1'. Finally, a large EQ device was required to absorb RBL coupling from WBL falling at the beginning of the cycle. The 6T control gets simpler, by adding independent pre-charge control to the uSA, WBL can remain low in standby and the NAND gate in the 3T/4T GSA can be converted into a simple inverter in 6T GSA shown in Fig. 8(b). Furthermore, adding the NFET footer, locally sources the RBL discharge current, allowing the inverter to be reduced in size. Without the WBL low transition at the beginning of the cycle, the RBL EQ device, no longer needs to absorb WBL down coupling and can also be reduced in size. These simplification and device size reductions in the GSA offset the area increase in the 6T μ SA. Fig. 9 accounts for the area changes in uSA/GSA architecture in terms of logic gate tracks. Although the 6T increase the uSA area by 5%, Design simplification and distribution row redundancy result in almost a 4% reduction, overall only 1.7% area increase in the 256 Kb Array was realized, which is negligible

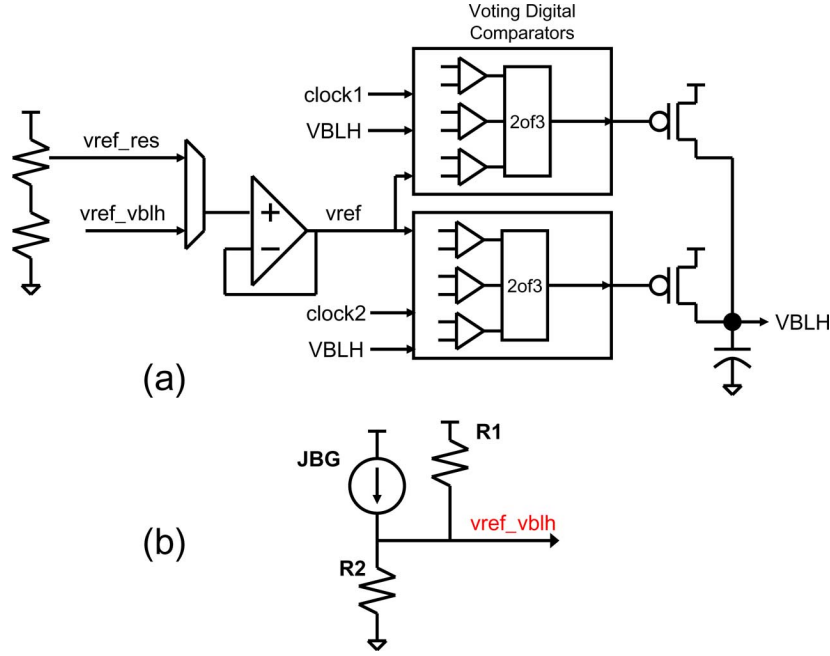


Fig. 11. BL high voltage generator: (a) regulator circuit, (b) reference voltage circuit.

