Nanomaterials/Nanotechnology: Nanoelectronics to improve energy efficiency

A white paper submitted to the NIST TIP

Nick Rizzo
nick.rizzo@everspin.com
480-347-1158
Everspin Technologies
Chandler, AZ
Nov. 11, 2009

Nanomaterials/Nanotechnology

Nanoelectronics to improve energy efficiency

Summary

US energy use is costly (1 trillion dollars annually), negatively impacts the environment, and causes the US to be dependent on unreliable foreign energy sources. A significant fraction of US energy use is driven by electronic devices used in the home and business, such as personal computers, DVRs, and smart phones. The energy demand of these devices is growing both because of their increased use, and because the underlying electronics are becoming less energy efficient. The microprocessors and memory in these devices are dissipating more energy due to leakage currents as the circuits shrink to nanometer size. What is needed is a high performance, low-power, nonvolatile, nanoelectronic memory that can be integrated with microprocessors to significantly reduce energy use. The only currently available memories are inadequate because of either high power or poor performance. Spin Torque Magnetoresistive Random Access Memory (ST-MRAM) is one attractive approach that could become a universal memory, replacing semiconductor cache memories, disk storage, and even enabling nonvolatile processors. The success of this project would meet the current Adminstration's goals of improving US energy efficiency, maintaining US leadership in nanotechnology, and creating a large number of high technology jobs for US workers. NIST TIP support is critical to accelerate the tremendously expensive development of advanced nonvolatile memory technologies, especially since most of the current funding for US energy research is focused on other areas.

Energy demand of electronic devices: a critical, national, and growing challenge

The worldwide demand for energy is already large, and increasing at an alarming rate; projections indicate energy use will increase by over 40% by the year 2030.[1] (See Fig. 1.) The US uses approximately 20% of the worldwide energy supply,[1] at a cost of over 1 trillion dollars annually, which is more than 5% of the US GDP.[2] The cost of energy is likely to increase as the worldwide demand increases, since almost all energy generation is currently from non-renewable sources having finite supply, such as oil, natural gas, and coal. The US is also heavily dependent on foreign sources of energy, importing more than 60% of its oil.[3] The dependence on foreign sources can present a security risk, particularly when many of the countries supplying the oil do not have democratic forms of government, and may have leaders whose interests run counter to those of the US. In addition, the US energy use makes it one of the world's largest producers of greenhouse gases, which are a major contributor to global warming and climate change. Therefore, for economic, national security, and environmental benefits, the US has a critical need to improve its energy generation and the efficiency of its energy use.

One specific cause of increased energy demand in the US is the increased use of microprocessors and memory in household devices such as personal computers, digital video recorders, and advanced mobile devices such as smart phones. Energy use by PC's can be estimated to be approximately 5% of the total household electricity use,[4] and is likely growing, given that over 60 million PCs are sold annually in the US.[5] Also, DVRs, which are similar to PCs in that they need a microprocessor and a disk drive, now are in approximately 25% of US households.[6] Smart phones with advanced microprocessors and memory are also growing tremendously in popularity— worldwide smart phone sales are projected to exceed those of desktop PC sales by 2011.[7] Given the growth of these consumer electronic applications, it is highly likely that they will account for up to 10% of US household electricity use in 2010.

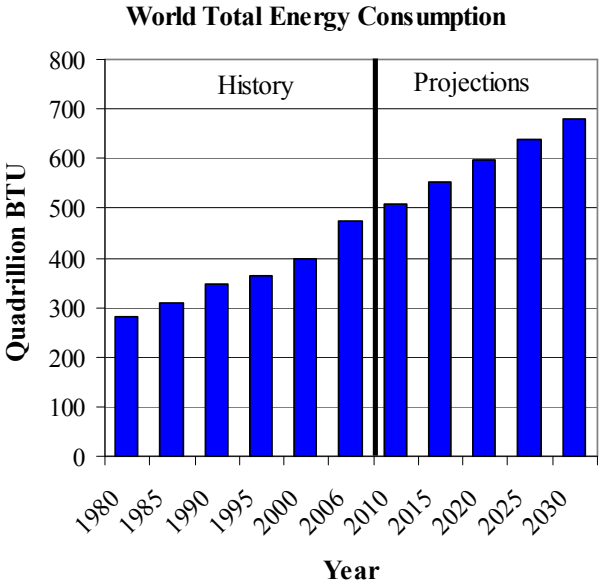

Figure 1. Worldwide Total Energy Consumption (in Quadrillion BTU) vs. year, from Reference 1.

