

MIT Open Access Articles

*An embedded energy monitoring circuit for a
128kbit SRAM with body-biased sense-amplifiers*

The MIT Faculty has made this article openly available. **Please share**
how this access benefits you. Your story matters.

Citation: Sinangil, Yildiz, and Anantha P. Chandrakasan. "An Embedded Energy Monitoring Circuit for a 128kbit SRAM with Body-Biased Sense-Amplifiers." 2012 IEEE Asian Solid State Circuits Conference (A-SSCC) (November 2012).

As Published: <http://dx.doi.org/10.1109/IPEC.2012.6522629>

Publisher: Institute of Electrical and Electronics Engineers (IEEE)

Persistent URL: <http://hdl.handle.net/1721.1/95786>

Version: Author's final manuscript: final author's manuscript post peer review, without publisher's formatting or copy editing

Terms of use: Creative Commons Attribution-Noncommercial-Share Alike



An Embedded Energy Monitoring Circuit for a 128kbit SRAM with Body-biased Sense-Amplifiers

Yildiz Sinangil and Anantha P. Chandrakasan

Massachusetts Institute of Technology, Cambridge, MA 02139 USA
e-mail: koken@mit.edu, anantha@mtl.mit.edu

Abstract— Embedded energy monitoring of critical system components can be used to enable better power management by capturing run time system conditions such as temperature and application load. In this work, an energy sensing circuit that provides digitally represented absolute energy per operation of a 128kbit SRAM is presented. Designed in a 65nm low-power CMOS process, SRAMs can operate down to 370 mV. Energy sensing circuit consumes $16.7\mu\text{W}$ during sensing at 1.2V (only 0.28% of SRAM active power at the same voltage). For improved performance, SRAMs utilize body-biased PMOS-input strong-arm type sense amplifiers that can achieve 45% tighter input offset distribution for only ~3.5% of total SRAM area overhead.

I. INTRODUCTION

With continuous reduction of transistor feature sizes, addressing the challenges of complex systems is becoming more difficult. Those complex systems usually need to meet multiple constraints such as a throughput requirement and a power budget. Furthermore, they often need to be optimized for conflicting goals. Additionally, those systems must perform efficiently under dynamically changing application loads and environmental changes. Self-aware systems can be used to achieve better system optimization by utilizing run time monitoring of critical system metrics and adapting to changes [1]. In this context, on-fly energy monitoring of various system components can be used for better allocation of system resources.

On-chip memories are responsible for a large portion of the total energy consumption of modern processors [2]. Moreover, technology scaling is fueling integration of larger on chip caches (50MB reported in [3]). Thus, an energy sensing circuit that is capable of generating a digital representation of absolute energy per operation (EOP) consumed by the caches would provide useful information to the system. However, energy sensing operation should be minimally intrusive to SRAM operation and need to introduce minimum overhead in terms of area and power.

On-chip caches not only dominate system power consumption but also limit system performance. Power consumption can be decreased by using voltage-scalability. However, scaling SRAM voltage degrades both SRAM performance and functionality. Thus, non-traditional bit-cell topologies and peripheral assist circuits are often used. Secondly, sense amplifiers are critical components of SRAMs

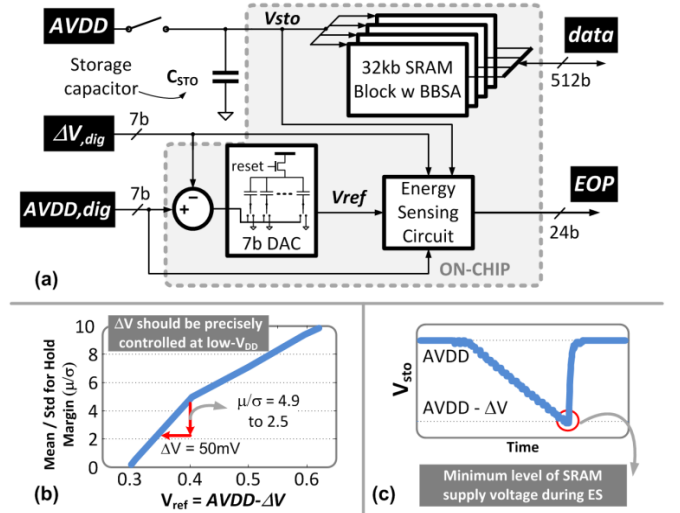


Figure 1. Chip block diagram. ΔV should be precisely controlled especially when the SRAM is operating at low-voltages.

as their input offset limit SRAM performance. Hence, reducing sense-amplifier input offset with minimum area penalty is significant to achieve better system performance.