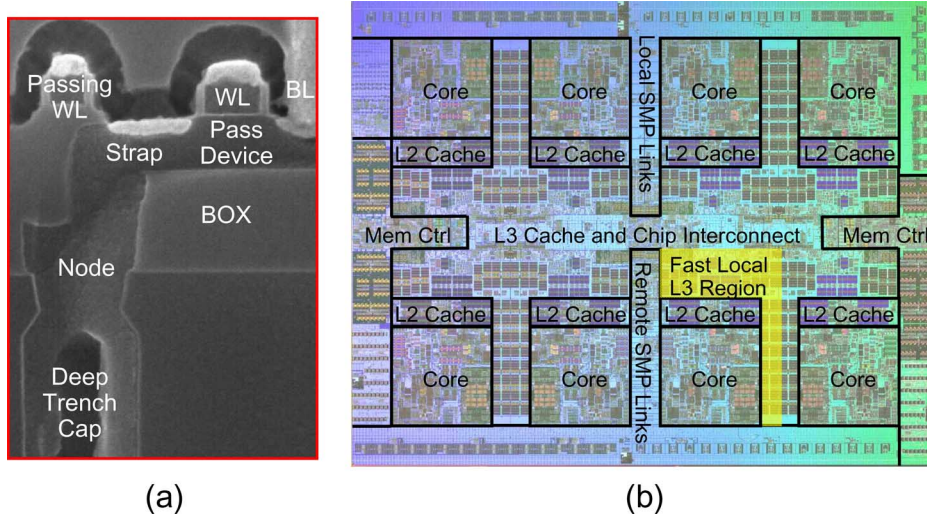


Fig. 12. (a) Cross section of deep-trench cell and (b) chip microphotograph for POWER7™ microprocessor.

at the microprocessor level. Fig. 10 accounts for active power savings realized by 1) The 4T PFET Header eliminating the unselected LBL (39%), while the 6T μ SA eliminated the WBL pre-charge power, representing a 45% Bit-Line Power Reduction over the 4T μ SA.

V. BL VOLTAGE GENERATOR

A BL high voltage (VBLH) generation is accomplished through a digital linear voltage regulator that modulates the gate of a PFET transistor connected between the power supply and the output voltage VBLH. The PFET gate is pulsed low if a comparator decides that VBLH is below the reference voltage. This operation charges the capacitance on VBLH and supplies some transient current to the circuit. Once charging has brought

the VBLH voltage above the reference voltage, the pulsing ceases and the PFET gate is driven high, turning it off.

The digital voltage regulator is composed of a resistor ladder, input reference mux, unity gain buffer, digital comparators and a pair of very wide output PFETs as seen in Fig. 11(a). The input reference mux is used to select an analog reference voltage that will be buffered to the digital comparators via an analog unity gain buffer. This analog reference voltage can be chosen using digital tune bits as either a tap off a resistor ladder between the power supply and ground, or an external analog voltage, vref_VBLH, which is based on both a constant voltage as well as the power supply.

Fig. 11(b) shows the external analog circuitry necessary to generate a highly accurate voltage with these characteristics. J_{BG} is a current which is generated by buffering a bandgap

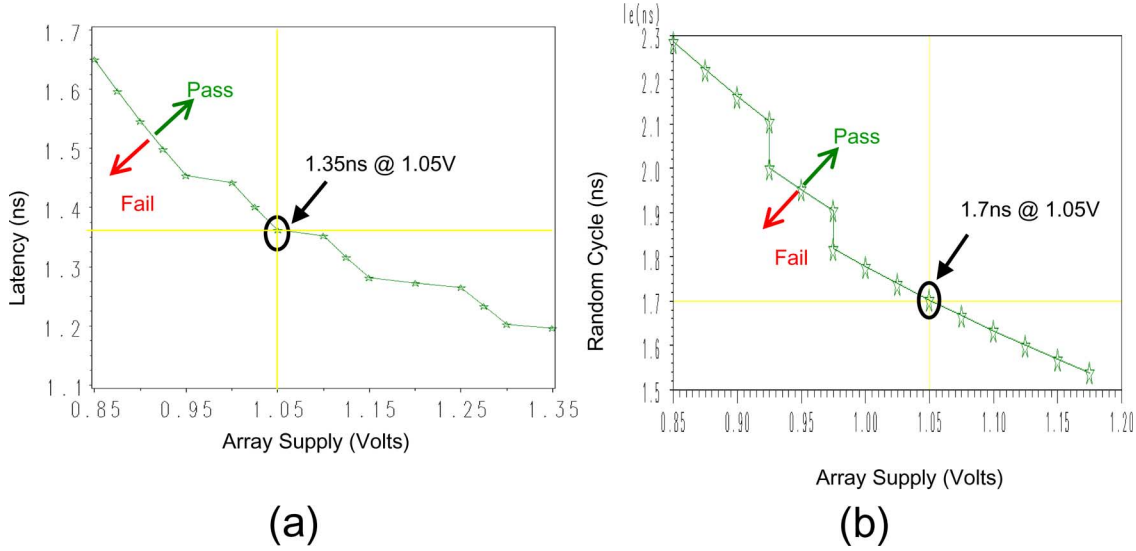


Fig. 13. Hardware results: (a) random access time, (b) random cycle time.

