Research Needs for Advanced Memory for 32nm Technology Node and beyond
September 14, 2005 Edition[1]

SRC Devices Sciences – Memory Task Force

Semiconductor Research Corporation
P.O. Box 12053
Research Triangle Park, NC 27709-2053
Phone: 919-941-9400

Introduction

Continued growth in memory products will be hampered by inability to further scale the majority of existing memory devices. It is questionable whether a stable SRAM cell can be scaled to 32nm node without sacrificing cell size. Similarly, DRAM is facing problems in how to store enough electrons in ever diminishing storage volumes. For traditional non-volatile memory devices like NOR flash, scaling could stop even before the 32nm node, while scaling of mass storage devices like NAND flash, is going to be very difficult beyond the 32nm node. Novel architectures may be needed to further improve data bandwidth in embedded as well as stand-alone memory applications. Cost is an important issue for memory devices. Solutions like 3D integration and stacking are viable alternatives to scaling. In order to realize stacked memory devices novel selection devices are needed.

The Device Sciences area of SRC is soliciting proposals for advanced memory. We acknowledge a wide breadth of device and material research that is being pursued at universities in the field of novel memories. While there is a plethora of publications describing new materials with memory effects, SRC is seeking research that will combine materials, device, and system research. Memory device applications are so broad that we foresee very different solutions, although a unified memory that would combine the speed of SRAM, density of DRAM, and non-volatility of flash would be desirable.

The purpose of this document is to offer industry-driven research needs vectors for variety of embedded and stand-alone memory devices. It is intended to address novel materials, devices concepts, and architectural aspects of memory devices addressing future memory applications.

Scope and Technology Requirements

The scope of this document is limited to those research needs related to extending memory devices to and beyond the 32nm node. The technology requirements are integration with CMOS, high functional density, high speed, low power dissipation, and low cost. The major technology barriers are stability, reliability, data retention, and disturb mechanisms. There is a significant interplay between requirements and barriers, and optimized trade-offs between these are within the scope of this solicitation.

---

[1] Includes inputs from 7 SRC Member Companies (AMD, Axcelis, Freescale, IBM, Intel, LSI-L, and TI) and SRC.

Research Needs

The Research Needs are organized into three major categories, as shown in Table 1: (1) Embedded Memory, (2) Non-volatile Memory, and (3) Selection Devices. Within each category we distinguish material/integration and device/architectural aspects. We do not include modeling and simulation that has been addressed by a separate solicitation.

Embedded Memory

The desire to integrate large quantities of memory on chip historically has been driven by four main needs: System Performance, Form Factor, Power Reduction, and Memory Granularity.

In the semiconductor industry today, there are three types of embedded memory beyond the easily-implemented ROM. SRAM is by far the most commonly embedded memory technology and typically dominates die size for many systems. Embedded DRAM has become more attractive if large quantity of on-chip, volatile memory is required in a given system. Finally, embedded flash in small, medium, and large densities is also common, especially in microcontroller technology. The research needs of each are discussed below.

Embedded SRAM

SRAM is by far the dominant form of embedded memory found in today's integrated circuits. In fact, the majority of transistors found in many integrated circuits are those utilized in the SRAM bit cells with the percentage die occupied by this type of memory approaching 75%-85% for certain integrated circuits. Because this memory is typically constructed from traditional CMOS devices, all of the issues associated with MOSFET scaling apply to scaling of SRAM. Issues such as dopant fluctuations, gate oxide leakage, control of short channel effects, contact resistance, abrupt and low spreading resistance junction technology must be resolved for continued scaling of the traditional SRAM bit cell.

Additionally, there is a desire to find a dense SRAM replacement that can substantially reduce the area occupied by the traditional SRAM bit cell (6T bit cell traditionally) while maintaining the performance offered by the current technology. Discovery of such a bit cell would have profound implications on the die cost of integrated circuits given the ever-increasing area occupied by this type of memory. In addition to area scaling of the bit cell, there is also a need to develop alternative bit cells that maintain stability while operating at low voltages, thus, allowing the industry to substantially reduce standby power consumption in the memory arrays. In all of this research, SRAM bit cells that maintain or increase read/write performance are required.

