Enabling a Reliable STT-MRAM Main Memory Simulation

Kazi Asifuzzaman*†, Rommel Sánchez Verdejo*†, Petar Radojković*
*Barcelona Supercomputing Center, Barcelona, Spain
†Universitat Politècnica de Catalunya, Barcelona, Spain
E-mail: {kazi.asifuzzaman, rommel.sanchez, petar.radojkovic}@bsc.es

Keywords*—*STT-MRAM, Main memory, High-performance computing.

I. EXTENDED ABSTRACT

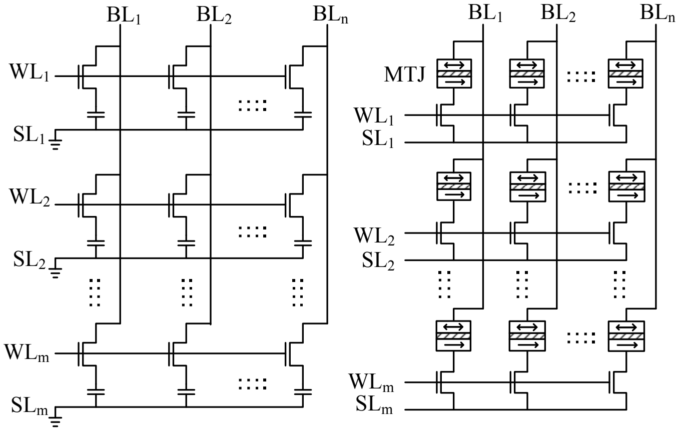
Memory systems are major contributors to the deployment and operational costs of large-scale HPC clusters [1][2], as well as one of the most important design parameters that significantly affect system performance. In addition, scaling of the DRAM technology and expanding the main memory capacity increases the probability of DRAM errors that have already become a common source of system failures in the field. It is questionable whether mature DRAM technology will meet the needs of next-generation main memory systems. So, significant effort is invested in research and development of novel memory technologies. A potential candidate for replacing DRAM is Spin Transfer Torque Magnetic Random Access Memory (STT-MRAM).

The main objective of this work is to understand and publish detailed STT-MRAM main memory timing parameters enabling a reliable system level simulation of the novel memory technology. The approach that we present converged through research cooperation with Everspin technologies Inc., one of the leading MRAM manufacturers, and it provides reliable STT-MRAM timing parameters while releasing no confidential information about any commercial products.

A. *STT-MRAM*

The storage and programmability of STT-MRAM revolve around a Magnetic Tunneling Junction (MTJ), see Figure 1(b). An MTJ is constituted by a thin tunneling dielectric being sandwiched between two ferro-magnetic layers. One of the layers has a fixed magnetization while the other layer's magnetization can be flipped. If both of the magnetic layers have the same polarity, the MTJ exerts low resistance therefore representing a logical "0"; in case of opposite polarity of the magnetic layers, the MTJ has a high resistance and represents a logical "1". In order to read a value stored in an MTJ, a low current is applied to it. The current senses the MTJ's resistance state in order to determine the data stored in it. Likewise, a new value can be written to the MTJ through flipping the polarity of its free magnetic layer by passing a large amount of current through it [3].

STT-MRAM main memory timing parameters has neither been standardized nor been released by any industry. This is perhaps due to the perpetual evaluation of the STT-MRAM technology that is constantly changing over a short duration of time. Memory manufacturers, who are developing STT-MRAM are judiciously not revealing these parameters ahead of time; so, at this point, we have to accept that there is no reliable information on how these timing parameters will change for the upcoming STT-MRAM devices.



