

Performance Impact of a Slower Main Memory: A case study of STT-MRAM in HPC

Kazi Asifuzzaman^{*†}, Milan Pavlovic^{*†}, Milan Radulovic^{*†}, David Zaragoza^{*†},
Ohseong Kwon[‡], Kyung-Chang Ryoo[‡] and Petar Radojković^{*}

^{*}Barcelona Supercomputing Center, Barcelona, Spain

[†]Universitat Politècnica de Catalunya, Barcelona, Spain

[‡]Samsung Electronics Co., Ltd., Memory Planning Group, Seoul, South Korea

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][3], 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).

In this paper, we explore whether STT-MRAM is a good candidate for HPC main memory systems. To that end, we simulate and analyze performance of production HPC applications running on large-scale clusters with STT-MRAM main memory, and compare the results with DRAM. Our results show that, despite being 20% slower than DRAM at the device level, STT-MRAM main memory delivers performance comparable to DRAM — for most of the applications under study, STT-MRAM introduces a slowdown below 1%.

A. STT-MRAM

The storage and programmability of STT-MRAM revolve around a Magnetic Tunneling Junction (MTJ). 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. As Figure 1(a) and (b) depict, 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 [4].

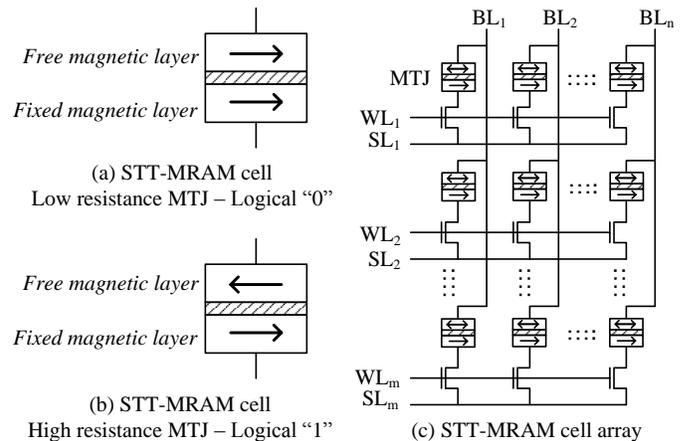


Fig. 1. STT-MRAM cell and cell-array

B. Experimental Environment

We evaluated STT-MRAM main memory on HPC applications included in the Unified European Application Benchmark Suite (UEABS) [8]. UEABS is the latest benchmark suite distributed by Partnership for Advanced Computing in Europe (PRACE) and it represents a good coverage of production HPC applications running on European Tier-0 and Tier-1 HPC systems. All UEABS applications are parallelized using Message Passing Interface (MPI) and they are regularly executed on hundreds or thousands of processing cores. UEABS also includes input data-sets that characterize production use of the applications. In our experiments, we executed UEABS applications with Test Case A, input data-set that is designed to run on Tier-1 sized systems, up to thousands of x86 cores. Table I summarizes applications used in the study.

We collected traces of UEABS applications running on MareNostrum supercomputer [10]. MareNostrum contains 3056 compute nodes (servers) connected with the Infiniband network. Each node contains two Intel Sandy Bridge-EP E5-2670 sockets that comprise eight cores operating at 2.6 GHz. Although Sandy Bridge processors support hyper-threading at core level, this feature is disabled, as in most of the HPC systems. Sandy Bridge processors are connected to main memory through four channels and each channel is connected to a single 4GB DDR3-1600 DIMM.

TABLE I. UEABS APPLICATIONS USED IN THE STUDY

Application	Scientific area	Selected number of cores
ALYA	Computational mechanics	1024
BQCD	Particle physics	1024
CP2K	Computational chemistry	1024
GADGET	Astronomy and cosmology	1024
GENE	Plasma physics	1024
GROMACS	Computational chemistry	1024
NEMO	Ocean modeling	1024
Quantum Espresso	Computational chemistry	256

We analyze main memory system in which the DRAM modules are simply replaced with STT-MRAM modules with same capacity and organization (4×4 GB DIMMs per socket), without requiring any modification in the rest of the system architecture [7]. Memory controller and channel latencies were set to 30 ns while the average DRAM device latency was simulated with 15 ns. These parameters correspond to average latencies of HPC applications running on real HPC systems [10]. Memory planning group of Samsung Electronics Co., Ltd. estimates that the high-density STT-MRAM main memory devices will be approximately 20% slower than conventionally used DRAM, so the average STT-MRAM access time was simulated with 18 ns¹. Like Suresh *et al.* [11] we used symmetrical read/write for STT-MRAM operation, which is in compliance with several scientific studies and products released recently [12][13][14].

