RESEARCH ARTICLE OPEN ACCESS

SPIN TORQUE TRANSFER MRAM AS A UNIVERSAL MEMORY

Jyoti Garg
ABES Engineering College, Ghaziabad, Uttar Pradesh, India

ABSTRACT
The current multi-core era has resulted in the integration of increasing numbers of cores into the microprocessors used to power computers and cell phones. Spin-Transfer Torque RAM (STT-RAM) is an emerging non-volatile memory technology with the potential to be used as universal memory. STT MRAM with read and write current is discusses here . In addition to that application of STT MRAM as a cache design is also discussed.

I. INTRODUCTION

As memory is a key component of computing systems, it has stringent design requirements as it faces various design challenges such as scalability and high leakage power [1]. An emerging storage device that fits well to these requirements is nano-magnetic random access memory (MRAM). Especially the Spin Transfer Torque memory (STT-MRAM) is very promising due to its advantageous features such as non-volatility, scalability, high density, low read latency, and CMOS-compatibility [2]. In a STT-MRAM, data is stored in a Magnetic Tunnel Junction (MTJ) cell by exploiting the magnetic orientation of two independent ferromagnetic layers. However, a high write current of several hundred µA is required to flip that magnetization [3] which is a major challenge for the establishment of STT-based storage devices in universal memories. The high write current infers a high power consumption of the memory (write energy is about 10x higher than in SRAM [4]) and also severe stress is imposed on the MTJ cell. As a result, various STT performance parameters such as the Tunnelling Magneto Resistance (TMR), write current, and latency as well as lifetime of the cell are degrading over time.

II. SPIN TRANSFER TORQUE MRAM

Spin-Transfer Torque RAM (STT-RAM) is an emerging NVM technology that has significantly higher write endurance (at least 10^{12} cycles). STT-RAM is being actively explored by industry and has the potential to be an Universal Memory technology[5]. Though not as dense as Flash and PCM, it is capable of high performance operation suitable for caches and can be integrated within the microprocessor. In particular, the combination of high endurance and the lack of cell leakage make it an ideal candidate for use in a wide range of applications from cell phones to servers. STT-RAM is a new, more efficient variant of magneto resistive RAM (MRAM) in which a single bit of data is stored in the magnetic orientation of the free layer of a magnetic tunnel junction (MTJ)[6]. MTJs consist of at least two ferromagnetic layers with an oxide barrier (insulator layer) between them, as shown in Figure 1. One of the two magnetic layers is called the hard, pinned, or fixed layer and has its magnetic orientation permanently set during fabrication The other is called the soft or free layer and has a weak magnetic orientation that can be changed dynamically. MRAM and STT-RAM are non-volatile because the free layer does not need an electric current to maintain its orientation.

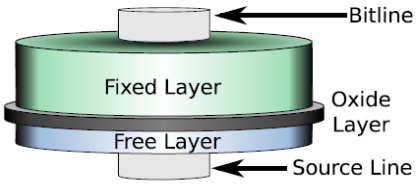

Figure 1.Structure of MTJ cell

MRAM never gained significant traction in the memory market due to its extremely high write energy requirements, caused by the use of rotating electric fields to change the free layer state. STT-RAM instead uses the spin-transfer torque effect to switch the free layer, which only requires passing a large electric current directly through the MTJ [7].
The presence of the oxide barrier between the ferromagnetic layers creates a noticeable resistance to electric current, dependent on the free layer orientation. When the two layers are oriented in the same direction, the MTJ is in the parallel (P) state and exhibits a low resistance (R_P), and when the two layers are oriented in opposite directions, it is in the anti-parallel (AP) state and exhibits a high resistance (R_{AP}). Figures 2(a,b) demonstrate the two free layer states. Reading the state of the MTJ is accomplished by using a small current through

the MTJ to estimate the resistance value. Performing a write requires holding the write current for a sufficient amount of time, which is called either the write pulse width or the MTJ write time to ensure the free layer has changed state[8].

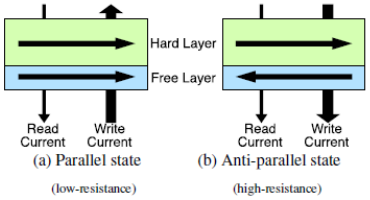

Figure 2(a,b): Magnetic tunnel junction (MTJ) operation

The Parallel configuration leads to a low resistive state R_P , whereas the antiparallel configuration leads to a high resistive state R_{AP}. Tunnel magnetoresistance (TMR), the ratio of the difference between RP and RAP, is a metric for determining the efficiency of the spintronic operation of an MTJ [9]. TMR is defined as

$$TMR = \frac{R_{AP} - R_P}{R_P} \quad \text{.....(1)}$$

A higher TMR is preferred for a reliable read operation as it will generate a larger signal difference between the two states. The physics and dependency of TMR on MTJ processing parameters can be found in [10].

