A Fully-Functional 90nm 8Mb STT MRAM Demonstrator Featuring Trimmed, Reference Cell-Based Sensing

John DeBrosse[1], Thomas Maffitt[1], Yutaka Nakamura[1], Guenole Jan[2], Po-Kang Wang[2]

[1]IBM Research Division, Essex Junction, VT, USA, [2]TDK, Milpitas, CA, USA

***Abstract* — Spin Transfer Torque Magnetoresistive RAM (STT MRAM) has uniquely attractive write performance and endurance characteristics. Nonetheless, little STT MRAM circuit hardware data has been published [1-4]. This paper describes a fully-functional 90nm 8Mb STT MRAM, identifies and describes solutions to the primary circuit challenges, and includes considerable circuit hardware data.**

***Index Terms* — STT, MRAM, sense, write, ECC.**

I. Introduction

Due to the truly electronic nature of its write mechanism, Spin Transfer Torque Magnetoresistive RAM (STT MRAM) is the only emerging memory technology with the potential write performance and write endurance required of cache and main memory applications. SRAM and DRAM face the same fundamental problem, which is that FET device off current scales in the wrong direction, resulting in ever higher standby current and refresh requirements. STT MRAM has the potential of zero array standby current, the density of a single FET cell, and operates at low voltage.

The fundamental device of MRAM is the Magnetic Tunnel Junction (MTJ), including a fixed or reference ferromagnetic layer, tunnel barrier and free ferromagnetic layer. The free magnetic moment is engineered to have two stable states, parallel (0) and anti-parallel (1) to the reference moment. The change in tunneling resistance between states is characterized by the parameter MagnetoResistance (MR), defined by R1=R0*(1+MR) and about 120% for this project, corresponding to a 2.2x resistance ratio between states.

Unlike earlier field-written MRAMs, the STT MRAM MTJ is written by applying a tunnel bias much larger than that during read [5]. Sufficient electrons tunneling from the reference to free layer write the parallel or 0 state. The opposite polarity writes the 1 state. Thermal energy perturbs the free moment, crucially initiating the precessional write process but causing disturbs during storage and read [6]. An STT MRAM design must therefore accommodate a certain level of soft errors, occurring during write, storage and read.

The primary circuit challenge of STT MRAM is sensing, done at low voltage to minimize read disturbs and accurate to about 1% to maximize read margin and yield. Secondly, writing requires a higher, but well-controlled MTJ bias. Thirdly, the design must correct soft errors and prevent them from accumulating. This paper presents solutions to these issues, and hardware data to demonstrate their effectiveness.

II. Sensing

The Sense Amplifier (SA) is illustrated in Fig. 1 and is an update of the conventional field MRAM SA [7-8]. The NFET source follower devices (N0-N2) force the data and two reference Bit Lines (BLs) to the target read voltage (Vbleq+Vread) while the Source Lines (SLs) are held at the BL standby level (Vbleq, 500mV). Hence, Vread (135mV) is applied from BL to SL but less than 100mV appears on the MTJ due to series losses. The PFET current mirror load (P0-P2) generates a voltage corresponding to the data cell current minus the average of the two reference cell currents (one of either state), which is sensed by the comparator.

The most significant change to the earlier field MRAM SA is the addition of the local feedback amplifier controlling the gate node of the NFET source followers. In field MRAM, this node is typically a global DC reference, and the forced BL voltage is thus dependent on the long-range variation of N0-N2. STT MRAM requires a much lower read voltage to minimize read disturbs, making this variation unacceptable. By introducing the local feedback amplifier, the forced BL voltage and the current comparison are dependent upon local matching only. Also for disturb reasons, transient overshoot of the loop must be minimized.

Despite the use of large, carefully-matched FETs and a 2.5V SA supply, 4 bit trimming of each SA is required to limit SA to SA offset to roughly 1% of the reference current. In decreasing order, this offset arises from FET mismatch in P0-P2, N0-N2 and the comparator. Trimming is achieved by enabling small PFETs in parallel with P0-P2.

III. WRITING AND ARRAY-LEVEL ARCHITECTURE

The target MTJ write voltage is approximately +/-600mV and must be tightly controlled to avoid underwrite

(insufficient to write cell) and overwrite (contributing to tunnel barrier breakdown). Write is most easily understood as a voltage divider between the MTJ resistance and the various series resistances, dominated by the array FET and the BL and SL wiring. The following techniques are used to minimize and control these series resistances.

The BL and SL are driven from opposite ends of the array, so that the sum of their resistances to the cell is relatively insensitive to Word Line (WL) address. Just above the MTJ on M4, the BL and its adjacent space occupy the full cell width. Below the MTJ, the SL must share the cell width with the metal Landing Pad (LP) connecting the MTJ down to the FET, and its resistance is thus much higher. Hence, the SL is wired on both M1 and M2 while M3 straps the WL. The design employs robust Column Decoders (CDs), Write Drivers (WDs), and write voltage (Vwr0, Vwr1) regulators and wiring to minimize loss and pattern dependence. A block diagram of one 128Kb segment or SA domain and an array cross-section are shown in Figs. 2 and 3.