In additional to household use, business use of microprocessors and memory in PCs and data servers are also creating significant demand for energy in the US. It has been estimated that over 100 million PCs are used in US businesses; when not in use, these machines are still powered, consuming energy that costs several billion dollars a year to produce and causing million of tons of CO_2 to be emitted in the process.[8] Data server farms are a particular business application that is producing significant energy demand. An individual data server is essentially just a microprocessor with memory and a disk drive. A data server farm is a collection of individual servers and is used for storing and quickly accessing large amounts of information. Data server farms can house up to thousands of individual servers. Server farms are used by many US businesses and government organizations to store information such as financial transactions, health records, manufacturing records, etc… The energy demand of US server farms was estimated to be 1.5% of US electricity use in 2006 and projected to grow to twice as large as that by the year 2011.[9] (See Fig. 2). Indeed, the growth of

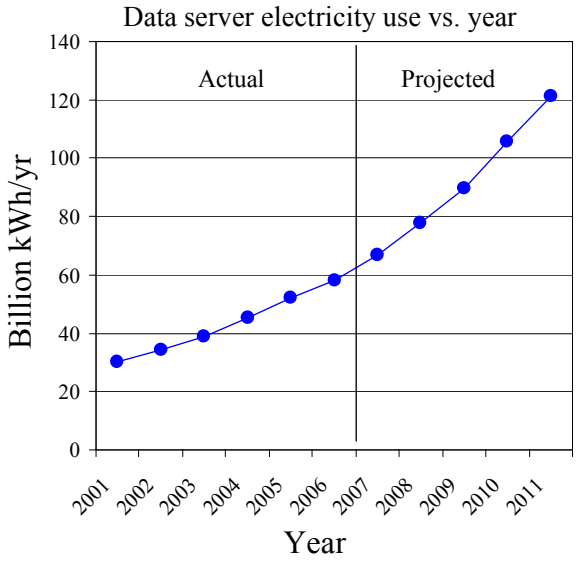

Figure 2. Date server farm electricity use (in billion kWh/yr) vs. year, from reference 9. The 2006 value was 1.5% of US electricity use and is projected to double by the year 2011.

"Nanoelectronics to improve energy efficiency" -- White Paper submitted to the NIST TIP by N.D. Rizzo, Everspin Technologies, Nov. 11, 2009.

server farms and their accompanying energy demand was recognized to be such a significant problem that Congress commissioned a special report by the EPA on data servers and their environmental impact.[9]

As pointed out above, microprocessors and associated memory are fundamentally responsible for significant energy demand the US. This energy demand will continue to grow not only because of the increased number of devices using advanced semiconductor technology, but also because the underlying semiconductor technology is becoming less energy efficient as the circuits shrink to nano-size. Leakage current of the transistors is responsible for an increasing share of the power consumed by advanced processors and stand-alone memory chips. The ITRS roadmap shows that leakage power per SRAM cell (Static power dissipation) is expected to increase 10X over the next decade (See Fig. 3).[10] Coupled with an expected increase of 10X in total number of logic and SRAM devices, the total power consumed by SRAM leakage in high-performance processors in 2019 could be 100X the already high values seen today. Leakage in SRAM cells is also exponentially dependent on temperature, increasing by approximately 3X between 25°C and 100°C.[11]

SRAM based cache memories are also becoming an increasingly large fraction of microprocessors for enterprise-class computing and networking. For example, the SRAM cache is 70% of the StrongARM[12] processor and 32% of the 45nm 8-core Intel Xeon processor[13] which has a shared 8MB-L3 cache with 731 million transistors. As a result, these memories and their associated leakage are among the main sources of power dissipation in such chips. It is common for the SRAM cache to already account for over 10% of the energy consumed, and in some cases it can be as high as 40%.[14]