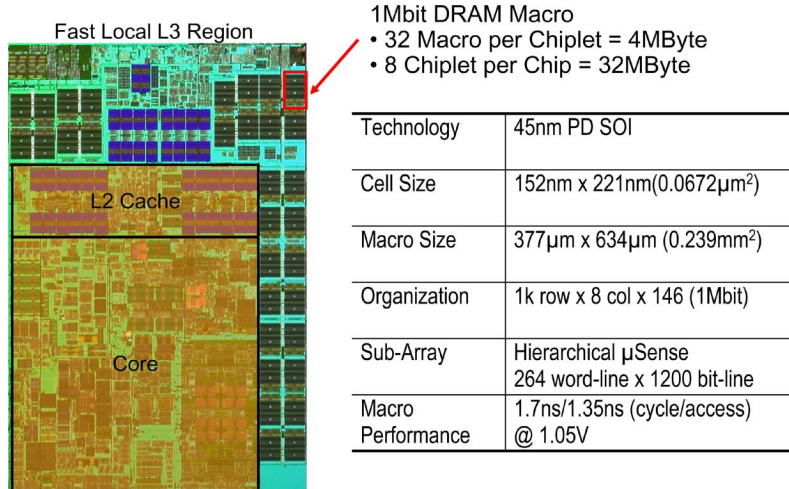


Fig. 14. POWER7™ single core microphotograph and embedded DRAM features.

voltage V_{BG} to a resistor R_{BG} and so is equal to V_{BG}/R_{BG} . Resistors $R1$ and $R2$ complete the circuit, creating the following voltage equation that is based on resistor ratio rather than absolute resistance. Note that VCS is the power supply voltage:

$$v_{ref\_vblh} = \left(\frac{V_{BG}}{R_{BG}} * [R1 || R2] \right) + \left(VCS * \frac{R2}{R1 + R2} \right)$$

This circuitry is located outside the DRAM macro in the charge pumps that generate VPP & VWL . In this way, one large bandgap generator that is already being used for the charge pumps can provide a v_{ref_VBLH} voltage for multiple DRAM macros without a huge area penalty.

The unity gain buffer is composed of a basic folded cascode differential amplifier and is used to provide a low impedance input voltage to the digital comparators. This is critical because the digital comparators will pre-charge for half a cycle before sensing the voltage difference in the next half cycle.

In order to reduce the tolerance of the comparators, three comparators vote on whether the regulated voltage was above

or below the reference voltage. This vote is pulsed to the PFET output, charging the output capacitance and increasing the output voltage. There are two banks of these voting comparators, each operating on opposite clock phases. This has the advantage of decreasing the sampling latency as well as reducing the ripple on the output.

Decoupling on the VBLH supply is provided by deep trench capacitance, similar to the structures in the DRAM. The capacitance is placed between the supply and ground and is spread out throughout the DRAM macro to provide a low impedance path to all the circuits requiring VBLH. The total area of the capacitance is determined by the transient current loading since most of the instantaneous charge must come from this capacitance. The regulator will restore this charge on the next regulator clock cycle. The VBLH generator lowers the POWER7™ VDD wide voltage supply window to stable eDRAM array voltage window. This will minimize body charging within the SOI NFETs of the memory cells, improving the embedded DRAM yield.

VI. HARDWARE RESULTS

POWER7™ architects took full advantage of this concept, using eDRAM to build a fast local L3 region very close to each core to maximize L3 performance. Fig. 12(a) shows an SEM of the Deep Trench Capacitor and Array Device. High capacitance is achieved by digging through the SOI and Buried Oxide, 4.5 μm into the P-type substrate. The expanded red box shows the Bit-Line Contact, Array Pass Device and silicide strap forming a low resistance connection to a 18 fF deep trench. The DRAM cell is fabricated on SOI and utilizes the same thick oxide as the base technology. Conveniently, the buried oxide provides isolation between the SOI and substrate, eliminating the need for the oxide collar required by bulk technologies to protect against vertical parasitic leakage. In addition to utilizing the trench for memory, POWER7™ also used the high capacitance trench for on-chip power supply decoupling, offering 25 $\times$ more capacitance than planar structures.