A double gate finger, 1.2V NFET is used to minimize the array FET resistance in a $50F^2$ cell (F=90nm). As illustrated in Fig. 4, holding the BLs at Vbleq in standby allows a higher value of WL voltage (Vpp) than if the BLs were grounded in standby. Only the selected array FET is exposed to the full Vpp and only during the Write 0 pulse, a 1000x lower duty.

IV. Error Correction and Chip-Level Operation

The interface is like a 3.3V x16 asynchronous SRAM. All cycles, Write or Read, include the following internal steps:

1.) Sense 80b, 1 per 128Kb segment or SA domain.

2.) Data redundancy reduces the 80b to 78b.

3.) ECC decoder corrects up to 2 errors in the 78b ECC word (BCH DEC code, 64 data + 14 check bits)

4.) For a Read cycle, multiplex out 16b of the corrected 64b to the DQ. Page mode Read of the corrected 64b is supported.

5.) For a Write cycle, multiplex in 16b from the DQ.

6.) Generate 14 new ECC check bits.

7.) Data redundancy expands the new 78b to the 80 segments of the chip.

8.) For each of the 78 active segments, write the selected cell if the bit has changed OR if the old bit was in error. If the old bit was in error, also concurrently refresh the selected reference cells of that segment.

This method prevents read disturbs, including reference cell disturbs, from accumulating in a 78 bit ECC word over more than one cycle. Write endurance and power are optimized by writing only when necessary. Somewhat like DRAM, a large page is sensed (allowing efficient on-chip ECC), followed by page mode IO and a delay after last write (allowing check bit generation and write).

The chip includes digitally-trimmed on-chip timing and voltage generation, and conventional row and data redundancy for hard fails, the ECC being intended for soft errors only. The redundancy and trimming "fuse" data is stored in an ECC-protected, twin cell extension of the non-volatile STT MRAM array. A chip micrograph is shown in Fig. 5, and Table I lists the chip features.

V. Hardware Results

For demonstration, the following data was collected with ECC disabled, though the ECC engine performs as expected. SA to SA offset is represented by the spread in the individual curves in the "Pre-Trim" plot of Fig. 6. The improvement of read margin due to SA trimming is illustrated in the "Post-Trim" plot of Fig. 6. Sensing and write performance and are demonstrated in Figs. 7 and 8. Additional technology-related data from this design, including thermal stability, temperature dependence of write, data retention and Bit Error Rate (BER), are available in [9-10].

VI. Conclusion

While much progress remains to be made in STT MRAM technology and circuits, this chip successfully demonstrated several design techniques for mitigating the challenges inherent to this exciting technology.

Acknowledgement

The authors wish to thank the many IBM and TDK contributors to this work, including Jack Morrish, Mark Drapa, Jerry Heath, Dennis Willmott, Richard Jordan, Chris Long, Luis Lastras, William Gallagher, Daniel Worledge, Karl Yang, Robert Chen, Son Le and Toshio Sunaga.

References

[1] M. Jefremow et al, "Time-differential Sense Amplifier for sub-80mV Bitline Voltage Embedded STT-MRAM in 40nm," *ISSCC*, 2013, pp. 216-217.

[2] T. Andre et al, "ST-MRAM Fundamentals, Challenges, and Applications," *CICC*, 2013, pp. 1-8.

[3] C. Kim et al, "A Covalent-Bonded Cross-Coupled Current-Mode Sense Amplifier for STT-MRAM with 1T1MTJ Common Source-Line Structure Array," *ISSCC*, 2015, pp. 134-135.

[4] H. Noguchi et al, "A 3.3ns-Access-Time 71.2μW/MHz 1Mb Embedded STT-MRAM Using Physically Eliminated Read-Disturb Scheme and Normally-Off Memory Architecture," *ISSCC*, 2015, pp. 136-137.
[5] J. C. Slonczewski, "Current-driven Excitation of Magnetic Multilayers," *J. Magn. Magn. Mater.*, vol. 159, pp. L1-L7, 1996.
[6] W. H. Butler et al, "Switching Distributions for Perpendicular Spin-Torque Devices Within the Macrospin Approximation," *IEEE Trans. Magn.*, vol. 48, pp. 4684-4700, 2012.
[7] J. Nahas et al, "A 4Mb 0.18μm 1T1MTJ Toggle MRAM Memory," *ISSCC*, 2004, pp. 44-45.
[8] J. DeBrosse et al, "A 16Mb MRAM Featuring Bootstrapped Write Drivers," *Symp. VLSI Circuits*, 2004, pp. 454-457.
[9] G. Jan et al, "Demonstration of Fully Functional 8Mb Perpendicular STT-MRAM Chips with sub-5ns Writing for Non-volatile Embedded Memories," *Symp. VLSI Tech.*, 2014. pp. 50-51.
[10] S. H. Kang, "Embedded STT-MRAM for Energy-efficient and Cost-effective Mobile Systems," *Symp. VLSI Tech.*, 2014, pp. 44-45.