First-generation MRAMs used field-induced magnetic switching (FIMS) to toggle the MTJ between its parallel and antiparallel states [11]. FIMS works by organizing word and bit lines into a crosspoint architecture. When a synchronized pulse of current is applied to the desired word and bit lines, a strong magnetic field is created at the intersection of the two wires. This magnetic field then causes the MTJ to switch to the desired state. A small access transistor is also required to read the state of each MRAM cell [12]. Aside from suffering from a serious write disturbance problem (the half-select problem), the major drawback of conventional MRAM is the increase in writing current as technologies scale [13]. The discovery of spin-transfer torque (STT)-based switching has enabled MRAMs to scale below 90 nm. Rather than using an indirect current to generate a magnetic field, STT uses a spin-polarized current through the MTJ to accomplish device switching [14]. Toggling of the MTJ is roughly determined by the current density. As the area of the MTJ device decreases, so does the writing current. Spin-transfer torque random access memories (STT-RAMs) have the added benefit of being architecturally much simpler than conventional MRAMs [15]. The simplest of STT-RAM architecture uses the one-transistor–one magnetic-tunnel-junction (1T1-MTJ) structure.

1T1-MTJ has three terminals namely source line (SL), bit line (BL) and word line (WL) as shown in Figure 3. It consists of a single MTJ cell and a single access transistor. The source of the access transistor is connected to the source line and the drain is connected to the pinned layer of the MTJ cell. The gate of the access transistor is connected to the word line to select the required bit-cell for the memory operations. The free layer of the MTJ cell is connected to the bit line terminal. In STT-MRAM, the read current is unidirectional and flows through the bit-cell to sense the stored value based on the resistance state. In contrast, the write current is bi-directional and the value to be written in the bit-cell is determined by the direction of the current as illustrated in Figure 1(b). As a consequence, the write current and the switching time are different depending on the value stored in the bit-cell, because of an inherent torque asymmetry of the MTJ device [16].

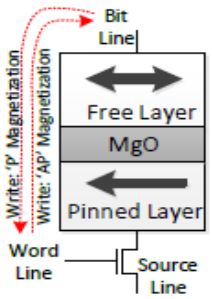

Figure 3: Bit cell Structure

III. WRITE CURRENT OPTIMIZATION

A major challenge for STT-MRAM is the high current of several hundred uA that is required to flip the magnetization of the free layer of the TJ cell [17]. Hence, the overall power consumption of the memory is negatively affected. Therefore, several research work introduced approaches to reduce this high write current. Early Write Termination (EWT) that is designed to reduce the power consumption of

STT-MRAM. The EWT technique presented in [18] terminates write operations as soon as these are detected to be unnecessary. To implement such a behavior, several additional circuits such as two conversion circuits using basic differential amplifiers, one sense amplifier, one latch, three multiplexers and one inverter are required. Due to these additional circuits, the proposed EWT architecture not only imposes a significant area overhead but also contributes to the total (write) power consumption of the memory. Furthermore, these additional circuits are part of each column of the memory bit-cell array,

so the ratio of the periphery area to that of the array circuits increases in direct relation with the number of bits per word.

STT-RAM operations can be done only on the row buffer without involving the sense amplifiers and memory cells (as also observed in [19]). Specifically, in STT-RAM, when a row buffer conflict occurs, a selective write-back can be performed, where the row buffer is not written back to the array if it is clean. If the row buffer is dirty, the row must be written back as the array contains stale data and the most recent version of the data is in the row buffer. This selective write optimization can improve performance by expediting row buffer conflicts and can improve energy by eliminating the cost of redundant array writes.

IV. READ OPTIMIZATION

Reads are completed by sensing the voltage differential in the two resistance states of the MTJ using a read current. Even with a relatively small reading current, there is still a probability to flip the MTJ states when applying a read current. Thus, for all reads to MTJs, there is a probability of disturbing the stored value. As the technology scales, the energy required for writing current (*Ic*) decreases, which is the property that leads to improved writing performance. Unfortunately, this reduction also increases the potential for read disturbance. To avoid increasing the probability of read disturbance, the ratio between read and write current must remain balanced. To address the trends of increasing sense time and uncorrectable error rate, we propose to use differential sensing. Differential sensing requires two MTJs to store both the value and its complement, and

by sensing the difference rather than comparing with a threshold, one doubles the sense margin. The sensing performance and reliability can be further improved by a better SA design. The sense amplifier was tuned for best possible performance (i.e., the sensing latency). Differential sensing has two main drawbacks [20]. The area required to store a value is doubled, and the dynamic power of write access is also effectively doubled. To efficiently leverage differential sensing and to minimize these side effects, we propose a configurable differential cell shown in Fig.3. Through a MUX and transmission gate, the cell can be configured into standard high-density mode (S-mode) by comparing the selected cell with *V*ref or in situations where read performance is critical to use extreme performance mode (X-mode) by sensing the voltage difference between adjacent cells.