Fig. 12(b) shows the chip micro-photograph of the POWER7™ microprocessor. It consists of 8 cores for a total of 32 MBytes of shared L3 cache in a 567 mm^2 die. The 1.2 Billion transistor design has the equivalent function of a 2.7B transistors processor due to the efficiency of 1T1C embedded DRAM. When compared to a typical processor that dedicates 40–50% of its die area to L3, embedded DRAM only consumes 11% of the die. One would ask why we did not add even more memory? The POWER7™ architects instead chose to use the area to balance the overall system performance by adding off chip bandwidth in the form of DDR3 controllers and SMP coherency links. POWER7™ boasts over 590 GB/s of total bandwidth and supports 20 thousand Coherent SMP Operations.

Fig. 13 shows a graph of random access and cycle times as a function of array supply voltage over the full address space, demonstrating 1.7 ns random cycle @1.05 V, which also corresponds to 1.35 ns access. The macro was characterized via a built-in self test engine. Fig. 14 shows the chip micrograph of a single POWER7™ Core and the 4 MB local portion of the L3, comprised of 32 DRAM macro instances. Fig. 14 also provides a table to summarize the features of the embedded DRAM macro.

VII. SUMMARY

We have developed a high density and high performance eDRAM macro for the highly parallel, scalable, next generation POWER7™ microprocessor. The evolution of the 6T μSA architecture improves sub-array latency by 10% and 0's timing margin by 500 ps. This results in a 1.35 ns random access time and a 1.7 ns random cycle time while reducing keep-alive power by 45% with a silicon overhead of 1.7%. Thirty-two embedded DRAM macros constructs 4 MB L3 cache memory per core, resulting in eight cores in the 567 mm^2 POWER7™ chip using 1.2B transistors in 45 nm SOI CMOS. This is the most advanced VLSI design using embedded DRAM so far reported. The integration of high density high performance eDRAM in microprocessor has just begun and is expected to open new era for next generation VLSI designs.

ACKNOWLEDGMENT

The authors thank East Fishkill Technology Development and Manufacturing Teams, Burlington Test and Characterization Teams, Yorktown T. J. Watson Research Support, and Poughkeepsie and Austin Design Centers.