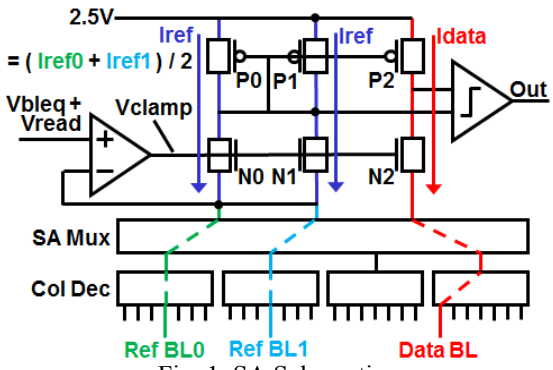

Fig. 1. SA Schematic.

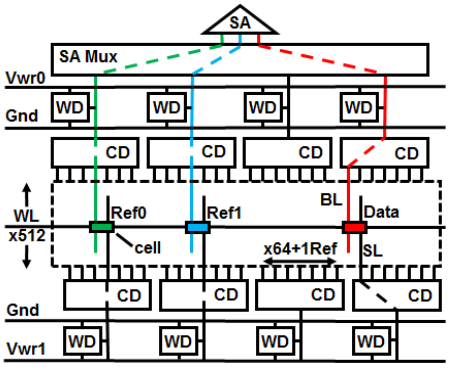

Fig. 2. 128Kb Segment or SA Domain.

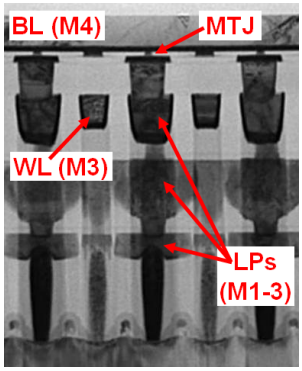

Fig. 3. Array Cross-Section.

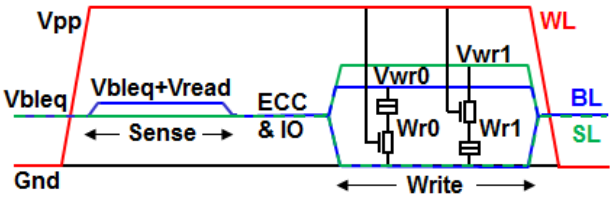

Fig. 4. Array Timing Diagram.

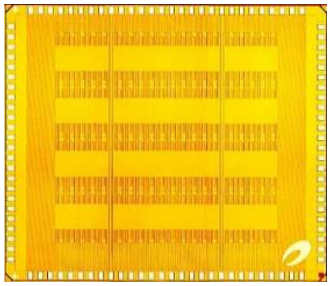

Fig. 5. Chip Micrograph.

Process	4Cu 90nm CMOS
Devices	1.2V - Array & Periph Logic 2.5V - SA, Row & Col Ckts 3.3V - IO & Power
MTJ	R0 ~ 2Kohm, MR ~ 120% Rd < 100mV, Wr ~ +/-600mV
Cell	50F², MTJ under M4
Interface	3.3V x16 Asynch SRAM-like
Cycle	70ns incl Sense, ECC & Wr
Temp	- 40C to 125C

Table I. Chip Features.

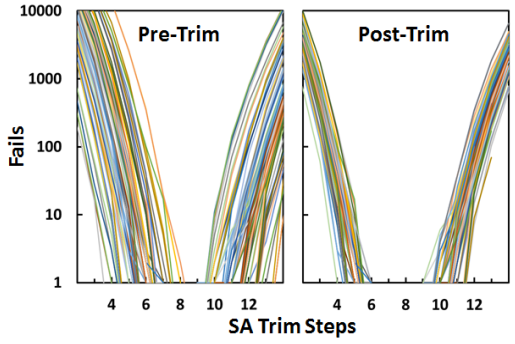

Fig. 6. SA Trim Shmoos (1 curve per SA).

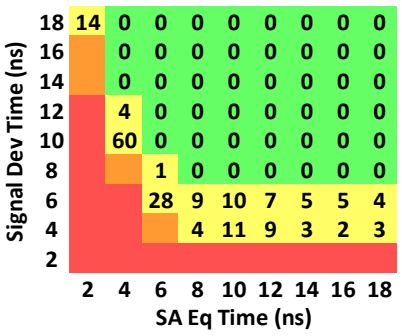

Fig. 7. 8Mb SA Timing Shmoo.

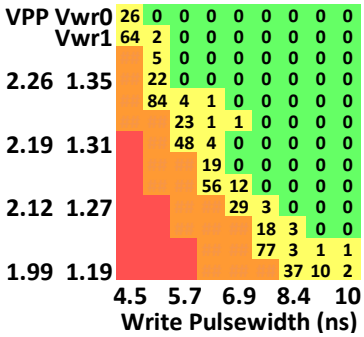

Fig. 8. 8Mb Write Shmoo.