In this work, we implemented an on-chip energy-sensing circuit for a 128kbit 8T bit-cell based SRAM design. This circuit provides a digital representation of absolute EOP of SRAM blocks. Furthermore, we propose an offset compensation technique using body-biasing to reduce input-referred offset of sense-amplifiers.

The rest of the paper is structured as follows: Section 2 presents on-chip energy sensing circuit architecture and its implementation. Then, in Section 3, 8T low-voltage SRAM design with offset compensated sense-amplifiers using body-biasing are discussed. Section 4 provides measurement results for the prototype test chip. Finally, section 5 concludes this paper.

II. ENERGY-SENSING CIRCUIT

The block diagram of our 128kbit SRAM with embedded energy sensing circuit is given in Fig. 1-a. When energy sensing is activated, an off-chip storage capacitor (C_{sto}) powers up SRAMs. Filtering capacitor of a power converter

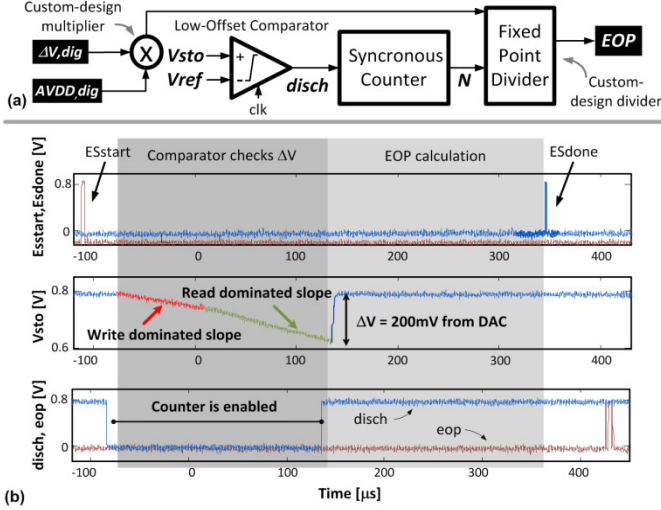


Figure 2. Energy sensing circuit (a) block diagram and (b) oscilloscope output of critical signals

can be used as C_{sto} . During the energy sensing period, the switch which is connected to power supply node (AVDD) stays OFF. When the voltage over C_{sto} (V_{sto}) drops by ΔV from V_1 to V_2 in N cycles (Fig. 1-c), EOP can be calculated as:

$$EOP = \frac{\frac{1}{2} \times C_{sto} \times (V_1^2 - V_2^2)}{N} \cong \frac{C_{sto} \times (V_1 \times \Delta V)}{N}$$

Previous work presents an energy metering circuit designed specifically for systems that use pulse frequency modulated switching regulators by observing the switching frequency variation of the regulator output [4]. Similarly, the authors in [5] introduced dynamically tracking the minimum energy point of an FIR by generating relative EOP numbers and without restricting the voltage drop of the supply voltage.

Absolute EOP numbers might be required in a system that operates with multiple energy sensing circuits. Furthermore, for SRAMs, ΔV needs to be precisely controlled in order not to degrade stability during energy sensing mainly at low voltages. As it can be seen from Fig. 1-b, a 50 mV voltage drop at $AVDD = 400$ mV changes hold margin μ/σ from 4.9 to 2.5 which would result into a drastic degradation of robustness due to the inability of many bit-cells to retain their states.

Fig. 2-a shows the energy sensing circuit block diagram. $AVDD_{dig}$ and ΔV_{dig} are 7-bit inputs. ΔV_{dig} can be set with 10 mV step sizes. It can be selected sufficiently large to capture an average EOP across many cycles or can be set small enough in order not to degrade SRAM operation especially at low voltages.

Multiplier and divider circuits are designed to be sequential to minimize the area overhead. The low-offset comparator is

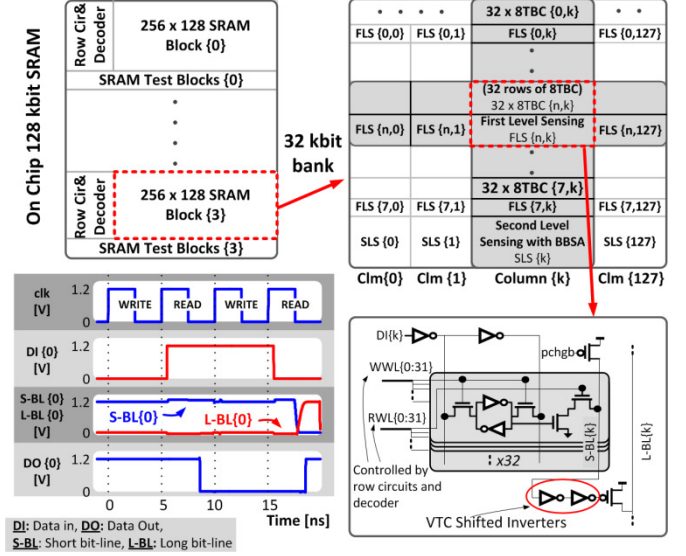


Figure 3. On-chip SRAM organization and simulated transient waveforms for the two-stage sensing scheme

carefully sized and can be digitally controlled to ensure an input offset voltage of less than 1mV.