REFERENCES

- [1] J. Friedrich, B. McCredie, N. James, B. Huott, B. Curran, E. Fluhr, G. Mittal, E. Chan, Y. Chan, D. Plass, S. Chu, H. Le, L. Clark, J. Ripley, S. Taylor, J. Dilullo, and M. Lanzerotti, "Design of the POWER6™ microprocessor," in *IEEE ISSCC Dig. Tech. Papers*, Feb. 2007, pp. 96–97.
- [2] S. Rusu, S. Tam, H. Muljono, J. Stinson, D. Ayers, J. Chang, R. Barada, M. Ratta, S. Kottapalli, and S. Vora, "A 45 nm 8-core enterprise XEON™ processor," *IEEE J. Solid-State Circuits*, vol. 45, no. 1, pp. 7–14, Jan. 2010.
- [3] H. Fujisawa, S. Kubouchi, K. Kuroki, N. Nishioka, Y. Riho, H. Noda, I. Fujii, H. Yoko, R. Takishita, T. Ito, H. Tanaka, and M. Nakamura, "An 8.1-ns column-access 1.6-Gb/s/pin DDR3 SDRAM with an 8:4 multiplexed data-transfer scheme," *IEEE J. Solid-State Circuits*, vol. 42, no. 1, pp. 201–209, Jan. 2007.
- [4] J. Barth, D. Anand, J. Dreibelbis, J. Fifield, K. Gorman, M. Nelms, G. Pomichter, and D. Pontius, "A 500-MHz multi-banked compliant DRAM macro with direct write and programmable pipeline," *IEEE J. Solid-State Circuits*, vol. 40, no. 1, pp. 213–222, Apr. 2005.
- [5] H. Pilo, D. Anand, J. Barth, S. Burns, P. Corson, J. Covino, and S. Lamphier, "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.
- [6] T. Okuda, I. Naritake, T. Sugibayashi, Y. Nakajima, and T. Murotani, "A 12-ns 8-MByte DRAM secondary cache for a 64-bit microprocessor," *IEEE J. Solid-State Circuits*, vol. 35, no. 8, pp. 1153–1158, Aug. 2000.
- [7] J. Barth, D. Anand, J. Dreibelbis, and E. Nelson, "A 300 MHz multi-banked eDRAM macro featuring GND sense, bit-line twisting and direct reference cell write," in *IEEE ISSCC Dig. Tech. Papers*, Feb. 2002, pp. 156–157.
- [8] S. S. Iyer, J. Barth, P. Parries, J. Norum, J. Rice, L. Logan, and D. Hoyaniak, "Embedded DRAM the technology platform for the BlueGene/L chip," *IBM J. Res. Dev.*, vol. 49, no. 2, 3, pp. 333–349, 2005.
- [9] J. Barth, D. Plass, E. Nelson, C. Hwang, G. Fredeman, M. Sperling, A. Mathews, W. Reohr, K. Nair, and N. Cao, "A 45 nm SOI embedded DRAM macro for POWER7™ 32 MB on-chip L3 cache," in *IEEE ISSCC Dig. Tech. Papers*, Feb. 2010, pp. 342–343.
- [10] D. Wendel, R. Kalla, R. Cargoni, J. Clables, J. Friedrich, R. Frech, J. Kahle, B. Sinharoy, W. Starke, S. Taylor, S. Weitzer, S. G. Chu, S. Islam, and V. Zyuban, "The implementation of Power7™: A highly parallel and scalable multi-core high-end server processor," in *IEEE ISSCC Dig. Tech. Papers*, Feb. 2010, pp. 102–103.
- [11] T. Kirihata, P. Parries, D. R. Hanson, H. Kim, J. Golz, G. Fredeman, R. Rajeevakumar, J. Griesemer, N. Robson, A. Cestero, B. A. Khan, G. Wang, M. Wordeman, and S. S. Iyer, "An 800 MHz embedded DRAM with a concurrent refresh mode," *IEEE J. Solid-State Circuits*, vol. 40, no. 6, pp. 1377–1387, Jun. 2005.
- [12] P. J. Klim, J. Barth, W. R. Reohr, D. Dick, G. Fredeman, G. Koch, H. M. Le, A. Khargonekar, P. Wilcox, J. Golz, J. B. Kuang, A. Mathews, J. C. Law, T. Luong, H. C. Ngo, R. Freese, H. C. Hunter, E. Nelson, P. Parries, T. Kirihata, and S. S. Iyer, "A 1 MB cache subsystem prototype with 1.8 ns embedded DRAMs in 45 nm SOI CMOS," *IEEE J. Solid-State Circuits*, vol. 44, no. 1, pp. 1216–1226, Apr. 2009.
- [13] J. Barth, W. R. Reohr, P. Parries, G. Fredeman, J. Golz, S. E. Schuster, R. E. Matick, H. Hunter, C. C. Tanner, J. Harig, H. Kim, B. A. Khan, J. Griesemer, R. P. Havreluk, K. Yanagisawa, T. Kirihata, S. S. Iyer, and S. S. , "A 500 MHz random cycle, 1.5 ns latency, SOI embedded DRAM macro featuring a three transistor micro sense amplifier," *IEEE J. Solid-State Circuits*, vol. 43, no. 1, pp. 86–95, Jan. 2008.
- [14] T. Kimuta, K. Takeda, Y. Aimoto, N. Nakamura, T. Iwasaki, Y. Nakazawa, H. Toyoshima, M. Hamada, M. Togo, H. Nobusawa, and T. Tanigawa, "64 Mb 6.8 ns random ROW access DRAM macro for ASICs," in *IEEE ISSCC Dig. Tech. Papers*, Feb. 1999, pp. 416–417.