Designers of advanced microprocessors are increasingly going to great lengths to manage power due to excessive leakage. In many instances, current SRAM designs include an enlarged memory cell size (from 6 transistor to 8 transistors) in an effort to reduce leakage power. Other designs include the use of on-chip microcontrollers solely for power management and multiple voltages to allow the use of low-leakage shut-down, standby, and drowsy modes.[13,15]

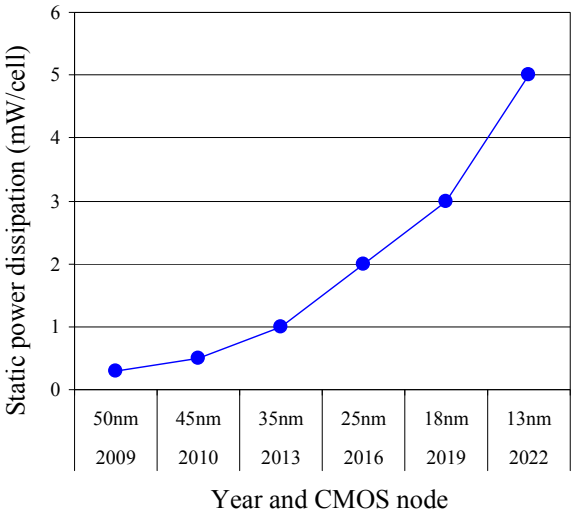

Figure 3. Static power dissipation of high performance logic (in mW/cell) vs. CMOS node and year, from the ITRS roadmap.[10]

Advanced nonvolatile memory: a path to low power electronics

To enable low power electronics, what is needed is a low power, fast, high density nonvolatile memory that can replace the energy inefficient SRAM cache memories currently in use. Such a memory would enable aggressive power management by allowing circuits to be shut down when not in use, without the need for saving and

restoring the state of the system. The boundary between nonvolatile memory and main memory would be eliminated, allowing greater flexibility in programming and adapting the same hardware to multiple uses, while also eliminating batteries in many uses. If the nonvolatile memory were of high enough performance and compatible with CMOS processing, then it would also be possible to incorporate it directly into the logic device, making for a nonvolatile CPU with zero static power dissipation. In fact, NEC (a Japanese company) is currently working on a low power CMOS flip-flop that incorporates nonvolatile magnetic bits similar to those used for MRAM.[16] In addition to nonvolatile logic and cache memory, a very high density nonvolatile memory would have the potential to replace high capacity disk storage, which would also provide significant energy savings. Disk drives are already being replaced by more energy efficient Flash memory drives in many energy sensitive applications. However, though Flash memory is a high density nonvolatile memory, it far too slow for cache or logic applications and wears out after as little as 10^6 cycles.

There are several different nanoelectronic memory technologies in development today that are vying to become a universal low power, high performance nonvolatile memory. Phase change memory (PCM) uses small amounts of chalcogenide material as the bit. The material changes phase from high resistance to low resistance depending on how the bit is heated by current pulses.[17] Although PCM is at a more advanced stage of development,[18] slow write times, limited write endurance, and poor reliability are preventing it from becoming a universal low power memory of choice. PCM would certainly still benefit from further research and innovation. Nano-ionic memory relies on growing or dissolving conductive nano-filaments between electrodes, depending on the voltage polarity applied.[19] There are still many technical hurdles to overcome for Nano-ionic memory to be commercialized. Spin Torque MRAM is the approach that Everspin technologies would like to develop as a universal, low power, high performance nonvolatile memory.