Fig. 2-b demonstrates the operation of the energy sensing circuit by showing the oscilloscope outputs for its critical signals. When energy sensing is initiated by $ESstart$, custom-design 7-bit capacitive DAC generates $V_{ref} = AVDD - \Delta V$. Every cycle, a comparator checks if V_{sto} is larger than V_{ref} and comparator output ($disch$) keeps a 10-bit synchronous counter enabled during this period to generate N . When V_{sto} drops below V_{ref} , a custom-design multiplier and fixed-point divider are used to calculate absolute EOP provided that value of C_{sto} is known.

III. SRAM WITH OFFSET COMPENSATED BBSSA

Fig. 3 shows the SRAM architecture used in this work. Test chip features four SRAM blocks, each structured as 256 rows and 128 columns. An 8T bit-cell is used to enable low-voltage operation. Write word-line (WWL) is powered up by a higher voltage ($AVDDH$) to improve write-ability below 600mV. SRAMs are measured to be operational down to 370mV.

To improve performance, this work implements a two-stage sensing scheme. The first-level sensing circuits (FLS) are placed at every 32 rows of each block. FLS is performed with two static inverters. These inverters are designed to have shifted VTC characteristics by 100mV in order to result in faster tripping. Second-level sensing (SLS) utilizes body-biased PMOS-input strong-arm type sense-amplifiers (BBSSA) that are proposed in this work.

The ability of a sense-amplifier to sense small input differences is limited by its input-referred offset. The authors

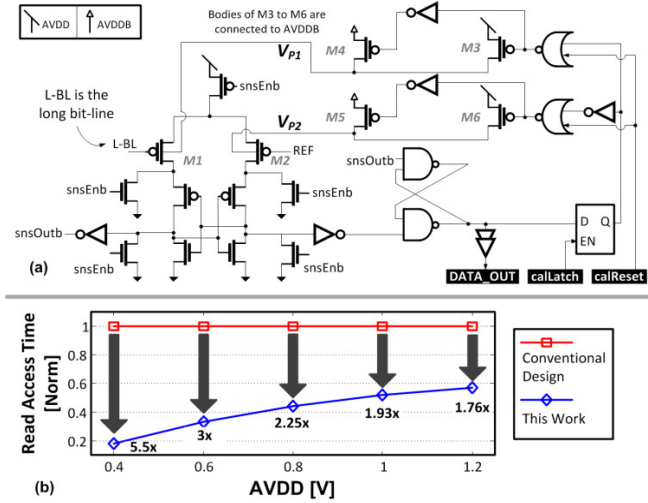


Figure 4 (a) BBSA circuit (b) Read access time improvement of this work over conventional small signal sensing.

in [6] and [7] utilize tunable SAs which select REF among multiple different voltage potentials to compensate for their offset. Authors in [8] use SA redundancy and select better performing sense amplifier during the calibration process. This scheme requires replicating the sense amplifiers as well as a calibration scheme at start-up.

In Fig. 4-a, the BBSA circuit implementation proposed in this work is shown. BBSA utilizes body biasing to reduce input-referred offset voltage of its input transistors (M_1 and M_2) by controlling their body voltages (V_{P1} and V_{P2}). M_3 to M_6 PMOS transistors drive V_{P1} and V_{P2} to either AVDD or AVDDB depending on the outcome of the calibration process.

A calibration process has to be performed once at the startup. The timing diagram of the calibration signals are shown in Fig. 5-a. At the beginning, calRst is kept high to make V_{P1} and V_{P2} of all BBSA circuits equal to AVDD. Then, snsEn is asserted while L-BL and REF are 0V. After sensing, DATA_OUT carries the information about positive or negative input referred offset of each BBSA. With calLatch signal assertion, this information is stored into a latch and body voltages are assigned to oppose the offsets. As shown in Fig. 5-b and Fig. 5-c, measured offset distribution of 512 sense-amplifiers can be made 45% tighter since the distribution drops from 190mV to 100mV by using AVDDB = AVDD + 150mV.