John Barth (M'04–SM'08) 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, Vermont. Mr. Barth is an IBM Distinguished Engineer currently developing SOI embedded DRAM macros for high performance microprocessor cache applications.

After completing his B.S. degree, he joined the IBM Development Laboratory in Essex Junction, VT, during which time he was involved the Design of a 16 mb DRAM product featuring embedded ECC and SRAM cache. His publication of "A 50 ns 16-Mb DRAM with 10 ns Data Rate and On-Chip ECC", received the IEEE JOURNAL OF SOLID-STATE CIRCUITS 1989–90 Best Paper Award. Following this, he was involved 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, standalone caches for P5 and P6 microprocessors and embedded caches for the Blue Gene/L super computer. Currently, he holds 43 US patents with 23 pending and has co-authored 21 IEEE papers. In 2002, he was co-recipient of the ISSCC Beatrice Award for Editorial Excellence for a 144 Mb DRAM targeted for standalone SRAM cache replacement. In 2007, he received ISSCC Best Paper Award for the three transistor micro sense amplifier, a derivative of which was used for the embedded DRAM on-chip cache for IBM's P7 microprocessor. He is an IEEE Senior Member, he served on ISSCC Memory Subcommittee from 2000–7 and currently server on the Technical Program Committee for the VLSI Circuits Symposium.



Don Plass is a Distinguished Engineer in the IBM Systems and Technology Group. He has been responsible for SRAM technology and designs for several generations of IBM server designs, including iSeries*, pSeries*, and zSeries* microprocessors, with a focus on the larger arrays. He joined IBM in 1978 at the Poughkeepsie facility, and in addition to CMOS SRAM, his research and development interests have included DRAM, gallium arsenide (GaAs), and BiCMOS. His recent accomplishments include bringing SOI eDRAM and dense SRAMs to

the product level for the 45 nm P7 microprocessor.



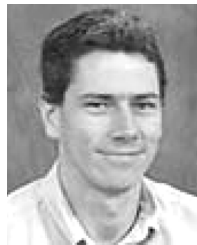
Erik Nelson started working at IBM shortly after earning a B.S.Ch.E. degree from Cornell University, Ithaca, N.Y. in 1982.

He works on embedded DRAM macro development and product qualification for IBM Systems and Technology Group, Burlington, Vermont. His first 10 years with IBM focused on bipolar process and product development including SRAM characterization. In 1993 he joined the IBM Seimens Toshiba DRAM Development Alliance and contributed to the creation of 64 Mb and 256 Mb DRAMs by delivering functional characterization results. In 2000 he applied his knowledge of process development and memory characterization to IBM's embedded DRAM project and helped make that offering an integral part of IBM's ASIC portfolio. In 2007 he focused on development of embedded SOI DRAM so as to enable its inclusion on the 45 nm processor chips for IBM servers.



Charlie Hwang received B.S. degree from National Taiwan University in 1987 and Ph.D. degree from Yale University in 1994.

He joined IBM Microelectronics in 1994 and worked on developing 0.25 μm DRAM technology. He joined DRAM design team in 1998 and worked on 0.18 μm 1 Gb DRAM design and a fast random cycle embedded DRAM design. He also worked on high speed serial link design from 2001 to 2002. In 2003, he joined IBM Server Group and worked on developing high-end processor designs for Power series and Mainframes. He is currently working on embedded DRAM designs for server processor chips.



Gregory Fredeman received the B.S. degree in electrical engineering technology from SUNY Utica/Rome, NY, in 1996, and 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 in 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.



Michael Sperling received the Bachelor of Science in electrical engineering from Carnegie Mellon University in 2002 and the Master of Science in electrical engineering from Carnegie Mellon University in 2003.