Spin Torque MRAM: the potential for a universal memory

An ST-MRAM bit consists of magnetic tunnel junction (MTJ) tri-layers, as shown in Figure 4. The tri-layer is made up of a free layer, whose magnetization can point in one of two directions, an insulating tunnel junction layer typically made up of MgO, and a fixed magnetic layer, whose magnetization direction is held fixed in one direction. The bits are programmed by passing a current *I* through the tri-layer and the pass transistor below the bit. A current in the down direction programs the free layer magnetization to be parallel to the fixed layer magnetization (0 state). A current in the up direction programs the free layer magnetization in the opposite direction (1 state). (The ST effect is a result of

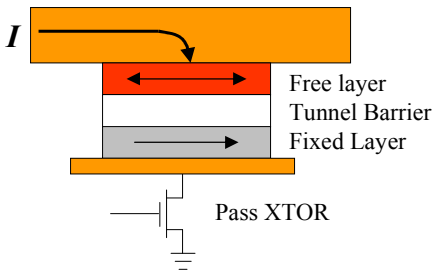

Figure 4. Schematic cross section of ST-MRAM bit. A current *I* is passed through a magnetic tunnel junction trilayer to program the 0-state or 1 state. The magnetization of the free layer (double arrow) points left or right depending on the polarity of *I*. The fixed layer magnetization remains fixed. To read the state: for parallel magnetization between free and fixed, the resistance is low; for antiparallel magnetization, the resistance is high.

the magnetic torque applied to the free layer due to the spin angular momentum associated with a spin polarized current.[20,21] For an excellent review of spin torque phenomena, see Refs. 22,23,24. See Appendix A for more detail on what determines the spin torque switching current.) The free layer magnetization direction is detected by measuring the resistance of the bit using a smaller read current. When the fixed and free layers are parallel, the resistance is low, and when the layers are antiparallel, the resistance is higher.

As Table I shows, an ST-MRAM has the potential to offer a set of attributes that are not found in any other memory technology: nonvolatility, low write energy, unlimited endurance, high density, and random access with fast read and write. Toggle MRAM, currently in production by Everspin, has similar attributes to ST-MRAM, with the exception of being lower density and having a larger write power. Even so, Toggle MRAM is already replacing other volatile and nonvolatile RAM solutions in a number of applications, and is enabling new capabilities in others. Clearly, ST-MRAM could provide significant advantages over SRAM in both cell size and in reduced active and static power dissipation. In addition, ST-MRAM could also replace NOR flash, since it has the potential for a similar cell size, but far superior performance. DRAM is volatile, and incompatible with CMOS processing, so that it too could potentially be replaced by a high performance ST-MRAM.

The goal of our ST-MRAM development is to reduce the current needed to program the free layer for two reasons: 1.) to achieve a high density memory 2.) to prevent tunnel barrier breakdown. Reducing the programming current is needed for high density because a smaller current allows the use of a smaller pass transistor in each cell. A typical logic transistor can pass approximately 600 μA per micron of transistor width, almost independent of technology node. A 90 nm × 180 nm elliptical bit with critical current density $J_c \approx 5$ MA/cm^2 would require just over 600 μA to switch, making the transistor width dominate the cell size. It would be desirable to achieve $J_c \sim 1$ MA/cm^2, or less, to enable a switching current $I_{sw} \sim 100$ μA, so that the size of the transistor is better matched to the MTJ size. In general, the goal is to reduce I_{sw} so that a minimum-size transistor can be used at each technology node, thus achieving the highest possible

Table I. Comparison of critical parameters for different memory types. Toggle MRAM is the only nonvolatile, high speed, infinite endurance memory currently in production. ST-MRAM has all the positive attributes of Toggle MRAM, but has the potential for much higher density and lower programming power.

	Toggle MRAM 180nm	Toggle MRAM 90nm*	ST MRAM 90nm*	FLASH 90nm†	DRAM 90nm†	SRAM 90nm†
cell size (μm²)	1.25	0.32	0.06#	0.064	0.06	**0.6**
Read time (ns)	35	10	10	10-50	10	<1
Program time (ns)	5	5	10	**0.1-100ms**	10	<1
Program energy/bit (pJ)	150	100	1	**10nJ**	5	5
Endurance	>10^{15}	>10^{15}	>10^{15}	>10^{15} Read **>10^5 Write**	>10^{15}	>10^{15}
Non-volatility	YES	YES	YES	YES	**NO**	**NO**
CMOS compatible	YES	YES	YES	YES	**NO**	YES

*90nm MRAM values are projected.