V. DESIGNING CACHE

Central processing units (CPUs) are composed of core and last level caches (LLCs), i.e., L2 and L3 (and L4) caches . The main power consumer of these components of CPUs are LLCs because most transistors inside CPUs are used for static RAMs (SRAMs) for LLCs. LLC capacity depends on the purpose of use and typically ranges from 1 to 100 MB. Individual SRAMs waste static power irrespective of whether jobs are present because leakage current flows through transistors as long as power is supplied. SRAM is "normally on type memory cell" design. Because larger cache capacities very effectively enhance CPU performance, there is strong demand for smaller SRAMs. However, smaller transistors induce larger leakage current. The power consumed by leakage current is the most serious problem in designing high performance processors. Analysis of SRAM based cache memory operation shows that LLC has very short standby time from 10 ns to 100 ns while CPU is active. For this standby time, more than 80% energy is consumed of the total cache memory energy. The most effective way to eliminate leakage is to use a new memory cell deign with "normally off type"[21].

VI. CONCLUSION

In this STT MRAM is discussed with its read and write current optimization. In addition Application of STT MRAM as a main memory has been discussed. There lies a scope of a lot of research work in this field

REFERNCES

[1] X. Bi et al., "Probabilistic design methodology to improve run-time stability and performance of stt-ram caches," in ICCAD, 2012, pp. 88– 94.

[2] R. Bishnoi et al., "Asynchronous asymmetrical write termination (AAWT) for a low power STT-MRAM," in Design, Automation and Test in Europe, Mar. 2014.

[3] M.-T. Chang et al., "Technology comparison for large last-level caches (l3cs): Low-leakage sram, low write-energy stt-ram, and refreshoptimized edram." in Symposium on High Performance Computer Architecture, 2013, pp. 143–154.

[4] X. Dong et al., "Circuit and microarchitecture evaluation of 3d stacking magnetic ram (mram) as a universal memory replacement," in Design Automation Conference, 2008, pp. 554–559.

[5] W. Guo et al., "SPICE modelling of magnetic tunnel junctions written by spin-transfer torque," Journal of Physics D: Applied Physics, p. 215001, May 2010.

[6] M. R. Guthaus et al., "Mibench: A free, commercially representative embedded benchmark suite," in International Workshop on Workload Characterization, 2001, pp. 3–14.
[7] M. Hosomi et al., "A novel nonvolatile memory with spin torque transfer magnetization switching: Spin-ram," in International Electron Devices Meeting, 2005, pp. 459–462.
[8] J. Hu et al., "Minimizing write activities to non-volatile memory via scheduling and recomputation," in Symposium on Application Specific Processors, 2010, pp. 101–106.
[9] A. Khvalkovskiy et al., "Basic principles of stt-mram cell operation in memory arrays," Journal of Physics D: Applied Physics, pp. 74 001– 74 020, 2013.
[10] Y. Kim et al., "Write-optimized reliable design of stt mram," in Proceedings of the International Symposium on Low Power Electronics and Design, 2012, pp. 3–8.
[11] D. Lee, S. K. Gupta, and K. Roy, "High-performance low-energy stt mram based on balanced write scheme," in Proceedings of the International Symposium on Low Power Electronics and Design, 2012, pp. 9–14.
[12] K. M. Lepak and M. H. Lipasti, "On the value locality of store instructions," in ACM SIGARCH Computer Architecture News, vol. 28, no. 2, 2000, pp. 182–191.
[13] H. Meng and J.-P. Wang, "Spin transfer in nanomagnetic devices with perpendicular anisotropy," Applied physics letters, pp. 172 506–172 506, 2006.
[14] M. Meterelliyoz, H. Mahmoodi, and K. Roy, "A leakage control system for thermal stability during burn-in test," in International Test Conference, 2005, pp. 10–pp.
[15] A. Nigam et al., "Delivering on the promise of universal memory for spin-transfer torque ram (stt-ram)," in Proceedings of the International Symposium on Low Power Electronics and Design, 2011, pp. 121–126.
[16] G. Panagopoulos, C. Augustine, and K. Roy, "Modeling of dielectric breakdown-induced time-dependent stt-mram performance degradation," in Device Research Conference, 2011, pp. 125–126.
[17] R. Sbiaa et al., "Reduction of switching current by spin transfer torque effect in perpendicular anisotropy magnetoresistive devices," Journal of Applied Physics, p. 07C707, 2011.
[18] G. Sun et al., "A novel architecture of the 3d stacked mram l2 cache for cmps," in Symposium on High Performance Computer Architecture, 2009, pp. 239–249.
[19] G. Sun et al., "Improving energy efficiency of write-asymmetric memories by log style write," in Proceedings of the International Symposium on Low Power Electronics and Design, 2012, pp. 173–178.
[20] S. A. Wolf et al., "The promise of nanomagnetics and spintronics for future logic and universal memory," Proceedings of the IEEE, pp. 2155– 2168, 2010.
[21] C. Xu et al., "Device-architecture co-optimization of stt-ram based memory for low power embedded systems," in Proceedings of the International Conference on Computer-Aided Design, 2010, pp. 463–470.

.