He joined IBM in 2003 and presently holds the position of analog circuit designer at the IBM Poughkeepsie development site. He has contributed to the design of phase locked loops, charge pumps, analog sensors and voltage regulators for the IBM family of microprocessors. He holds five patents with 11 pending and is currently working on adaptive voltage regulation for embedded DRAM and microprocessors.



Abraham Mathews received the B.S. degree in electrical engineering

He joined IBM in 1992 where he worked as a logic and circuit designer in the area of Graphics, ASICs and SoC designs. He also made several contributions to the design of dynamic latches, charge pumps and arrays. Prior to joining IBM, he worked as a designer of switch mode power supplies and micro controllers for a Sanyo collaboration company. He holds four patents. He is currently working on SOI embedded DRAM products for various server platforms.



Toshiaki Kirihiata (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 Tokyo Research Laboratory where he developed 22-ns 1-Mb and a 14-ns 4-Mb high-speed DRAM. In 1992 he joined the low-power DRAM design project at IBM Burlington Laboratory in Essex Junction, VT. In 1993 he served as Lead Engineer for the IBM Toshiba Seimens 256-Mb DRAM development at IBM, East Fishkill, NY. He joined IBM Research T. J. Watson Research Center in November 1996 and continued work on the 390- mm^2 1-Gb DDR and 512-Mb DDR2 SDRAMs as a Product Design Team Leader. In 2000 he transferred to the IBM

Semiconductor Research and Development Center IBM East Fishkill where he served as manager for the development of high performance embedded DRAM technology, during which he produced noteworthy designs including a 2.9 ns random cycle embedded DRAM in 2002, an 800 MHz embedded DRAM in 2004, a 500 MHz random cycle SOI embedded DRAM in 2007, and A 1 MB cache subsystem prototype development in 2008. He is currently a senior technical staff member for IBM Systems and Technology Group where he manages the embedded DRAM design department for high performance embedded DRAM and 3-D memory.

Mr. Kiriata presented papers at the ISSCC 1998, 1999, 2001, and 2004 conferences entitled “220 mm², 4 and 8 bank, 256 Mb SDRAM with single-sided stitched wordline architecture”, “390 mm² 16 bank 1 Gb DDR SDRAM with hybrid bitline architecture”, and “A 113 mm² 600 Mb/sec/pin DDR2 SDRAM with folded bitline architecture”, and “An 800 MHz embedded DRAM with a concurrent refresh mode”, respectively. He was a coauthor for ISSCC paper entitled “A 500 MHz Random Cycle, 1.5 ns Latency, SOI Embedded DRAM macro Featuring a Three Transistor Micro Sense Amplifier” which received the Lewis Winner outstanding paper award.



William R. Reohr received the Bachelor of Science in electrical engineering from the University of Virginia, Charlottesville, VA, in 1987 and the Master of Science 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. 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 Everspin, Inc. and others. Most recently, he has participated in the development of an embedded DRAM macro for Silicon-On-Insulator (SOI) 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. He holds over 42 patents with 26 pending covering various areas of VLSI logic, circuits, and technology.



Kavita Nair received the B.Sc. and M.Sc. degrees in electronics from the University of Pune, India. She later received the M.S. and Ph.D. in analog and mixed signal design from the University of Minnesota in 2001 and 2005, respectively.

She joined IBM in July 2005 where she worked on SRAM designs for high performance servers in Poughkeepsie, NY. She is currently working on SOI embedded DRAM designs for different server platforms.



Nianzheng Cao received the B.S. and M.S. degrees in mechanics from Peking University, Beijing, China, in 1982 and 1984, respectively. He received the Ph.D. degree in mechanical engineering from City University of New York in 1993.

Since he joined IBM Research in 1996, he has contributed to the development of high performance processors such as IBM Power4, Power5, Power6 and Power7 for various units in both core and cache area. He is a Research Staff Member currently working on high performance lower power VLSI circuit design in cache memories for IBM next generation microprocessors.