Fig. 4-b compares the statistical simulation results of read access time of this work and a conventional SRAM design. The latter has 256 cells per bit-line and uses small-signal strong-arm type SAs without offset compensation. This work brings up to 5.5x read time improvement due to two-level sensing and offset reduction using BBSA.

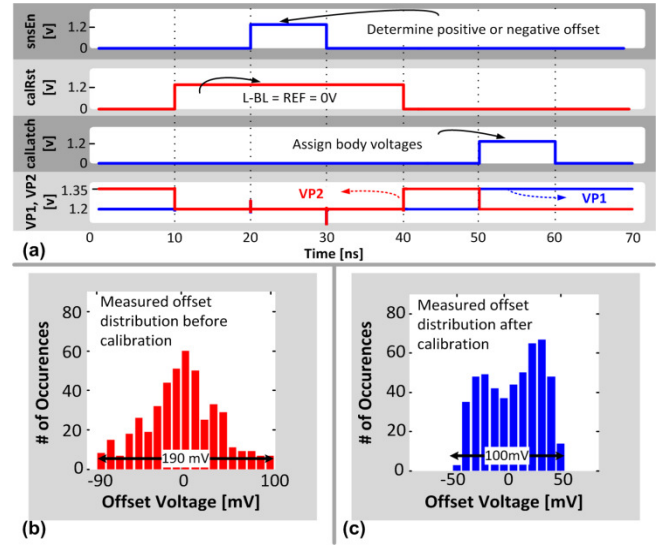


Figure 5 (a) Timing diagram of critical calibration signals of BBSA. (b) Measured BBSA input referred offset before calibration (c) Measured BBSA input referred offset after calibration. Input offset decreases by almost 1.9x with offset compensation using AVDDB = 150mV + AVDD.

By placing NMOSs of BBSA between different n-well potentials and carefully designing the layout of the circuit, total SRAM area overhead due to BBSA is kept below 3.5%.

IV. MEASUREMENT RESULTS OF THE TEST-CHIP

Body-biased sense amplifiers for offset-compensation and on-fly fine grained energy monitoring circuit are implemented in a 128kbit 65nm SRAM test-chip. Fig 6 shows the die photo and the specifications of the test chip. Die size is 1.4mm x 1.4mm. SRAMs are measured to achieve functionality down to 0.37V with no errors.

Fig. 7 shows the EOP numbers using three different methods:

- 1- Reading the digital EOP output of the energy monitoring circuit.
- 2- Measuring energy consumption values of the test-chip.
- 3- Using transistor-level extracted simulation.

As it can be seen from the figure, EOP values are reduced with voltage scaling. EOPs calculated by the energy sensing circuit are accurate and always within 9.95% of the measured values.

Fig. 8 shows measured EOP calculated by the on chip energy sensing circuit across different temperature and by varying total read and write operation ratios. Read accesses result in larger energy consumption and the ratio of total read and write operations can change significantly from one application to the other. Moreover, higher temperature results in larger leakage and larger overall EOP. Depending on these conditions, EOP can change by more than 2x, so systems can greatly benefit from run time energy sensing calculations.

Process	65nm low-power CMOS
Chip Size	1.4 mm x 1.4 mm
Organization	256 x 128 (4 blocks)
DAC resolution	7bits (10mV steps)
Supply Voltages	DAC: 1.2 V
	SRAM:
	AVDD: 0.37 – 1.2 V
	AVDDH: 0.6 – 1.2 V
	AVDDB: AVDD+150mV
Area	REF: 100mV
	ES: 0.026 mm ²
	DAC: 0.059 mm ²
	SRAM: 0.524 mm ²
EOP (data = 512bits)	108 pJ at 1.2V
	49.2 pJ at 0.8V
	21.2 pJ at 0.5V

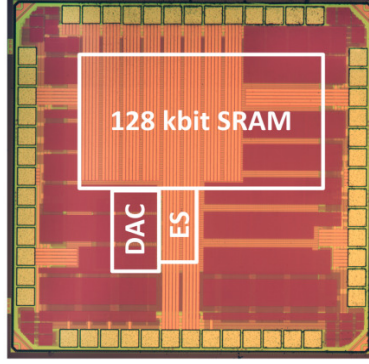


Figure. 6 Chip specifications and die photograph of the test-chip fabricated in 65nm low-power CMOS process.