Embedded DRAM

Embedding DRAM into a CMOS process flow has become more popular over the last decade as a means for integrating large quantities of memory on chip. Typically, this technology is based on a variant of the 1 transistor -1 capacitor (1T1C) bit cell found in stand-alone DRAM. As such, the technology as it exists today is plagued by many of the scaling issues associated with stand alone DRAM. In particular, there has historically been a need to maintain a minimum capacitance associated with the capacitor of the bit cell. The actual value of the capacitance is dictated by the memory array architecture but is traditionally in the range of 20-40fF. The need to maintain this level of capacitance in an

ever-shrinking planar area occupied by the bit cell has led to development of exotic techniques for increasing surface area of the capacitor in the bit cell. This has taken the form of developing high aspect ratio posts or deep trenches that provide the surface area required for achieving the desired capacitance. Going forward, there is a need to develop improved high dielectric constant materials for use in the DRAM capacitor that can alleviate the high aspect ratios associated with today's technology, while still maintaining the low leakage current densities required for DRAM operation. At the transistor level, a low leakage pass transistor is required for DRAM implementation, and maintaining this low off state leakage is becoming more challenging as the transistor is scaled to smaller gate lengths. Certain novel transistor designs such as those offered by double gated devices may be useful for implementation in DRAM applications.

As a result of the difficulty associated with scaling DRAM technology, there is a need to develop alternative bit cell technologies that can maintain or increase the read/write/refresh performance of embedded DRAM while maintaining the small bit cell area typical of DRAM. In the case of embedded DRAM, the ease with which a new bit cell can be integrated into a conventional CMOS process flow should be taken into account. In principal, if the technology performs at an adequate level, a dense embedded DRAM technology could serve the same application space as embedded SRAM.

Embedded Flash

Embedding Flash Nonvolatile Memory into CMOS process flows is common place in the microcontroller industry. Embedded Flash faces different scaling challenges to those found in stand alone Flash technology due to the relatively small number of bit cells that are embedded. Because the number of bit cells embedded is limited to values typically less than 16MB, memory array area is dominated by the high voltage peripheral devices required to write and erase the bit cells (typically +/- 9V). As such, for embedded Flash applications, it is critical to find technologies that allow scaling of the voltages required to operate the memory array. Recent work has shown that reducing the operating voltage to values of +/- 6V could reduce typical embedded Flash module sizes by 50% at the 90nm technology node. Opportunities for research in this area include development of nanocrystal based floating gate memories and dielectric charge storage based memories (e.g. SONOS). Development of alternative memory technologies based on resistance changes associated with phase changes in materials systems (e.g. chalcogenide-based memories with resistance changes between the amorphous and crystalline state) or memories based on resistance changes in magnetic tunnel junctions are of interest. Memory technologies that combine the best aspects of SRAM (speed, ease of integration), DRAM (density), and Flash (nonvolatile) are of interest for embedded applications.

Non-volatile Memory

The leading non-volatile memory (NOR flash for code storage and NAND flash for data storage) is facing severe scaling problems stemming from the inability to scale down the tunneling dielectric. Since NAND flash has longer read times, which corresponds to a more relaxed short channel control, it is estimated that it can be scaled to the 32nm node. We anticipate that research in the following area could extend the life of conventional flash devices: tunneling barrier engineering, new material for storage layers, devices based on asymmetric tunneling, nanocrystal storage elements, devices with multi bit spatial resolution,

devices with distinguishable multiple layers, 3-D stacking of conventional non-volatile memory devices in a cost-effective manner.

Research on alternative non-volatile memory devices (phase change memory, MRAM, FeRAM) that shows scaling path beyond 32nm node is also part of this solicitation. New materials offer numerous opportunities to develop new non-volatile memories, which can be used as stand alone or embedded. Complex metal oxides, various perovskites with correlated electronic structure can be used for resistive RAM (RRAM). Ionic conduction or conduction bridges or wires can be exploited in materials like solid electrolytes. NEMS and CNT devices using electromechanical effects can be explored for memory applications.

While we acknowledge a breadth of new materials that can be used for memory applications, we emphasize research that combines material research with device characterization. Different architectural solutions that tend to the future product applications are also sought.

Selection Devices:

The established semiconductor memory technologies (SRAM, DRAM and Charge Storage/Trapping) are expected to approach their scaling limits within the next two or three technology nodes [1]. Several new materials and mechanisms are currently being investigated to develop the next generation memory devices. Most of these approaches rely on non-Si/SiO_2 based memory storage elements [2]. This opens up the possibility of stacking several layers of memory arrays on top of a Si substrate containing the CMOS support circuitry – yielding closer to 100% array efficiency and a factor of **n** increase in memory density, (where n is the number of stacked layers at a given technology node).