†These values are from the ITRS roadmap.

#This area is obtained assuming the minimum pitch compatible with the back end bit cell; it does not take into account potential increases in area due to front end CMOS requirements.

"Nanoelectronics to improve energy efficiency" -- White Paper submitted to the NIST TIP by N.D. Rizzo, Everspin Technologies, Nov. 11, 2009.

memory density. A lower programming current also prevents tunnel barrier breakdown because of the lower bias voltage across the tunnel barrier. The switching voltage of the free layer $V_{sw} = I_{sw}R$, must be much less than the breakdown voltage of the tunnel barrier V_{bd}, otherwise the tunnel barrier will be damaged. There has been substantial progress in reducing J_c over the past several years, but further reductions are needed.

The Possibility of Transformational Results

As made clear by the discussion above, the US development of a nanoelectronic low-power, or even zero-power (nonvolatile) microprocessor and/or associated memory would have a transformational impact given the ubiquitous use of these devices in US society. First and foremost, US energy consumption would be significantly reduced, saving tens of billions of kWh of electricity over the next 20-30 years. Second, such US-based technology would generate thousands of high-technology jobs and a leadership position in next-generation semiconductor nanotechnology for the US. Based on Gartner Dataquest and internal analyses, the global semiconductor memory market is anticipated to exceed $53 billion by 2011. Due to the size of these markets, a successful program that serves any significant portion has the potential for huge job creation, including multiple billion-dollar factories, increased US based R&D efforts across the high tech semiconductor and systems industry, including numerous support jobs that surround global manufacturing efforts

Maps to Administration Guidance

"…Energy supplies are growing scarcer, energy demands are growing larger, and rising energy use imperils the planet we will leave to future generations. And that's why the world is now engaged in a peaceful competition to determine the technologies that will power the 21st century…The nation that wins this competition will be the nation that leads the global economy…And I want America to be that nation."

President Barack Obama
Address to MIT
Cambridge, MA
Oct. 23rd, 2009

Improving our energy technologies is clearly in-line with the goals of the current Administration. The US is a world leader in nanotechnology research and recognizes the importance of applying advances in nanotechnology to help meet our energy needs. Indeed, "Nanoscience research for energy needs" was listed as one of the Grand Challenges of the National Nanotechnology Initiative (NNI) in 2004.[25] (The NNI is the organization that coordinates and directs the nanotechnology research and development throughout the various departments of the US government.) In addition, both energy and, more specifically, improved energy efficiency of CMOS devices were listed as a "High Impact Application Opportunity" and a "Critical Research Need" in the NNI strategic plan in 2007.[26] NIST itself claims to be in support of research for national energy needs, and specifically cites its support for "a new generation of spintronic nanoelectronic

devices that promise to be 10 times faster and 100-times more energy efficient than conventional electronics," in the NNI supplement to President's 2010 budget.[26]

The Justification for Government Attention

Current US research in the area of novel nonvolatile memories is fragmented, being carried out by a small number of small groups, and is lacking sufficient funding. Bringing a new CMOS technology to market is tremendously expensive, requiring significant resources. For example, the total cost of commercializing Everspin's Toggle MRAM has easily exceeded $200M dollars. Government attention and financial assistance is therefore desperately needed to coordinate and boost the US effort to achieve the ambitious goal of developing a new nonvolatile memory that is capable of displacing one or more of the incumbent technologies over a broad range of applications.

Financial assistance from the government can be a key accelerator in developing costly high risk, high reward technologies. Companies are more willing to engage in more speculative, longer term research when the financial cost/risk is mitigated or reduced by support from an outside agency. In fact, the successful commercialization of Toggle MRAM by Everspin Technologies is a great example of the positive effect of government support on nanoelectronics research. The DARPA Spintronics program provided funding of more than $2M/year for 5 years to Motorola beginning in 1995 for the development of MRAM. Motorola (whose MRAM program became Everspin Technologies) had no prior experience in magnetic devices before this program and would likely not have engaged in such a risky venture. Now, slightly more than 10 years later, Everspin introduced the first commercial MRAM part in 2006. Everspin now employs over 50 high-tech workers, has sales approaching $10M annually and is on track for significant growth. Thus, the government funding played a key role in enabling a US company to be a world leader in nanoelectronic devices.

