White paper, Technology Innovation Program

Critical National Need Idea Title: **Unified Memory/Storage – UMS**

Submitting Organization: **Integrated MagnetoElectronics (IME)**

Contributing Organizations: **None**

Contact:
Richard Spitzer, PhD
IME
1214 Oxford Street
Berkeley CA 94709-1423
Voice: 510-841-3585
Fax: 510-841-2405
richard.spitzer@transpinnor.com

Key words: electrical energy computer storage data centers storage farms unified memory/storage

White paper, Technology Innovation Program

Unified Memory/Storage – UMS

Critical National Need

The gross national product of the United States increasingly depends on electronic transactions across the entire spectrum of commerce. But this trend towards electronic ordering and delivering of goods and services, and towards consumer communications is seriously threatened by one critical element: the inefficient use of electrical power, which is outpacing all other demands on the system. The "Report to Congress" by the Environmental Protection Agency, 2 August 2007, specifies five pressures on commerce, all having to do with electrical-energy efficiency:

- electronic transactions
- internet communications and entertainment
- electronic medical records
- global commerce
- satellite navigation and shipment tracking

This white paper identifies and defines a specific energy-related problem whose resolution has the potential to enhance significantly the economic growth of the nation: the soaring size and electrical-energy requirements of computer storage in data centers and server farms. This is the single most pressing problem facing that industry. The challenge is twofold: 1) to reduce the energy consumption of the computer-storage industry by developing a more energy-efficient and smaller storage technology than that presently in use: hard-disk drive (HDD); 2) to meet the greatly expanding need for storage projected for the future for data centers and server farms.

The magnitude of the problem is evidenced by data on energy usage, and on the rapid growth of this usage, in the computer-storage industry.

A study from Lawrence Berkeley National Laboratory (J. G. Koomey, *Estimating Total Power Consumption by Servers in the U.S. and the World*, LBNL, February 15, 2007) shows that total power used by servers (including cooling and auxiliary infrastructure) represented about 1.2% of total U.S. electricity consumption in 2005. The total corresponding power demand is equivalent (in capacity terms) to about five 1000 MW power plants for the U.S. and 14 such plants for the world. The total electricity bill for operating those servers and associated infrastructure was about $2.7 B and $7.2 B for the U.S. and the world, respectively.

The total energy consumed in this sector in the U.S. has grown from 1.0 % to 1.5 % in the first five years of this decade, and is predicted to reach 2.0 % by 2010. The total cost (new servers together with power and cooling) of server rooms based on HDD in U.S. market has risen from $20B in 2001 to $60B in 2008, with $80B projected in 2009. The corresponding fraction of total cost for power and cooling has risen from 10% in 2001 to 50% in 2008, with 70% projected for 2009 (R. F. Freitas and W. W. Wilcke, IBM J. of Research and Development, July 2008, vol. 52, issue 4, p. 439-447; www.research.ibm.com/journal/rd52-45.html).

The magnitude of the problem is so eminently vast as to justify government action if the attendant societal challenges that need to be overcome are not being adequately addressed, as is in fact the case (see ***Why Government Support is Needed***, below).

The heart of the technical problem is the great disparity between memory (fast, expensive, volatile) and storage (slow, cheap, nonvolatile). It can be addressed through high-risk, high-reward research of the kind proposed in ***Potential Technical Solution***, below. The proposed technical solution has the potential to eliminate the bulk of the present and projected cost of energy consumption based on HDD storage and of the associated capital component of the storage hardware – a reduction in energy and space by factors of 100-10,000.

Why Government Support is Needed

Major manufacturers of data-storage equipment are currently pursuing an increasing number of storage options. This was emphasized by T. Coughlin in the opening talk at the IDEMA (International Disk Drive Equipment and Materials Association) Technical Symposium, "The Future of Nonvolatile Technologies", held 11 December 2008 in Milpitas, Calif. But the problems presented by data storage and access are currently being addressed by industry in piecemeal fashion.

Current methods and strategies for reducing power consumption in computer storage involve incremental improvements. For example, a typical large server, such as Fujitsu-Siemens', costs about $1M; with redesign it has reduced energy consumption from 16Kw to 10Kw daily. Sun has done the same. Micron Technology offers a line of low-voltage servers that result in lower power demand. Further, at least two companies have bundled flash devices as solid-state disks to lower power consumption in data centers (see ***Potential Technology Solution***, below). The societal challenge to reduce power consumption is not being addressed by such incremental approaches to its manifestation in computer storage. Failure to address this challenge will inhibit the Nation’s economic growth (see ***Consequence of Inaction***, below)**.** An approach to the problem that enables a disruptive change, over and above the current approaches, is clearly called for to meet satisfactorily the expanding storage requirements and attendant energy needs.

These product innovations do, however, attest to the urgency of the energy problem. They also indicate why a more basic approach has gone unfunded, and is unlikely to be funded in the future, by private capital. Venture capital firms avoid investing in companies interested in preserving their existing products, with market presence, and with the manufacturing and sales capability to maintain their market share. Coughlin expressly highlighted in his IDEMA presentation the 'extreme difficulty of displacing individual companies pursuing incremental advances in their own products'. The large barrier presented by the current methods and strategies prevents the successful development of a solution to the area of critical national need represented by the growing demand for energy in computer storage. Government action is justified by its potential to overcome this barrier.