V. CONCLUSION

In this paper, a 128kbit 8T SRAM design that can work down to 0.37V is presented. SRAMs utilize body-biased sense-amplifiers for offset-compensation to achieve 1.9× tighter input offset distribution. In addition, an energy metering circuit that can measure run time absolute energy per operation numbers of the SRAM blocks is shown. Energy sensing circuit and DAC introduce 16.3% area compared to total SRAM area but this overhead will be smaller for larger SRAM sizes. Active power consumed by energy sensing circuit and DAC is 0.28% of the SRAM active power. The small area and power overhead of the energy sensing circuit makes it possible to be used for system-level power management.

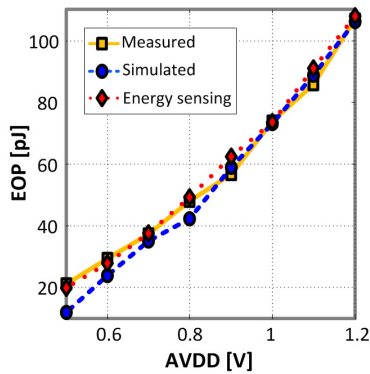


Figure. 7 EOP vs. AVDD (measured using power sources, simulated and measured by energy monitoring circuit). EOP calculated with on-chip energy sensing circuit is within 9.95% of the measured value across different voltages.

Measured Energy per Operation at VDD=800mV, f=1MHz

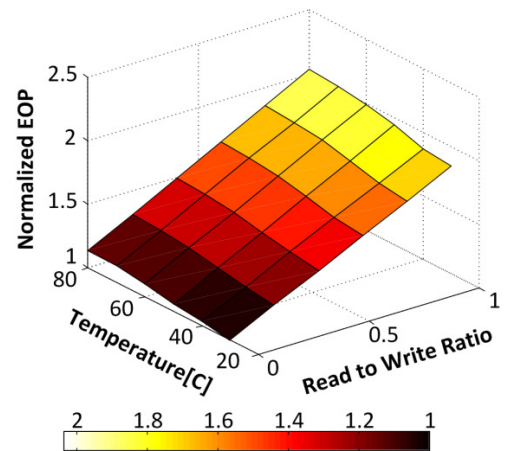


Figure. 8 Measured EOP under different read operation to write operation ratios and temperatures. EOP changes by more than 2x across different conditions.

ACKNOWLEDGMENT

This work was funded by the U.S. Government under the DARPA UHPC program. The views and conclusions contained herein are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of the U.S.

The authors thank TSMC University Shuttle Program for chip fabrication.

REFERENCES

- [1] H. Hoffmann, J. Holt, G. Kurian, E. Lau, M. Maggio, J. E. Miller, S. M. Neuman, M. Sinangil, Y. Sinangil, A. Agarwal, A. P. Chandrakasan, S. Devadas, "Self-aware Computing in the Angstrom Processor", Proceedings of the 49th Design Automation Conference (DAC), June 2012
- [2] N. Ickes, Y. Sinangil, F. Pappalardo, E. Guidetti, A. P. Chandrakasan, "A 10 pJ/cycle ultra-low-voltage 32-bit microprocessor system-on-chip," *ESSCIRC*, pp.159-162, September 2011.
- [3] R. J. Riedlinger, R. Bhatia, L. Biro, B. Bowhill, E. Fetzer, P. Gronowski, T. Grutkowski, "A 32nm 3.1 billion transistor 12-wide-issue Itanium® processor for mission-critical servers," *ISSCC*, pp.84-86, February 2011.
- [4] P. Dutta, M. Feldmeier, J. Paradiso, D. Culler, "Energy Metering for Free: Augmenting Switching Regulators for Real-Time Monitoring," *IPSN*, pp.283-294, April 2008.
- [5] Y. K. Ramadass, A. P. Chandrakasan, "Minimum Energy Tracking Loop With Embedded DCDC Converter Enabling Ultra-Low-Voltage Operation Down to 250 mV in 65 nm CMOS," *JSSCC*, vol.43, no.1, pp.256-265, January 2008.
- [6] S. Cosemans, W. Dehaene, F. Catthoor, "A 3.6pJ/access 480MHz, 128Kbit on-Chip SRAM with 850MHz boost mode in 90nm CMOS with tunable sense amplifiers to cope with variability," *ESSCIRC*, pp.278-281, September 2008.
- [7] M. E. Sinangil, N. Verma, A. P. Chandrakasan, "A 45nm 0.5V 8T column-interleaved SRAM with on-chip reference selection loop for sense-amplifier," *A-SSCC*, pp.225-228, November 2009
- [8] N. Verma, A. P. Chandrakasan, "A 256 kb 65 nm 8T Subthreshold SRAM Employing Sense-Amplifier Redundancy," *JSSCC*, vol.43, no.1, pp.141-149, January 2008