C. Results

The performance comparison between STT-MRAM and DRAM main memory is presented in Figure 2. For each application, different bars correspond to different simulated CPI of 0.5, 1 and 2. The solid bars represent the average STT-MRAM slowdown over DRAM, and the error bars show the standard deviation for various application processes and main-loop iterations. For ALYA and GROMACS, we detect almost no performance difference between STT-MRAM and DRAM main memory systems. Four out of the remaining six applications, CP2K, GADGET, QE and BQCD, experience less than 1% slowdown. Finally, GENE slowdown ranges between 1.5% and 2%, while the slowdown of NEMO is around 2.5%. Overall, the impact of higher STT-MRAM latency on the HPC application performance is very low — for six out of eight

¹Micro-architecture and detailed timings of Samsung high-density STT-MRAM main memory devices can not be disclosed due to confidentiality issues. Samsung memory planning group estimates that capacity of high-density STT-MRAM devices will be comparable with DRAM modules.

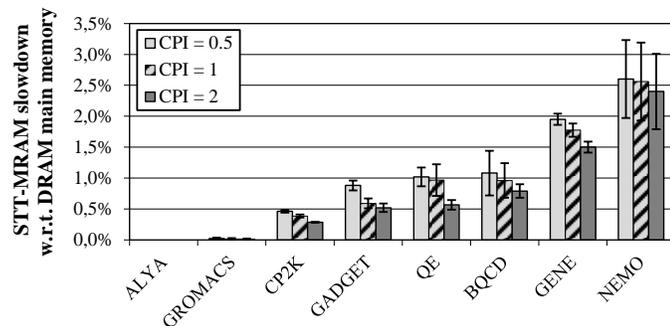


Fig. 2. STT-MRAM slowdown with respect to DRAM main memory

applications the slowdown is below 1% and it is only 2.6% in the worst case.

D. Conclusion

We simulate and analyze performance of production HPC applications running on large-scale clusters with STT-MRAM main memory and compare the results with DRAM. Our results reveal that, although being 20% slower than DRAM at the device level, STT-MRAM main memory induces a performance degradation below 1% for most of the HPC applications under experiment.

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] R. Stevens *et al.*, “A Decadal DOE Plan for Providing Exascale Applications and Technologies for DOE Mission Needs,” Presentation at Advanced Simulation and Computing Principal Investigators Meeting, Mar. 2010.
- [4] Y. Xie, “Modeling, Architecture, and Applications for Emerging Memory Technologies,” *IEEE Design Test of Computers*, 2011.
- [5] C. Kim *et al.*, “Magnetic Random Access Memory,” 2013.
- [6] H. Kim *et al.*, “Magneto-resistive memory device including source line voltage generator,” 2013.
- [7] H. Oh, “Resistive Memory Device, System Including the Same and Method of Reading Data in the Same,” 2014.
- [8] *Unified European Applications Benchmark Suite*, Partnership for Advanced Computing in Europe (PRACE), 2013.
- [9] A. Rico *et al.*, “On the Simulation of Large-scale Architectures Using Multiple Application Abstraction Levels,” *ACM Trans. Archit. Code Optim.*, 2012.
- [10] Barcelona Supercomputing Center, “MareNostrum III System Architecture,” <http://www.bsc.es/marenostrum-support-services/mn3>, 2013.
- [11] A. Suresh *et al.*, “Evaluation of Emerging Memory Technologies for HPC, Data Intensive Applications,” in *IEEE International Conference on Cluster Computing (CLUSTER)*, 2014.
- [12] H. Noguchi *et al.*, “A 250-MHz 256b-I/O 1-Mb STT-MRAM with Advanced Perpendicular MTJ Based Dual cell for Nonvolatile Magnetic Caches to Reduce Active Power of Processors,” in *Symposium on VLSI Technology (VLSIT)*, 2013.
- [13] R. Nebashi *et al.*, “A 90nm 12ns 32Mb 2T1MTJ MRAM,” in *IEEE International Solid-State Circuits Conference*, 2009.
- [14] Everspin Technologies, Inc., “Everspin Enhances RIM Smart Meters with Instantly Non-Volatile, Low-Energy MRAM Memory,” <http://www.everspin.com/everspin-embedded-mram>.



Kazi Asifuzzaman was born in Dhaka, Bangladesh. He received his Bachelor of Science (BSc) degree in Computer Engineering from North South University (NSU), Bangladesh in 2008. The following year, he worked at the IT department of Shimizu Denetsu Kogyo Co. Ltd (SEAVAC) in Japan. He completed his Master of Science (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 Universitat Politècnica de Catalunya (UPC), Spain. His research primarily focuses on analyzing suitability and feasibility of using emerging non-volatile memories as the main memory of high performance computing (HPC) systems.