Consequence of Inaction

Based on market trends over the last five years, the demand for storage is projected to exceed HDD capacity by the end of this decade (Coughlin, IDEMA Symposium). Already in 2008, existing HDD data-center facilities were often unable to support storage-capacity requirements (H. Smith, Director, Enterprise Storage Marketing at Samsung Electronics; IDEMA Symposium). Smith presented data indicating that retrofitting of existing HDD data centers as well as construction of new such centers to meet the increasing demands for data storage are both prohibitively expensive.

Furthermore, there is neither enough worldwide fabrication capacity to replace HDD by silicon storage, nor enough funding to build such capacity (S. Hetzler, Manager, Storage Architecture Research at IBM; IDEMA Symposium). Clearly, a disruptive technology is needed to fill future storage requirements cost effectively and energy efficiently.

Failure to develop a UMS technology capable of keeping up with the growing demand for storage capacity will severely curtail the economic growth attendant to meeting this demand.

Potential Technical Solution

The potential technical solution to energy-efficient, cost-effective storage lies in a single nonvolatile memory/storage technology that will combine the advantages and minimize the respective disadvantages of solid-state memory and hard-disk storage: Universal Memory/Storage (UMS) to replace active memory (DRAM), cache (SRAM), and storage (hard disk and flash). UMS blurs the aforementioned distinction between memory and storage. System requirements for UMS applications include nonvolatility, endurance (> 10^{12} cycles), speed (~ DRAM like), low cost per bit (no more than 3-5 x cost of HDD storage; < $1/GB by 2015), low power (at least 100 times lower than HDD storage); (R.F. Freitas, IDEMA Symposium).

The IBM study cited in ***Critical National Need***, above, states that use of UMS (IBM refers to this technology as "Storage Class Memory", SCM) in server farms will save factors of 1000 to 10,000 in space and energy. In effect, the bulk of substantially the entire present and projected cost of energy consumption and the capital cost of storage elements of HDD data centers and server farms can be eliminated using UMS. The savings for data-center and server-farm applications, extrapolated to 2020 from the data cited in the IBM study, are quantified in Table 1.

Table 1. Comparison of HDD and UMS characteristics – Extrapolation to 2020

	Data centers		Server farms	
	disk	UMS	disk	UMS
Number of devices	1.3 M disks	406 modules	5 M disks	8 K modules
Space	4500 sq ft	85 sq ft	16,500 sq ft	12 sq ft
Power	6000 Kw	41 Kw	22,000 Kw	1 Kw

The urgency of the soaring energy and capacity requirements of server farms and data centers has been underscored by the rapid market entry, by companies such as Fusion-io and Texas Memory Systems, of flash devices bundled as solid-state disks to address these requirements. But there is neither the fabrication capacity nor the funding to build such capacity for silicon storage (see ***Consequence of Inaction***). Thus, though flash-based UMS meets the first challenge identified in ***Critical National Need***: a more energy-efficient and more compact storage technology than HDD, it fails to meet the second: satisfy the greatly expanding future need for computer storage. The flash technology used by Fusion-io and Texas Memory Systems is thus only a stopgap measure, and a technology other than silicon-based is needed to fill future storage requirements more cost effectively, energy efficiently, and compactly.

The high risk of UMS development is evidenced by the fact that some of the largest electronics companies in the world have spent more than 20 years and billions of dollars in seeking a solution to various aspects of the memory-storage dichotomy, with results that provide only partial solutions, none adequate to meet the existing and rapidly increasing crisis. The high reward of successful UMS development is attested to by its wide-ranging implications, far beyond those of addressing the crisis that exists in computer storage in data centers and server farms – the subject of this white paper.

These wide-ranging and long-term implications of the potential project outcome assure a transformational result. Disruptive changes will extend to many other applications in addition to data-center and server-farm storage, including all mobile devices, where GMR RAM's low power consumption translates into significantly longer battery life; embedded memories in microprocessors, to replace the presently used combination of SRAM and flash by a single technology; specialized military and space applications, to take advantage of the inherent radiation hardness of IME's GMR RAM; and a fundamental transformation of the hardware, operating system, and operation of computers to provide instant-on capability. The last alone – some 150 million laptop computers and another 100 million desktops are manufactured each year – guarantees a disruptive change.

The IBM study cited in ***Critical National Need***, above, concluded that the impact of UMS will be to usher in a 'major new industry' – a potential project outcome that enables disruptive changes over and above current methods and strategies, and challenges the status quo of research approaches and applications.

Audience for Competition

Several technologies are presently competing to become the UMS of choice: magnetic RAM, ferroelectric RAM, phase-change RAM, resistive RAM, solid electrolytes, polymer/organic, advanced flash. Development of most of these technologies has been ongoing for 20 years, some even longer.

Nearly 50 companies world-wide are active in developing these seven types of nonvolatile memory technologies; many in more than one technology, about one-third in the U.S. (R.F. Freitas, IDEMA Symposium). All U.S. companies in these fields – both established and innovative startups – are potential candidates for developing proposal submissions to TIP for UMS.