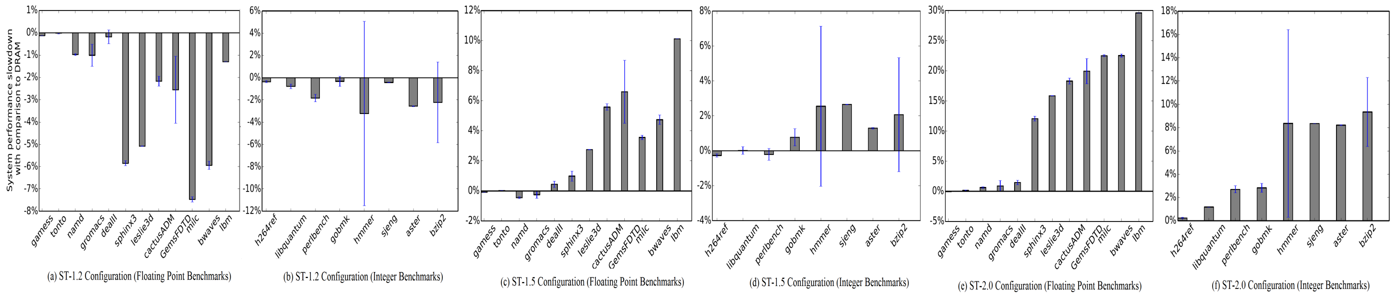
(a) DRAM cell array (b) STT-MRAM cell array

Fig. 1. STT-MRAM cell and cell-array

Industrial patents [4][5][6] suggest STT-MRAM manufacturers are adopting STT-MRAM technology in to DDRx interface and protocols in order to enable a seamless integration into rest of the system. STT-MRAM memory devices are DDRx compatible, with the same or very similar organization and CPU interface, as the conventional DRAM. Also, both, DRAM and STT-MRAM main memory devices use row buffer as an interface between the cell-arrays and the memory bus. Since the circuitry beyond the row buffer for DRAM and STT-

TABLE I. DRAM AND STT-MRAM PARAMETERS ASSOCIATED WITH ROW OPERATION (DDR3-1600 CYCLES)

Timing Parameters	Description	DRAM	ST-1.2	ST-1.5	ST-2.0
tRCD	Row to column command delay	11	14	17	22
tRP	Row precharge	11	14	17	22
tFAW	Four row activation window	24	29	36	48
tRRD	Row activation to Row activation delay	5	6	8	10
tRFC	Refresh cycle time	208	1	1	1



(a) ST-1.2 Configuration (Floating Point Benchmarks)

(b) ST-1.2 Configuration (Integer Benchmarks)

(c) ST-1.5 Configuration (Floating Point Benchmarks)

(d) ST-1.5 Configuration (Integer Benchmarks)

(e) ST-2.0 Configuration (Floating Point Benchmarks)

(f) ST-2.0 Configuration (Integer Benchmarks)

Fig. 2. STT-MRAM slowdown with respect to DRAM main memory

MRAM is essentially the same — once the data is in the row buffer, STT-MRAM timing parameters for the consequent operations are the same as DRAM. Therefore, the values of all the timing parameters that are not associated with row operations do not change from DRAM to STT-MRAM.

The only fundamental difference in STT-MRAM and DRAM main memory is their storage cell technology (see Figure 1), MTJ and capacitor, respectively. Due to the difference in the cell access mechanism of these two memory technologies, the timing parameters associated with STT-MRAM row operations would deviate from DRAM, and there is no reliable information on how these timing parameters will change for the upcoming STT-MRAM devices. Therefore, a sensitivity analysis is performed on timing parameters that would deviate from DRAM. In this study, we selected three set of timings naming ST-1.2, ST-1.5 and ST-2.0 with deviations of 1.2x, 1.5x and 2x from respective DRAM timing parameters as summarized in Table I. All the timing parameters that are *not* listed in this table will be same as DRAM. The presented methodology converged through our research cooperation with Everspin Technologies Inc.

B. Experimental Environment

STT-MRAM main memory was evaluated on a set of eight integer and twelve floating point benchmarks from the SPEC CPU 2006 suite [7]. We used ZSim [8] system simulator for the experiments. The simulated hardware platform comprises a detailed model of Sandy Bridge-EP E5-2670 cache hierarchy [9]. This Sandy Bridge E class processor has eight cores, dedicated L1 instruction and data cache of 32 KB each, dedicated L2 cache of 256 KB and a shared L3 cache of 20 MB. Both DRAM and STT-MRAM main memory is simulated with DRAMSim2 [10].