Much of the literature on these 3D approaches envisages a cross-point memory architecture - each memory storage layer is sandwiched between an array of orthogonal word lines and bit lines. However, an optimally designed addressable/readable memory array implementation without disturbs and with good sense margin, requires the insertion of an access device for each memory bit. For a 3D memory, this calls for the ability to fabricate stackable access devices – an aspect not quite receiving the requisite attention. The attributes required in this device include, aside from the ability to be stacked, adhesion to underlying and overlying layers, low thermal budget for fabrication and dopant activation (within temperatures tolerated by the memory cell materials e.g.~ 400 C), small geometry, small Ron, low leakage current and, for some applications, bi-directional (bias polarity) operation. *The use of a diode as the access device would allow for the smallest possible cell size - $4F^2$ (where F is the technology node feature size).* It would also result in fewer technology scaling limitations and issues compared to those faced when using a transistor as an access device, where the transistor itself usually becomes the scaling limiter (through gate leakage, SCE etc.).

A thin film diode fabricated by low temperature (< 400 C) deposition of polycrystalline or amorphous films and relying on tunable (via deposition conditions) intrinsic defects for majority carrier generation could meet the stackability and thermal budget constraints outlined above.

A desired research approach would be to develop a thin film diode that meets the defined pass device characteristics, and is easily integrated into a low temperature backend CMOS

flow. The preferred device would have a low, controlled 'breakdown' voltage for forward and reverse fields but would appear asymmetrical in it's I/V characteristic. This diode would be suitable for use where a memory material requires a reverse polarity field to Erase compared to that used to Program the memory. An alternative device would have a low breakdown in one direction only. This second device is limited for use with memory materials that can be both Programmed and Erased with a uni-directional field.

References:

1. H. Goronkin and Y. Yang, 'High-performance Emerging Solid State Memory Technologies,' MRS Bulletin, Nov. 2004, pp. 805-808.
2. G. Muller et al., 'Status and Outlook of Emerging Nonvolatile Memory Technologies,' IEDM Technical Digest, 2004, pp. 567-570.

Legends for Table 1: SRC member companies assigned different priorities to various research topics. The prioritization is based on product roadmaps and understanding the state of memory research in the first half of 2005. SRC member companies also express their opinion about the total level of funding available for particular research.

Priority Legend

High Priority Research Topic	(red)
Med. Priority Research Need	(yellow)
Low Priority Research Topic	

Funding Gap Legend

Large Funding Gap	(red)
Medium Funding Gap	(yellow)
Small Funding Gap	

Table 1: Memory Technology of Prioritized Research Needs and Funding Gaps

	Embedded memory		
SRAM	new memory for SRAM replacement (stability, performance, density, power constraints)		
	Innovative device/circuit solutions to enable SRAM scaling (stability, performance, density, power constraints)		
DRAM	new memory for embedded DRAM replacement (stability, retention, performance, density, power constraints)		
	Innovative device/circuit solutions to enhance DRAM scaling (stability, retention, density, performance, power constraints)		
embedded dense memory	new ideas for highly scalable embedded volatile and non-volatile memories addressing issues related to stability, retention, density, performance, power constraints, endurance		
	Subtopics include: embedded MRAM, spin injection MRAM, embedded PCM, and 3-D FRAM		
	Non-volatile memory		
device	scaling of tunneling dielectric (increasing writing and reading speed, improving data retention)		
	3D memory cell (TFT memory devices, FeRAM capacitors, FinFET)		
	nanocrystal memory (single and multilayer, integrated in ONO layers)		
	trap storage		
	multiple bit and multi level storage		
	high T data retention for flash		
material	new material for tunneling barrier engineering (blocking layer, storage layer, tunneling dielectric, asymmetric tunneling)		
	new material for electrodes		
	new memory material systems		
	Subtopics include: metal gates, complex metal oxides, perovskite/correlated oxide memory, solid electrolyte based memory, molecular memory, polymer memory, new material with memory features, and carbon nanotubes with crossbar material		
integration	nano crossbar memories		
	Selection device		
	process integration (low T (<400°C) activation)		
	architecture (memory cell dependent addressing, minimizing leakage path in cells)		
	TFT transistors		
	back-end diode (high current density, reverse breakdown voltage)		