Although US funding for energy research in increasing overall, nanoelectronics research for energy still needs much more support to produce significant results. Much of the increased funding is going to projects to improve energy generation in photovoltaics or thermoelectrics, and energy storage in fuel cells or supercapacitors. For example, the DOE list of SBIR technical topics does not include any topic under which one could propose to improve the energy efficiency of nanoelectronic devices. The recently announced ARPA-E projects also did not award funding to any companies attempting to solve the critical problem of power dissipation in nanometer sized electronic devices. As the NNI supplement to the FY2010 budget notes, "…much of the increase in 2010 [in DOE funding for nanoscience] results from the initiation of Energy Innovation Hubs focusing on electrical energy storage and solar fuels." The Semiconductor Research Council (www.src.org), a worldwide consortium of companies and universities, has established a program called the Nanoscale Research Initiative (NRI) to address the overall problems (including leakage) with scaling semiconductor circuits to the nanoscale. The goal of the NRI "is to discover the next switch, a new mechanism for computing that goes beyond simply improving today's transistor." Unfortunately, the NRI funding is primarily for University based research groups and the timeframe to achieve their goal of going "beyond CMOS" is by the year 2020, which too far out for significant direct involvement by industry.

What the NIST TIP needs to provide is enough funding for 5 to 10 companies to pursue various approaches to developing advanced, nano-sized, low power memory and microprocessors. Based on our previous experience, it is expected that any significant, realistic effort to develop a new memory technology will cost each company at least $2-4M/year, so that at least $15-20M would be needed annually for this program. By comparison, the magnitude of such a program would be only a fraction of the budget for similar programs in Japan. In 2008, the Japanese Government's New Energy and Industrial Technology Development Organization (NEDO) provided $150M annually in support of Electronics and Information Technology overall, including programs on Next Generation Semiconductor Materials and Process technology and a Spintronics Nonvolatile Devices Project.[27] If similar US programs were supported by NIST TIP funding, US nanotechnology development will be significantly accelerated, US leadership in nanotechnology will be maintained, and US energy use will be reduced.

Appendix A – ST-MRAM switching current

The spin torque switching current for a bit in zero field can be shown to be approximately:

$$I_{sw} = \left(\frac{2\hbar}{e}\right)\left(\frac{\alpha M_s V}{g}\right)\left(H_k + 2\pi M_s\right) \quad (1)$$

where $\hbar$ is Planck's constant, e is the electron charge, α is the Gilbert damping constant of the free layer, M_s is the saturation magnetization, V is the free layer volume, g is the spin torque efficiency, and H_k is the in-plane anisotropy of the free layer. The $2\hbar$ term is the angular momentum change per electron. The proportionality to α originates from the spin torque needing to overcome the damping torque for a magnetization reversal to occur. (The damping constant α is proportional to the magnetization loss of angular momentum, primarily to the lattice.) The term in parentheses reflects the magnetic anisotropy that the spin torque must overcome to produce a free layer magnetization reversal. The $2\pi M_s$ term derives from the out-of-plane demagnetizing field for a magnetic thin film. For the geometry shown in Fig. 1, the direction of the spin torque rotates the free layer magnetization out of the film plane. Therefore the out-of-plane demagnetizing field must be overcome for a switch to occur. Since usually $2\pi M_s >> H_k$, the demagnetizing anisotropy is the dominant term.

Equation 1 implies that a path to lower I_{sw} would be to reduce the bit volume V or the magnetization M_s of the free layer, however this path is limited for several reasons. First, the free layer thickness t cannot be thinner than ≈2-3nm or the magnetic properties of the film degrade. Second, reducing M_s or V increases the likelihood of data loss due to thermal fluctuations, since the energy barrier (E_b) to magnetization reversal caused by thermal fluctuations proportional to $M_s^2 V$. For a memory with a typical 10 year nonvolatility requirement, one needs $E_b/k_bT \geq 50$. Therefore, other approaches to lower I_{sw} are needed.