C. Results

Figure 2 shows overall system performance impact of STT-MRAM configurations on SPEC integer benchmark. The vertical bars represent system performance deviation from DRAM for the corresponding benchmarks listed at *X* axis. Both Floating point (Figure 2(a)) and Integer (Figure 2(b)) benchmarks with ST-1.2 configuration experience a speedup with the STT-MRAM main memory. This is due to the operation sequence of STT-MRAM, which is different from DRAM. Unlike DRAM, STT-MRAM has a non-destructive read which does not have to write-back; meaning it can issue precharge command sooner [11]. For ST-1.5 configuration (Figure 2(c)(d)) system performance degradation ranges from -0.2% (h264ref) to 10.1% (lbm). For the most pessimistic configuration ST-2.0 (Figure 2(e)(f)), slowdown ranges between 0.2% (h264ref) and 29.6% (lbm).

D. Conclusion

In this study, we publish reliable timing parameters of STT-MRAM main memory and measure its performance degradation w.r.t DRAM. We believe this study will enable researchers to perform a reliable STT-MRAM main memory simulation.

II. ACKNOWLEDGMENT

This work bas been published in proceedings of the International Symposium on Memory Systems (MEMSYS), 2017 [12].

REFERENCES

[1] P. Kogge *et al.*, "ExaScale Computing Study: Technology Challenges in Achieving Exascale Systems," DARPA, Sep. 2008.

[2] A. Sodani, "Race to Exascale: Opportunities and Challenges," Keynote Presentation at the 44th Annual IEEE/ACM International Symposium on Microarchitecture (MICRO), Dec. 2011.

[3] Y. Xie, "Modeling, Architecture, and Applications for Emerging Memory Technologies," *IEEE Design Test of Computers*, 2011.

[4] H. Kim *et al.*, "Magneto-resistive memory device including source line voltage generator," 2013.

[5] H. Oh, "Resistive Memory Device, System Including the Same and Method of Reading Data in the Same," 2014.

[6] C. Kim *et al.*, "Magnetic Random Access Memory," 2013.

[7] J. L. Henning, "SPEC CPU2006 Benchmark Descriptions," *SIGARCH Comput. Archit. News*, 2006.

[8] D. Sanchez and C. Kozyrakis, "Zsim: Fast and accurate microarchitectural simulation of thousand-core systems," in *Proceedings of the 40th Annual International Symposium on Computer Architecture*, ser. ISCA, 2013.

[9] Intel, "Intel 64 and IA-32 Architectures Optimization Reference Manual," http://www.intel.com/content/www/us/en/architecture-and-technology/64-ia-32-architectures-optimization-manual.html.

[10] P. Rosenfeld *et al.*, "DRAMSim2: A Cycle Accurate Memory System Simulator," *IEEE Computer Architecture Letters*, 2011.

[11] J. Wang *et al.*, "Enabling High-performance LPDDRx-compatible MRAM," ser. ISLPED, 2014.

[12] K. Asifuzzaman *et al.*, "Enabling a reliable stt-mram main memory simulation," in *Proceedings of the International Symposium on Memory Systems*, ser. MEMSYS '17.

Kazi Asifuzzaman received his BSc degree in Computer Engineering from North South University (NSU), Bangladesh in 2008. The following year, he worked at the IT department of Shimizu Densetsu Kogyo Co. Ltd (SEAVAC) in Japan. He completed his MSc degree in Electronic Design from Lund University, Sweden in 2013. Since 2014, he has been with the Memory Systems group of Barcelona Supercomputing Center (BSC) as well as a PhD student at the department of computer architecture of Universitat Politècnica de Catalunya (UPC), Spain.