[1] *International Energy Outlook 2009*, Energy Information Administration, Report#DOE/EIA-0484(2009), May 27, 2009.

[2] *AnnualEnergy Outlook 2009*, Energy Information Administration, Report#DOE/EIA-0383(2009), March, 2009.
[3] *U.S. Crude Oil Supply & Disposition*, Energy Information Administration, www.eia.doe.gov, accessed Nov. 9, 2009.
[4] *2005 Residential Energy Consumption Survey*, Energy Information Administration, www.eia.doe.gov; Assuming a typical PC power use of 100W (excluding monitor) and always on, the energy use is more than 8% of total electricity use per household.
[5] Gartner press release, www.gartner.com., Jan. 15, 2009
[6] "*Nielsen reports TV, Internet, and Mobile Usage among Americans*," Nielsen press release, en-us.nielsen.com, July 8, 2008.
[7] **"***More Smartphones Than Desktop PCs by 2011*", Lexton Snol, PC World, Sept. 3, 2009.
[8] "*Report From 1E and the Alliance to Save Energy Finds That Organizations Waste Billions of Dollars Running Idle Computers*," Press release, Alliance to Save Energy, ase.org, March 25, 2009.
[9] Report to Congress on Server and Data Center Energy Efficiency Public Law 109-431, U.S. Environmental Protection Agency ENERGY STAR Program, August 2, 2007.
[10] *International Technology Roadmap for Semiconductors*, 2008 Update, Table SYSD4a. www.itrs.net.
[11] D.-H. Lee, D.-K. Kwak, and K.-S. Min, *J. Comp*. **3**, 39 (2008).
[12] V.G. Moshnyaga, and K. Inoue, in *Low-Power Electronics Design*, ed. C. Piguet (CRC Press, 2005), p. 25-1.
[13] S. Rusu, S. Tam, H. Muljono, J. Stinson, D. Ayers, J. Chang, R. Varada, M. Ratta, S. Kottapalli, *ISSCC Dig. Tech. Pap*. 3.1, 56 (2009).
[14] H. Noori, M. Goudarzi, K. Inoue, K. Murakami, *IEICE Trans. Elec*. **E91-C** 418 (2008).
[15] "*A Family of 45nm IA Processors*," R. Kumar and G. Hinton, ISSCC Dig. Tech. Pap. 3.2, 58 (2009).
[16] "*NEC develops a nonvolatile magnetic flip flop that enables standby-power-free SoCs*," NEC press release, www.nec.com, Jan. 5, 2009.
[17] "*Full Integration of Highly Manufacturable 512Mb PRAM based on 90nm Technology*," J.H Oh, et al., International Electron Devices Meeting 2006, p. 1-4, San Francisco, CA, (2006).
[18] "*Samsung moves phase-change memory to production*", Peter Clarke, EE Times, Sept. 22, 2009.
[19] "*Nanoionics based resistive switching memories*," R. Waser and M. Aono, Nature Mat. **6**, 833 (2007).
[20] J.C. Slonczewski, *J. Magn. Magn. Mat*. **159**, L1 (1996).
[21] L. Berger, *Phys. Rev. B* **54**, 9353 (1996).
[22] D.C. Ralph and M.Stiles, *J. Magn. Magn. Mater*. **320**, 1190 (2008).
[23] J. A. Katine and E. E. Fullerton, *J. Magn. Magn. Mater*. **320**, 1217 (2008).
[24] J. Z. Sun and D. C. Ralph, *J. Magn. Magn. Mater*. **320**, 1227 (2008).
[25] Nanoscience Research for Energy Needs, Report of the National Nanotechnology Initiative Grand Challenge Workshop, March 16-18, 2004, www.nano.gov.
[26] www.nano.gov.
[27] www.nedo.go.jp.